Oligonucleotide Directed Sequence Specific Recognition and Alkylation of Double Helical DNA by Triple Helix Formation

Thesis by

Thomas J. Povsic

In Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy

California Institute of Technology

Pasadena, California

1992

(Submitted August 6, 1991)

Acknowledgments

I would like to first acknowledge Dr. Peter Dervan. His guidance made this work possible; his standards of excellence challenged me during my stay here, and will continue to do so wherever I go.

The Dervan group has had many members whose friendship I greatly treasured. I thank them for their company and faith in me.

I thank TG, Tim, Pat, Joe and Terry, John Gibboney, Andy, Diann and Betsy, Scott Singleton, the Strobels: Scott, Lynette, Ben, and Sarah, Ramy, and Marc, as well as the many others whom I have overlooked. No one can be a failure with friends such as you.

I thank Dawn. Few people are ever as blessed as I was to have you in my life.

Finally, I dedicate this thesis to my family. Your love is my foundation, and your lives of dedication, service and perseverance are an example to me forever.

Abstracts

Chapter II

Footprinting of Oligonucleotides on Double Helical DNA using MPE•Fe(II), DNase I, and Dimethyl Sulfate

Pyrimidine oligonucleotides can, when equipped with the thymidine-EDTA•Fe(II) analogue (T^*), recognize and subsequently cleave double helical DNA at binding sites >15 base pairs in size. If binding affinities of unmodified oligonucleotides are to be determined under conditions relevant to those *in vivo*, alternate methods of detecting oligonucleotide-directed triple helix formation are required. The footprinting of short (up to 15 base pairs) triple helical regions on restriction fragment size DNA has been undertaken. Techniques for the determination of oligonucleotide binding to double helical DNA using **MPE**•Fe(II), DNase I, and dimethyl sulfate have been developed. **MPE**•Fe(II) allows for the determination of binding site size, and has shown that oligonucleotide binding to DNA is cation concentration, solvent, and oligonucleotide length dependent. DNase I footprinting was conducted under conditions optimal for DNase I activity (10 mM in each Mg^{+2} and Ca^{+2}), demonstrating that oligonucleotide-directed triplexes are capable of interfering with protein activity at the oligonucleotide binding site under physiological conditions, and that divalent cations can stabilize triple helix formation. Footprinting using dimethyl sulfate reveals that a single guanine 3' to the binding site becomes hyperreactive to methylation upon triplex formation. This suggests that the triplex-duplex junction involves a change in DNA conformation which is largely limited to a single base pair.

DMS footprinting reveals that the oligonucleotide CT-15 ($T_5(CT)_5$) does not bind the terminal 2 base pairs of the binding site in plasmid pDMAG10. DMS footprinting can be used to analyze the binding of oligonucleotides to DNA under conditions not amenable to MPE•Fe(II) or DNase I activity, and to assay the kinetics of oligonucleotide binding. DMS and DNase I footprinting techniques were used to assay for the effect of oligonucleotide concentration and base composition on binding affinity.

Chapter III

Oligonucleotide-Directed Triple Helix Formation using Oligonucleotides with Increased Binding Affinities

The specificity offered by the triple helix motif might provide a method for the artificial repression of gene expression and viral diseases. Changes in oligonucleotide structure could be used to control oligonucleotide affinity under *in vivo* conditions, where temporal and spatial intracellular pH (7.0 - 7.4) and ionic strength are strictly regulated and cannot be altered. Substitution at position 5 of pyrimidines alters the hydrophobic driving force, base stacking, and the electronic complementarity of the Hoogsteen base pairing for triple helix formation. Incorporation of 5-substituted pyrimidines offers a method of modulating binding affinity without changing the hydrogen bonding pattern and sequence specificities of pyrimidine oligonucleotides. Replacement of 2'-deoxycytidine with 5-methyl-2'-deoxycytidine increases the oligonucleotide affinity and extends the pH range for binding. Substitution of 5-bromo-2'-deoxyuridine for thymidine increases binding affinity. Oligonucleotides constructed with 2'-deoxyuridine show lower binding affinities. Pyrimidine oligonucleotides constructed from 5-iodo-2'-deoxyuridines and 5-ethynyl-2'-deoxyuridines display increased binding affinities relative to thymidine, but decreased relative to 5-bromouridine containing oligonucleotides. Substitution by ethyl, pentyl, pentynyl, 2-phenyl-ethynyl, or fluoro at the 5 position of 2'-deoxyuridine or bromo at the 5 position of 2'-deoxycytidine residues results in oligonucleotides with decreased binding affinities for double helical DNA.

Chapter IV

Efficient, Base-Specific Alkylation of DNA using *N*-Bromoacetyloligonucleotides

The attachment of a non-specific diffusible cleaving functionality to a DNA binding molecule allows for the elucidation of the structural principles for DNA recognition, a technique termed affinity cleaving. Once these principles have been determined, it becomes possible to design and attach structural domains designed to carry out a desired DNA modification. The development of a thymidine derivative capable of efficient and base specific DNA modification is reported. *N*-bromoacetyloligonucleotides are capable of near quantitative double strand modification of double helical DNA at a single guanine position in a manner which produces ends which are ligatable with compatible ends produced by conventional restriction enzyme digestion. The products thus produced are capable of transforming bacterial cell lines. *N*-bromoacetyloligonucleotides modify double helical DNA with specificities great enough to produce efficient (>90%) chemical cleavage at a single site within a yeast chromosome 340 kbp in size.

The acceleration obtained by tethering a reactive moiety to a DNA binding unit has been estimated. The rate of alkylation of DNA by *N*-bromoacetyloligonucleotides and bromoacetamide has been measured. Comparison of these rates indicates that an effective molarity of 2-3 M is obtained upon tethering the bromoacetyl moiety to an oligonucleotide to effect triple helix mediated DNA alkylation.

The utility of the bromoacetyl moiety as a reporter group is shown in studies concerning the effect of oligonucleotide length on binding affinity and the cooperative interaction between oligonucleotides binding abutting sites is reported.

Table of Contents

List of Figures and Tables

Chapter I

Tables

Figures

Chapter II

Chapter III

Tables

Figures

Chapter I

Introduction to Oligonucleotide-Directed Recognition of Double Helical DNA and Triple Helical DNA Complexes

The recognition of specific DNA sequences, the folding of proteins into functional domains, signal transduction via the recognition of messenger molecules follwed by induced allosteric changes in protein structure, the discrimination between antigens and "self" epitopes by the immune system: these are a few of the fundamental biological processes controlled and directed by the formation of ensembles of individually weak interactions. The organic chemist has just begun to turn his attention in earnest to the study of how weak non-covalent interactions impart to biological systems the specificities and affinities which enable them to be sensitive to small changes in environment and to avidly bind minute quantities of the molecule of interest in the presence of myriads of other, often similar, structures.

The problem of DNA recognition can only be solved once the rules for "non-covalent" chemistry are understood. The tools of molecular biology can be used to analyze large numbers of potential binding sites, allowing the rapid analysis of specificity and affinity. As is often the case, the scientist takes as an example the exquisite molecules constructed by nature. The living organism depends on the ability of molecules to recognize and bind sequences of DNA with high specificity and with binding constants which enable their actions to be felt under physiological conditions. The modern era of molecular biology

was ushered in with the discovery and isolation of molecules which can modify DNA quantitatively at specific sequences.[1,2] Without reagents such as kinases, polymerases, and restriction endonucleases, common procedures such as gene isolation, DNA sequence determination, and DNA recombination and cloning would remain intractable tasks.

The use of restriction enzymes to quantitatively cut DNA with almost absolute specificity at the recognition sequence is the foundation for laboratory DNA manipulation, enabling the study of not only the DNA, but the gene product of interest.[1,2] The effectiveness of restriction enzymes, both as natural defense mechanisms and as tools in the laboratory, lies largely in their selectivity for the correct binding site and ability to effect quantitative reaction at that site. For instance, upon introduction of any incorrect base pair into the EcoR I binding site, the rate of cleavage decreases by five orders of magnitude.[3]

Recent interest in the manipulation and sequencing of genomic DNA, however, has made the sequence specificities of commonly available restriction enzymes (4-8 base pairs) inadequate (Table 1).[4,5] Molecules which recognize 8 base pair sequences cleave statistically once every 32,000 base pairs. Cleavage of human chromosomal DNA (average size = 130,000,000 base pairs) would result in the production of over 4,000 pieces. If manipulation of chromosomal DNA (up to 10^8 base pairs) is to become a reality, the recognition of unique sites > 14 base pairs in size will be necessary (Table 1.1).[4]

The need for "rare cleaving" molecules capable of such specific recognition is reflected in the number of strategies which have been developed to extend the specificity of restriction enzymes to the 8-12 base pair level.[6-14] Molecules with >16 base pair specificities would allow the recognition of single sequences within human genomic size DNA, and might

additionally possess therapeutic significance as agents for specific gene repression or activation. The challenge to build molecules possessing arbitrary DNA sequence specificities and which can effect desired modifications has been answered by chemists, not only because a solution might usher in a new era of rational drug design, but to obtain a conceptual understanding of how non-covalent forces dictate the function of biologically relevant macromolecules.

base pairs	unique sites	base pairs	unique sites
4	136	13	33,554,432
5	512	14	134,225,920
6	2,080	15	536,870,912
7	8,192	16	2,147,516,416
8	32,896	17	8,589,934,592
9	131,072	18	68,719,607,808
10	524,800	19	137,438,953,472
11	2,097,152	20	549,756,338,176
12	8,390,656	21	2,199,023,300,000

Table 1.1: Relation between binding-site size and number of unique sequences.[4] A binding site size of 16 base pairs is sufficient to define unique sequences with DNA the size of human chromosomal DNA.

During the past 15 years, work in this laboratory has centered on elucidation of the structural principles responsible for the sequence specific recognition of DNA by small minor groove binding molecules, as well as major groove binding proteins and oligonucleotides. It is hoped that an understanding of these principles might allow for the design of novel DNA binding molecules with specificities for arbitrary sequences 6 base pairs or more in length.

DNA recognition by minor groove binding molecules. Studies of small minor groove binding molecules have centered on analogues of netropsin and distamycin, natural products which recognize A•T base pairs in the minor groove of DNA.[4,5,15-22] These studies have exemplified the power of synthesis as a means of structural analysis and of creating novel structures with pre-designed functions. For example, an understanding of the structural requirements for DNA recognition by this class of molecules has allowed the design of stereospecific[15] and metallo-regulated[16] derivatives of these natural products. Despite extensive success in the recognition of stretches of dA•dT DNA up to 16 base pairs in length,[23] extension to the recognition of G-C base pairs from the minor groove has been slow.[24]

Protein DNA binding motifs. Our understanding of DNA recognition by proteins has been greatly enhanced by solution NMR,[25,26] and X-ray crystallography data.[27-31] A number of structural motifs used by DNA binding proteins have been identified.[25,26] Searches of protein data banks indicate that each of these motifs (leucine zipper, zinc finger, helix-turn-helix and helix-loop-helix) is used by a family of proteins to recognize different target DNA sequences. The affinity cleaving technique developed in this group has provided a means by which models of protein-DNA structures can be tested in solution.[32-36] Despite these advances, no generalized binding scheme for the recognition of base pairs by "complementary" amino acids has or is likely to emerge.[37] The *de novo* synthesis of DNA binding proteins with arbitrary DNA binding specificities is therefore limited by the complexity inherent in the design and prediction of protein super-secondary structure.[38-40]

Three stranded DNA complexes. The concept of oligonucleotide-directed triple helix formation as a motif for the recognition of double helical DNA was inspired by the observation that polynucleotides form not only

double, but triple helical complexes under proper stoichiometric ratios and salt conditions. The first such complex, poly(U):poly(A):poly(U), was reported over 30 years ago based on UV mixing studies.[41] Rich et al. suggested that the third uridine strand could be accomodated in the major groove of DNA with the formation of Hoogsteen base pairing.[42,43] Triple helical complexes between a variety of DNA and RNAs have subsequently been found. The most common such triple helical complexes are composed of 2 pyrimidine strands and a single purine strand (including U•A-U,[43-45] dT•dA-dT,[46] dT•A-U,[47] dT•A-T,[47] C+G-C,[48-51] dT•A-dT,[47,52] U•dA-U,[48-51] and U•dA-dT,[53]) although purine•purine•pyrimidine complexes (I-A-dT,[47] G-G-C,[48,50] and dI-dC-dI[54]) and purine•purine•purine triplexes (I-A-I,[43] and I-I-I[55]) have also been reported.

It was recognized that within the pyrimidine•purine•pyrimidine family of triple helices, isomorphous base triplets T•AT and C+GC could be formed between any purine-pyrimidine duplex and a sequence specific (in a Hoogsteen sense) third pyrimidine strand (Fig. 1.1). The DNA duplex poly (dT-dC)•poly (dG-dA) was subsequently shown to associate with poly (U-C)[53] or poly (dT-dC)[56] to afford triple helical complexes.

Fiber diffraction studies obtained in the 1970's indicated that triplexes of this type most closely resemble an A' form helix as is commonly formed by double helical RNA under high salt conditions (Fig. 1.2).[57-61] These studies indicated that:[60] (i) the bases are predominantly flat with respect to each other (base tilt ~8.5°) (ii) Hoogsteen base pairing is involved; (iii) that all three strands are composed of sugars in a C-3' endo sugar conformation, with anti-glycosidic bonds; (iv) a 30° rotation per residue around the helical axis; (v) an axial rise of 3.26 Å per residue; and (vi) a pitch height of 39.1 Å.

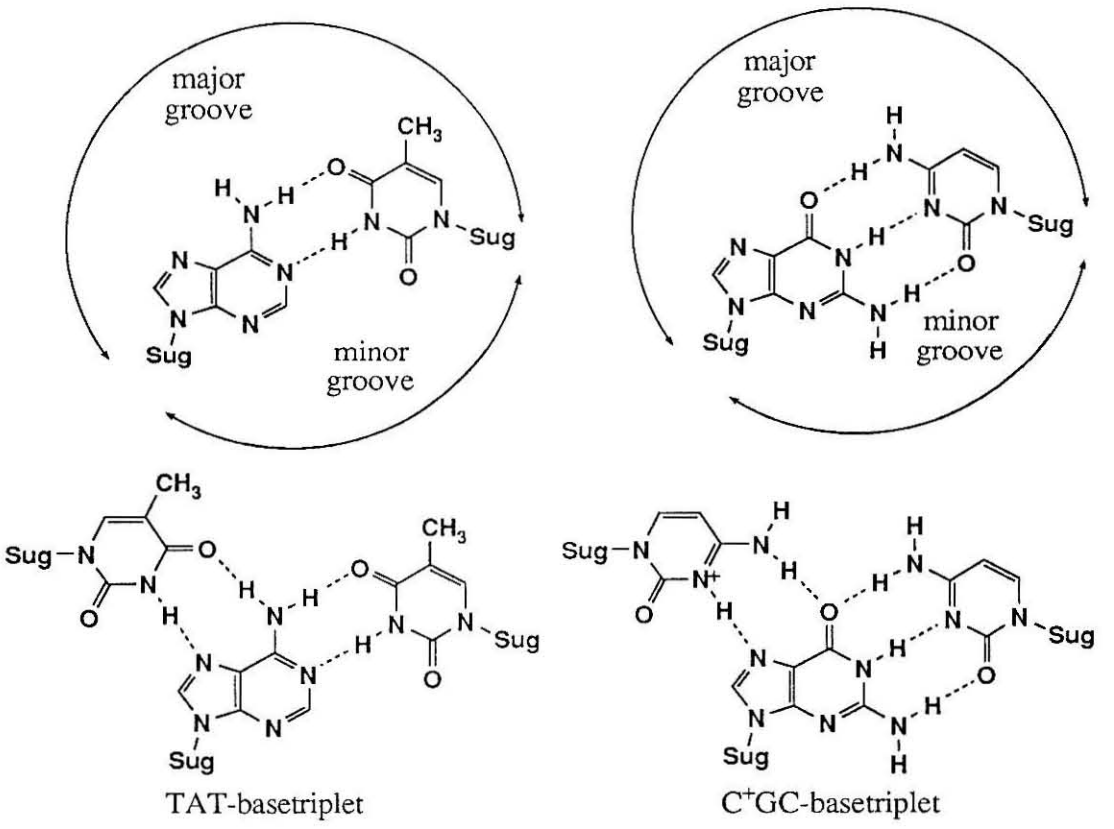
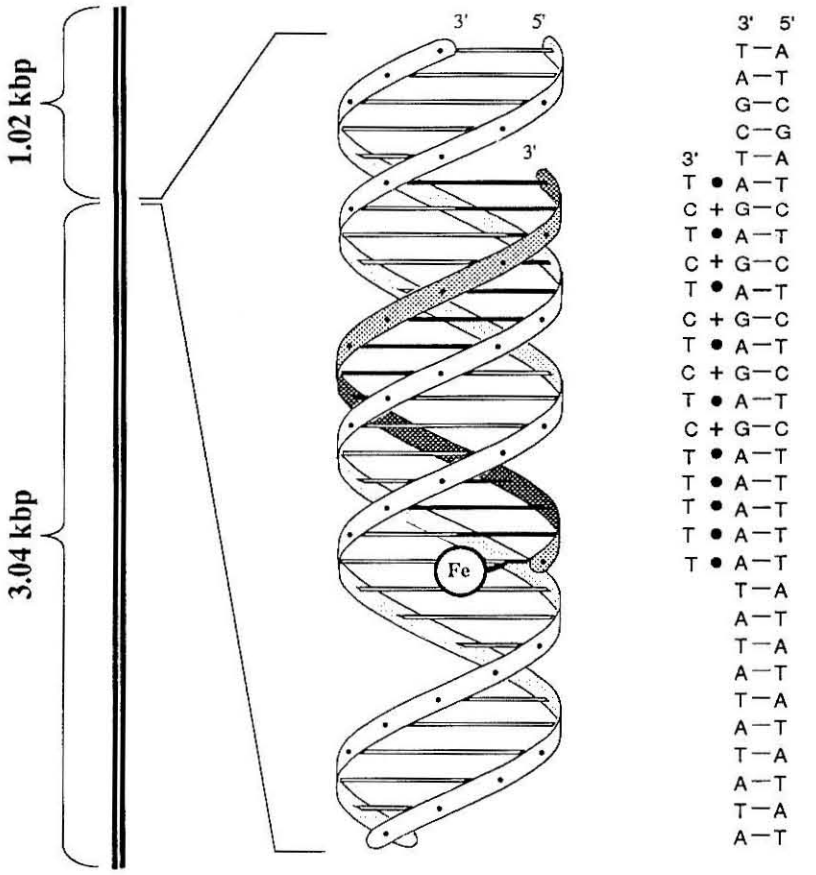


Figure 1.1. (Top) Watson-Crick base pairs in double helical DNA. Major and minor grooves offer different hydrogen bonding patterns for recognition by DNA binding molecules. (Bottom) Isomorphous base triplets formed by incorporation of a third pyrimidine strand in the major groove of double helical DNA parallel to the Watson-Crick purine strand via Hoogsteen base pairing.

Recently, solution NMR structures of intramolecular triplexes[62] at low pH indicate that these parameters are predominantly correct.[63-67] An A-form helix with 3'-endo sugar puckers is suggested; however, Feigon has proposed that the purine strand exists in a different conformation (C-2' endo sugar pucker).[63-66] Triple helical structures have also been studied by a variety of other techniques, including electron microscopy,[68] circular dichroism,[69,70] and Raman spectroscopy.[71,72]

Oligonucleotide-directed recognition of double helical DNA: pyrimidine•pyrimidine-purine motif. The use of oligonucleotides to form a local triple helical complex (oligonucleotide-directed triple helix formation) offers a motif for the recognition of duplex DNA at the > 15 base pair level. This makes the potential specificity of this motif for the recognition of double helical DNA in principle 10^4-10^5 times greater than that attainable by restriction enzymes (6-8 base pair specificity). The recognition of a large numbers of sequences is made possible by the existence of a "triple helical code", much like the genetic code, whereby thymidines recognize A-T base pairs (T•A-T base triplets) and protonated cytosines recognize G-C base pairs (C+G-C base triplets) via Hoogsteen hydrogen bonding.[43,53,73] Furthermore, issues of folding and secondary structure, which complicate the modeling and design of protein domains, are largely avoided, making the modeling and design of new molecules with altered specificities easier and more accurate.

Moser and Dervan were the first to demonstrate that pyrimidine *oligonucleotides* complementary to a purine sequence could bind double helical DNA in a sequence specific manner.[73] The double helical cleavage of plasmid DNA (4.06 kbp) at a single site using oligonucleotides containing a modified thymidine-EDTA•Fe(II) (T*)[74] (Fig. 1.3) was achieved.[73] Using affinity cleaving, Moser and Dervan were able to positively discern the orientation of the third strand as parallel to the purine strand, eliminating D-looping as a mode of binding.[73] The asymmetry in the cleavage pattern confirmed the location of the third strand in the major groove of the duplex.[73] The importance of high salt and neutral to acidic pH as conditions conducive to oligonucleotide-directed triple helix formation was stressed. Oligonucleotide binding was restricted to solutions containing the polycations spermine or $Co(NH_3)_6^{+3}$ at neutral to acidic pH. This is consistent with an



Helix type:	A'
Helix symmetry:	12
Pitch Height:	39.1Å
Axial rise per residue	3.26Å
Rotation per residue	30.0°
Dislocation from helix axis	2.80Å
Base tilt	8.5°
Sugar puckering:	C-3'-endo

Fig. 1.2. Oligonucleotide-directed triple helix formation allows the recognition and cleavage of a 15 base pair site within plasmid DNA as described by Moser and Dervan. A simplified model of the triple helical complex between the Hoogsteen-bound DNA-EDTA•Fe(II) at a single site within 4.06 kbp of plasmid DNA is shown. Specificity is imparted by Hoogsteen hydrogen bonded base triplets T•AT and C+GC.

accumulation of negative charge upon accomodation of a third polyanion in the major groove and the requirement for protonation at N-3 of cytosine for the formation of Hoogsteen hydrogen bonds. Binding of a pyrimidine oligonucleotide to a purine-pyrimidine tract in plasmid DNA was also reported to occur at low pH, as assayed by gel mobility changes.[75] Improving the stability of oligonucleotide-directed triplexes may be required before oligonucleotides can be expected to form stable complexes under *in vivo* conditions.

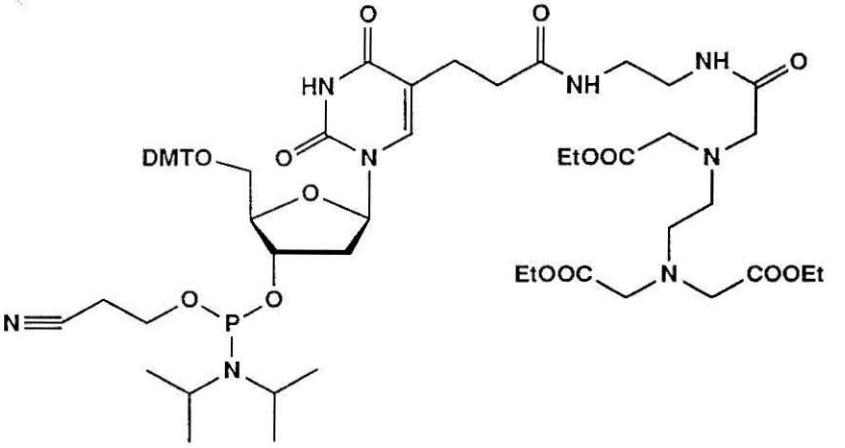


Figure 1.3. DMT-protected T*-phosphoramidite-triethylester[74] compatible with automated synthesis techniques for the construction of oligonucleotide-EDTA for cleavage of double helical DNA via triple helix formation.

Strobel, Moser and Dervan extended this study to the cleavage of large DNA. Cleavage at a single site within the genome of the phage lambda (48,500 base pairs) using thymidine-EDTA•Fe(II) containing oligonucleotides was achieved in 25% yield;[76] however, cleavage of an engineered site within chromosome III (340 kbp) of the *Saccharomyces cerevisiae* genome could

only be accomplished in low yield (5%), and cleavage at secondary sites was observed.[77] The presence of secondary cleavage sites is thought to arise due to the known over-representation of purine sequences in genomic DNA.[78,79] This study indicates that there are limits to the specificity and cleavage efficiency obtainable on large DNA using oligonucleotides equipped with non-specific DNA modification moiety. Use of a specific cleavage moiety may result in the production of DNA cleaving molecules with increased specificities.

Work by Maher, Wold and Dervan has shown the ability of local triplexes to disrupt protein function.[80] These studies indicate that pre-formed triple helices can inhibit methylases, restriction endonucleases, and transcription factors when the triplex and the protein binding sites overlap.[80] Inhibition of restriction endonuclease activity (i. e. restriction endonuclease footprinting) can be used as a method of assaying for triple helix formation.[80,81] This work demonstrated the feasibility of using triple helix formation to direct restriction endonuclease cleavage of DNA via an "Achilles heel" strategy.[13,14] In this strategy, triple helix formation is used to inhibit methylation at a methylase/triplex site. Complete methylation of intact DNA renders all methylase/restriction enzyme sites not protected by the oligonucleotide resistant to cleavage by the endonuclease. Disruption of the triplex exposes the purine/methylase site for cleavage by a complementary restriction endonuclease.

Strobel and Dervan have shown the utility of the triple helix mediated "Achilles heel" strategy by using oligonucleotides and EcoR I methylase/restriction enzymes to effect the enzymatic cleavage of a yeast chromosome at a single engineered purine/methylase/restriction site.[82] In this study, near quantitative cleavage of a 340 kbp yeast chromosome (94%)

was achieved at a single site. No cleavage of any other chromosome was detected. This work demonstrates the feasibility of using triple helix strategies to impart >16 base pair specificity to restriction enzymes to realize the cleavage of single sites within chromosomal DNA. Although a triplex/methylase/restriction site was specifically engineered into chromosomal DNA for the purpose of this study, the generalizability of the triple helix motif allows the recognition and cleavage of sequences present in genetic markers of interest.

A method for the screening of large sequences of DNA for triple helix methylation-restriction sites using oligonucleotides degenerate in the nucleosides 5-bromouridine and 5-methylcytidine has made the determination of endogenous sites cleavable by the Achilles heel strategy possible.[83] This technique has been used by Strobel and Dervan to determine the sequence of a purine/restriction enzyme site within an unsequenced genetic marker near the Huntington's disease gene. Specific cleavage of a human chromosomal DNA at this site has been achieved.[83]

The specificity of triple helix formation offers the tantalizing possibility of using these molecules as specific gene repressors or activators *in vivo*. Recent work indicates that triplexes can lower the efficiency of transcription *in vitro* when formed near the site of initiation, both when the triplex does or does not block specific transcription factors.[84] This suggests that triple helix formation can effect transcription both through structural changes imparted to the DNA, and by inhibiting the binding of required transcriptional promoters.[84]

Characterization of the thermodynamics underlying triple helix formation has been undertaken using UV melting,[85] NMR,[86] CD,[87,88] and differential scanning calorimetry methods.[89] Estimates for the enthalpic

contributions for triplex formation indicate that, at pH 6.5 in 200 mM NaCl, the ΔH^0 for triplex formation is approximately 2.0 kcal/mol base triplet (the ΔH^0 for duplex formation was determined at 6.1 kcal/mol bp),[89] but that triplex formation is entropically much less unfavorable than is duplex formation. Melting of the triple helix is entropically disfavored by approximately 6.5 cal/mol•°K (the duplex transition is approximately 18.5 cal/mol °K).[89] Triplex formation at 25 °C, close to neutral pH (6.5), and 200 mM NaCl is favored by 1.3 kcal/mol for a 15 base pair oligonucleotide.[89] These studies are conducted using oligonucleotide Watson-Crick duplexes. Binding of oligonucleotides to large DNA may differ in several aspects, not least of which is the production of a triplex to duplex transition within the target DNA. These values are expected to be sequence dependent. The study of a large number of sequences will be required before general parameters for the *a priori* prediction of triple helix stability will be possible.

Oligonucleotide-directed recognition of double helical DNA: purine•purine-pyrimidine motif. A less well understood motif (purine•purine-pyrimidine) allows for recognition of G-C base pairs with guanines, and A-T base pairs with either thymine or adenine bases (Fig. 1.4).[90-95]. Binding of the third purine strand is anti-parallel to the Watson-Crick purine strand, again eliminated D-looping as a mode of recognition.[90] Preliminary data indicate that triplexes of this type are driven by the formation of stable G•G-C triplets.[90,92,95] In the study of Cooney et al., a 27 base long oligonucleotide parallel to the purine strand (incorrect orientation) was able to bind a target sequence due to the high G content (20 out of 28 positions were G).[95]

Since protonation is not required for formation of G•G-C or A•A-T triplets, these triplexes are much less pH dependent than are

pyrimidine•purine-pyrimidine triplexes.[90] An intramolecular triple helix of this family has been reported.[91] Studies of this motif are in their infancy, and little has been accomplished to extend the recognition alphabet of this motif to all four possible base pairs (A-T, T-A, G-C, C-G) of the double helix. The two triplex families are complementary: pyrimidine oligonucleotides bind with higher affinities to A-T rich sequences; purine G-rich oligonucleotides can be used to recognize predominantly G-C containing sequences.

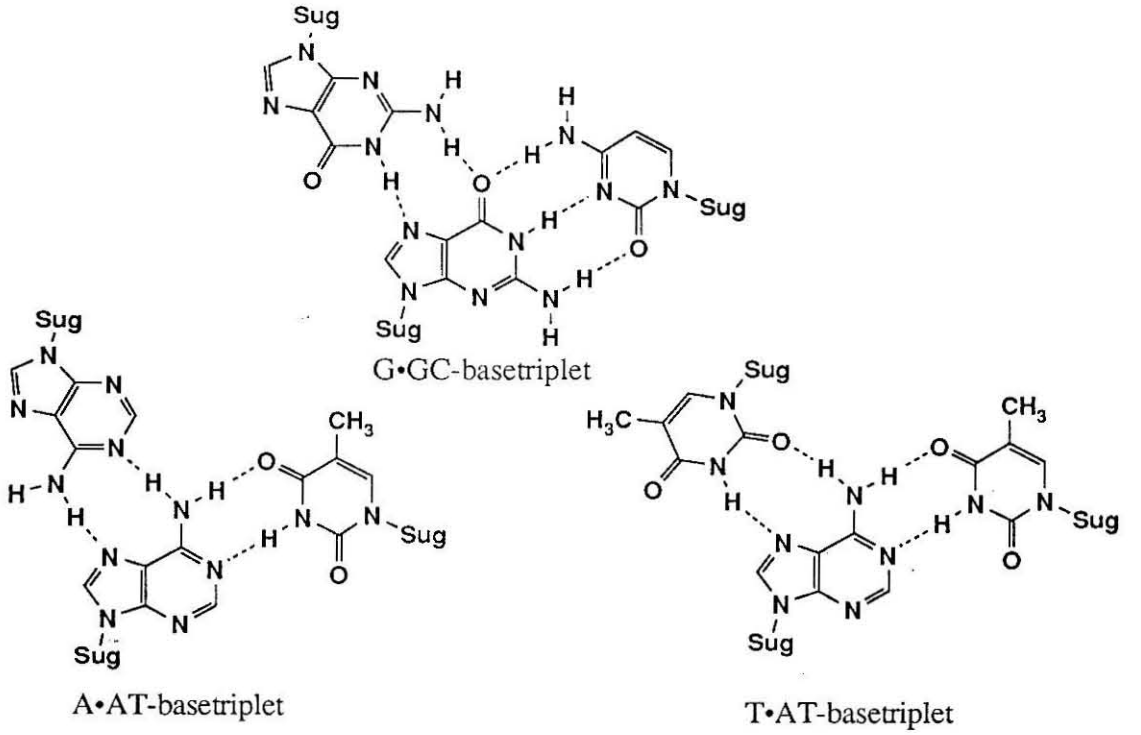


Fig. 1.4. Models proposed by Beal and Dervan[90] for G•GC, A•AT, and T•AT base triplets within a purine•purine-pyrimidine triple helix motif where the third strand is antiparallel to the purine Watson-Crick strand and the bases are in the anti conformation.

General triple helical solutions to DNA recognition. The above studies clearly demonstrate the power of oligonucleotide directed triple helix

formation as a motif for the recognition of specific sequences of double helical DNA; however, these studies have been limited to the use of purine-pyrimidine sites. Extension of triple helix formation as a motif for the recognition of arbitrary sequences containing all four base pairs has focussed on five strategies: (i) the search for novel natural base triplets, such as G•TA;[96] (ii) the design and synthesis of oligonucleotides which can cross-over the major groove, making the recognition of purine-pyrimidine type sequences possible;[97] (iii) the use of "null" bases to skip over the "unreadable" base pairs C-G and T-A;[98] (iv) the study of novel triple helical motifs which are[99,100] or are not[90] protein mediated; and (v) the design of new heterocycles which might allow the recognition of all four base pairs within a triple helical motif.[101]

Natural roles for triple helices. Since the discovery of triple helices, roles for these structures *in vivo* have been postulated.[102-104] Anti-bodies raised against triple helical DNA are reported to bind eucaryotic chromosomes *in vivo*.[105]

Purine runs have been located at the 5' ends of many genes.[94,95,106-128] Long purine runs are found to be statistically over-represented in mammalian DNAs.[78,79,117,129] Many of these runs display differential reactivity to DNA modification agents and nucleases. Several structures were proposed to account for the changes in reactivity observed. A triple helical structure was first proposed by Christophe, Vassart and co-workers to explain the sensitivity of these sequences under superhelical stress to S1 nuclease.[108]

Kamenetskii and co-workers studied purine runs more extensively, and named the intramolecular triple helical complex formed by extended purine sequences H-form DNA.[130-136] These workers postulated that half of the

pyrimidine strand folds back to Hoogsteen hydrogen bond with one half of the purine sequence, leaving half of the purine run in a single stranded form.

The reactivity of such sequences to chemical DNA modification reagents has been extensively studied, and models concerning the exact nature of the loops of the DNA have been proposed.[137-139] Modelling indicates that H-form DNA preferentially forms asymmetrically such that the 3' half of the polypyrimidine strand loops back to become Hoogsteen hydrogen bonded.[137,138] This conformer results in the release of maximal superhelical tension.[137] Chemical reactivity data as a function of superhelical stress support this idea.[137] This structure generally best explains the reactivity of these sequences *in vivo* to nucleases and *in vitro* to chemical modification reagents.[106,108,114-128,139-142]

Alternate "H-form" DNAs which result in the formation of purine-purine-pyrimidine triplexes and a "free" pyrimidine strand have also been postulated.[92] Some sequences are reported to form either of these types of "H-form" DNAs, the structure formed being dependent on the concentration of Mg^{+2}.[92] The homopyrimidine-homopurine sequence $d(CT/GA)_{22}$ was recently shown to undergo transition to a *H-form DNA, different from either of the triplexes above.[143] This transition is promoted by zinc ions, and a structure has not yet been assigned.[143]

These motifs have been suggested to play roles in the regulation of downstream genes *in vivo*.[94,95,108,109,144] Oligonucleotides complementary in a Hoogsteen sense to the purine site prevent formation of H-form DNA,[75] offering a possible mechanism for oligonucleotide control of gene expression.[94,95] Hogan *et al.* have shown that the sensitivity of the sequence

5' to the *c-myc* gene to S1 nuclease increases in the presence of a RNA factor. *In vitro*, a G-rich purine oligonucleotide forms a triplex at this site.[95] The site is rich in G-C base pairs, and only a purine rich (i. e. G rich) oligonucleotide is capable of altering gene activity.[95] Triple helical regions have also been implicated in a variety of other biological functions, including viral integration[110] differential methylation of sequences,[127] and DNA inversions.[107]

Other triplex work. Not surprisingly, the number of workers in the triple helix field has grown substantially within the last few years. A strategy for the high yield site-specific cleavage of double helical DNA via a triple helix strategy, as demonstrated by Pei and Schultz, is the use of semi-synthetic nuclease-oligonucleotide conjugates.[145] Although this strategy offers higher yields than those obtained with oligonucleotide-EDTA•Fe(II), no additional specificity is obtained using a non-specific nuclease as a cleaving functionality.

Helene and co-workers have equipped triple helix forming oligonucleotides with intercalators,[146,147] DNA cross-linking reagents,[148,149] and DNA cleaving reagents (1,10 phenanthroline•Cu(II)).[150] Triplex formation between several types of modified oligonucleotides and native DNA has also been investigated.[151-155] Triplex formation with phosphorothioates, phosphorodithioates, methyl phosphonates has been reported.[154,155] Haner has shown that pyrimidine oligonucleotides made with 2'-OMe nucleosides bind with higher affinity than their DNA analogues.[156] Helene and co-workers have constructed nuclease resistant oligonucleotides from [α]-deoxynucleotides.[151,152] One of the interesting properties of these molecules is their binding orientation, which is the opposite of that normally observed in oligonucleotide•duplex DNA complexes. Modified

oligonucleotides have properties such as nuclease resistance and increased permeability through phospholipid membranes which may make them attractive for use *in vivo* and as possible therapeutic agents.

Description of work. This thesis describes experiments designed to improve our understanding of oligonucleotide-directed triple helix formation, to improve the stability of pyrimidine•pyrimidine-purine triplexes, and to equip triple helix forming oligonucleotides with reactive moieties capable of efficient and base specific DNA modification.

Chapter II describes the development of footprinting techniques which can be used to detect triple helix formation. Affinity cleaving requires the synthesis and incorporation of modified nucleosides into the oligonucleotide. Incorporation of such nucleosides may change oligonucleotide binding characteristics, and limits binding assays to those conditions compatible with EDTA chemistry. The development of conditions for footprinting using several different reagents (MPE•Fe(II), DNase I, dimethyl sulfate) allows for the analysis of triple helix formation under a variety of salt conditions, over a large range of pH values and a variety of time intervals.

Chapter III focuses on the incorporation of modified nucleosides to afford oligonucleotides with increased binding affinities. Oligonucleotides which are resistant to denaturation at slightly alkaline pH might be important if they are to be used in environments, such as those within living systems, where cation concentration and pH are strictly regulated and cannot be altered.[157-160] Substitution at the 5-position of pyrimidines affects the hydrophobic driving force, base stacking, and electronic complementarity of these nucleosides in triple helix formation.[161,162] Oligonucleotides containing

5-methylcytosine and 5-bromouracil bases display increased binding affinities over an extended pH range. Modified oligonucleotides have been repeatedly used in applications where high binding affinities are required. Preliminary communication of these results has appeared.[163]

Chapter IV describes the development of a nucleoside derivative capable of highly efficient and base specific DNA modification.[164] Incorporation and derivatization of this nucleoside affords *N*-bromoacetyloligonucleotides capable of efficient chemical modification and cleavage of chromosomal DNA at a single nucleotide position. Due to the high cleavage efficiency possible with this moiety, it is an ideal reporter group to assay for small changes in binding affinity. Studies concerning the ability of *N*-bromoacetyloligonucleotides to bind to abutting sites in a cooperative manner and to inhibit transcription downstream from the initiation site will be reported.

References

(1) Modrich, P. *Crit. Rev. Biochem.* **1982**, *13*, 287-323.

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

(3) Lesser, D. R.; Kurpiewski, M. R.; Jen-Jacobson, L. *Science* **1990**, *250*, 776-786.

(4) Dervan, P. B. *Science* **1986**, *232*, 464-471.

(5) Dervan, P. B. In *Nucleic Acids and Molecular Biology*; F. Eckstein and D. M. J. Lilley, Ed.; Springer-Verlag: Heidelberg, 1988; Vol. 2; pp 49-64.

(6) Barlow, D. P.; Lehrach, H. *Trends Genet.* **1987**, *3*, 167-171.

(7) McClelland, M.; Kessler, L. G.; Bittner, M. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 983-987.

(8) McClelland, M.; Nelson, M.; Cantor, C. R. *Nucl. Acids Res.* **1985**, *13*, 7171-7182.

(9) Patel, Y.; Van Cott, E.; Wilson, G. G.; McClelland, M. *Nucl. Acids Res.* **1990**, *18*, 1603-1607.

(10) Weil, M. D.; McClelland, M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 51-55.

(11) Perlman, P. S.; Butow, R. A. *Science* **1989**, *246*, 1106-1109.

(12) Wenzlau, J. M.; Saldanka, R. J.; Butow, R. A.; Perlman, P. S. *Cell* **1989**, *56*, 421-430.

(13) Koob, M.; Grimes, E.; Szybalski, W. *Science* **1988**, *241*, 1084-1086.

(14) Koob, M.; Szybalksi, W. *Science* **1990**, *250*, 271-273.

(15) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1986**, *108*, 5008-5009.

(16) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 6840-6842.

(17) Youngquist, R. S.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565-2569.

(18) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528-5529.

(19) Schultz, P. G.; Dervan, P. B. *J. Biomolec. Struct. Dyn.* **1984**, *1*, 1133-1147.

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

(21) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861-6863.

(22) Schultz, P. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6834-6837.

(23) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 7564-7566.

(24) Wade, W. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 1574-1575.

(25) Struhl, K. *Trends Biochem. Res.* **1989**, *14*, 137-140.

(26) Brennan, R. G.; Matthews, B. W. *Trends Biochem. Res.* **1989**, *14*, 286-290.

(27) Aggarwal, A. K.; Rodgers, D. W.; Drottar, M.; Ptashne, M.; Harrison, S. C. *Science* **1988**, *242*, 899-907.

(28) Anderson, J. E.; Ptashne, M.; Harrison, S. C. *Nature* **1985**, *316*, 596-601.

(29) Anderson, J. E.; Ptashne, M.; Harrison, S. C. *Nature* **1987**, *326*, 846-852.

(30) Jordan, S. R.; Pabo, C. O. *Science* **1988**, *242*, 893-899.

(31) Otwinowski, Z.; Schevits, R. W.; Zhang, R.-G.; Lawson, C. L.; Joachimiak, A.; Marmorstein, R. Q.; Luisi, B. F.; Sigler, P. B. *Nature* **1988**, *335*, 321-329.

(32) Sluka, J. P.; Bruist, M.; Horvath, S. J.; Simon, M. I.; Dervan, P. B. *Science* **1987**, *238*, 1129-1132.

(33) Sluka, J. P.; Griffin, J. H.; Mack, J. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 6369-6374.

(34) Mack, D. P.; Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7572-7574.

(35) Graham, K. S.; Dervan, P. B. *J. Biol. Chem.* **1990**, *265*, 16534-16540.

(36) Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847-850.

(37) Caruthers, M. H. *Acc. Chem. Res.* **1980**, *13*, 155-160.

(38) Degrado, W. F. *Adv. Prot. Chem.* **1988**, *39*, 51-125.

(39) King, J. *Chem. Eng. News* **1989**, *April 10, 1989*, 32-54.

(40) Dill, K. A. *Biochemistry* **1990**, *29*, 7133-7155.

(41) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

(42) Hoogsteen, K. *Acata Cryst.* **1959**, *12*, 822-823.

(43) Rich, A. *Nature* **1958**, *181*, 521-525.

(44) Blake, R. D.; Massoulie, J.; Fresco, J. R. *J. Mol. Biol.* **1967**, *30*, 291-308.

(45) Massoulie, J. *Eur. J. Biochem.* **1968**, *3*, 439-447.

(46) Cassani, G. R.; Bollum, F. J. *Biochemistry* **1969**, *8*, 3928-3936.

(47) Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1960**, *46*, 1044-1053.

(48) Lipsett, M. N. *Biochem. Biophys. Res. Comm.* **1963**, *11*, 224-228.

(49) Howard, F. B.; Frazier, J.; Lipsett, M. N.; Miles, H. T. *Biochem. Biophys. Res. Comm.* **1964**, *17*, 93-102.

(50) Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256-1260.

(51) Thiele, D.; Guschlbauer, W. *Biopolymers* **1971**, *10*, 143-157.

(52) Riley, M.; Maling, B.; Chamberlin, M. J. *J. Mol. Biol.* **1966**, *20*, 359-389.

(53) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63-80.

(54) Inman, R. B. *J. Mol. Biol.* **1964**, *10*, 137-146.

(55) Rich, A. *Biochim. Biophys. Acta* **1958**, *29*, 502-509.

(56) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucl. Acids Res.* **1979**, *6*, 3073-3091.

(57) Arnott, S.; Bond, P. J. *Nat. New Biol.* **1973**, *244*, 99-101.

(58) Arnott, S.; Hukins, D. W. L.; Dover, S. D.; Fuller, W.; Hodgson, A. R. *J. Mol. Biol.* **1973**, *81*, 107-122.

(59) Arnott, S.; Selsing, E.; Smith, P. J. C. *Nucl. Acids Res.* **1976**, *3*, 2459-2470.

(60) Arnott, S.; Bond, P. J.; Selsing, E.; Smith, P. J. C. *Nucl. Acids Res.* **1976**, *3*, 2459-2470.

(61) Hattori, M.; Frazier, J.; Miles, H. T. *Biopolymers* **1976**, *15*, 523-531.

(62) Haner, R.; Dervan, P. B. *Biochemistry* **1990**, *29*, 9761-9765.

(63) Rajagopal, P.; Feigon, J. *Nature* **1989**, *239*, 637-640.

(64) Sklenar, V.; Feigon, J. *Nature* **1990**, *345*, 836-839.

(65) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859-7870.

(66) de los Santos, C.; Rosen, M.; Patel, D. *Biochemistry* **1989**, *28*, 7282-7289.

(67) Umemoto, K.; Sarma, M. H.; Gupta, G.; Luo, J.; Sarma, R. H. *J. Am. Chem. Soc.* **1990**, *112*, 4539-4545.

(68) Stokrova, J.; Vojtiskova, M.; Palecek, E. *J. Biomolec. Struct. Dyn.* **1989**, *6*, 891-898.

(69) Steely, H. T., Jr.; Gray, D. M.; Ratliff, R. L. *Nucl. Acids Res.* **1986**, *14*, 10071-10091.

(70) Antao, V. P.; Gray, C. W.; Gray, D. M.; Ratliff, R. L. *Nucl. Acids Res.* **1986**, *14*, 10091-10113.

(71) Benevides, J. M.; Wang, A. H.-J.; Rich, A.; Kyogoku, Y.; van der Marel, G. A.; van Boom, J. H.; Thomas, G. J., Jr. *Biochemistry* **1986**, *25*, 41-50.

(72) O'Connor, T.; Bina, M. *J. Biomolec. Struct. Dyn.* **1984**, *2*, 615-625.

(73) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.

(74) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968-972.

(75) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D.; Cantor, C. R. *Nucl. Acids Res.* **1988**, *16*, 2165-2179.

(76) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927-7929.

(77) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73-75.

(78) Beasty, A. M.; Behe, M. J. *Nucl. Acids Res.* **1988**, *16*, 1517-1528.

(79) Behe, M. J. *Biochemistry* **1987**, *26*, 7870-7875.

(80) Maher, L. J., III; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-730.

(81) Maher, L. J., III; Dervan, P. B.; Wold, B. *Biochemistry* **1990**, *29*, 8820-8826.

(82) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172-174.

(83) Strobel, S. A. Ph. D. Thesis, California Institute of Technology, 1991.

(84) Maher, L. J., III; Dervan, P. B.; Wold, B. J. *Biochemistry* submitted.

(85) Shea, R. G.; Ng, P.; Bischofberger, N. *Nucl. Acids Res.* **1990**, *18*, 4859-4866.

(86) Mooren, M. M. W.; Pulleyblank, D. E.; Wijmenga, S. S.; Blommers, M. J. J.; Hibers, C. W. *Nucl. Acids Res.* **1990**, *18*, 6523-6529.

(87) Xodo, L. E.; Manzini, G.; Quadrifoglio, F. *Nucl. Acids Res.* **1990**, *18*, 3557-3564.

(88) Manzini, G.; Xodo, L. E.; Gasparotto, D.; Quadrifoglio, F.; van der Marel, G. A.; van Boom, J. H. *J. Mol. Biol.* **1990**, *213*, 833-843.

(89) Plum, G. E.; Park, Y.-W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Biochemistry* **1990**, *87*, 9436-9440.

(90) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.

(91) Chen, F.-M. *Biochemistry* **1991**, *30*, 4472-4479.

(92) Kohwi, Y.; kohwi-Shigematsu, T. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3781-3785.

(93) Broitman, S. L.; Im, D. D.; Fresco, J. R. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 5120-5124.

(94) Boles, T. C.; Hogan, M. E. *Biochemistry* **1987**, *26*, 367-376.

(95) Cooney, M.; Czernuszewica, G.; Postel, E. H.; Flint, S. J.; Hogan, M. E. *Science* **1988**, *241*, 456-459.

(96) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-970.

(97) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435-2437.

(98) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* submitted.

(99) Rao, B. J.; Dutreix, M.; Radding, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2984-2988.

(100) Hsieh, P.; Camerini-Otero, C. S.; Camerini-Otero, R. D. *Genes and Development* **1990**, *4*, 1951-1963.

(101) Griffin, L. C. Ph. D. Thesis, California Institute of Technology, 1990.

(102) Minton, K. W. *J. Exp. Pathology* **1985**, *2*, 135-148.

(103) Hoffman, S.; Witowski, W. *Z. Chem.* **1976**, *16*, 442-446.

(104) Hopkins, R. C. *Comments Mol. Cell. Biophys.* **1984**, *2*, 153-178.

(105) Lee, J. S.; Burkholder, G. D.; Latimer, L. J. P.; Haug, B. L.; Braun, R. P. *Nucl. Acids Res.* **1987**, *15*, 1047-1061.

(106) Wells, R. D.; Collier, D. A.; Hanvey, J. C.; Shimizu, M.; Wohlrab, F. *FASEB J.* **1988**, *2*, 2939-2949.

(107) Wohlrab, F.; McLean, M. J.; Wells, R. D. *J. Biol. Chem.* **1987**, *262*, 6407-6416.

(108) Christophe, D.; Cabrer, B.; Bacolla, A.; Targovnik, H.; Pohl, V.; Vassart, G. *Nucl. Acids Res.* **1985**, *13*, 5127-5144.

(109) Kinniburgh, A. J. *Nucl. Acids Res.* **1989**, *17*, 7771-7778.

(110) Shih, C.; Burke, K.; Chou, M.-J.; Zeldis, J. B.; Yang, C.-S.; Lee, C.-S.; Isselbacher, K. J.; Wands, J. R.; Goodman, H. M. *J. Virology* **1987**, *61*, 3491-3498.

(111) Fowler, R. F.; Skinner, D. M. *J. Biol. Chem.* **1986**, *261*, 8994-9001.

(112) Cantor, C. R.; Efstratiadis, A. *Nucl. Acids Res.* **1984**, *12*, 8059-8072.

(113) Lapidot, A.; Baran, N.; Manor, H. *Nucl. Acids Res.* **1989**, *17*, 883-900.

(114) Evans, T.; Efstratiadis, A. *J. Biol. Chem.* **1986**, *261*, 14771-14780.

(115) Nickol, J. M.; Felsenfeld, G. *Cell* **1983**, *35*, 467-477.

(116) Htun, H.; Lund, E.; Dahlberg, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 7288-7292.

(117) Hentschel, C. C. *Nature* **1982**, *295*, 714-716.

(118) Finer, M. H.; Fodor, E. J. B.; Boedtker, H.; Doty, P. *Proc. Natl. Acad. Sci. U.S.A.* **1984**, *81*, 1659-1663.

(119) Collier, D. A.; Grifffin, J. A.; Wells, R. D. *J. Biol. Chem.* **1988**, *263*, 7397-7405.

(120) Iacono-Connors, L.; Kowalski, D. *Nucl. Acids Res.* **1986**, *14*, 8949-8962.

(121) Larsen, A.; Weintraub, H. *Cell* **1982**, *29*, 609-622.

(122) McCarthy, J. G.; Heywood, S. M. *Nucl. Acids Res.* **1987**, *15*, 8069-8085.

(123) Pulleyblank, D. E.; Haniford, D. B.; Morgan, A. R. *Cell* **1985**, *42*, 271-280.

(124) Ruiz-Carillo, A. *Nucl. Acids Res.* **1984**, *12*, 6473-6492.

(125) Schon, E.; Evans, T.; Welsh, J.; Efstratiadis, A. *Cell* **1983**, *35*, 837-848.

(126) Yu, Y.-T.; Manley, J. L. *Cell* **1986**, *45*, 743-751.

(127) Parniewski, P.; Kwinkowski, M.; Wilk, A.; Klysik, J. *Nucl. Acids Res.* **1990**, *18*, 605-611.

(128) Davis, T. L.; Firulli, A. B.; Kinniburgh, A. J. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9682-9686.

(129) Manor, H.; Sridhara Rao, B.; Martin, R. G. *J. Mol. Evol.* **1988**, *27*, 96-101.

(130) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomolec. Struct. Dyn.* **1985**, *3*, 327-338.

(131) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomolec. Struct. Dyn.* **1986**, *3*, 667-669.

(132) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomolec. Struct. Dyn.* **1987**, *5*, 275-282.

(133) Lyamichev, V. I.; Mirkin, S. M.; Danilevskaya, O. N.; Voloshin, O. N.; Balaskaya, S. V.; Dobrynin, V. N.; Filippov, S. A.; Frank-Kamentskii, M. D. *Nature* **1989**, *339*, 634-637.

(134) Mirkin, S. M.; Lyamichev, V. I.; Drushlyak, K. N.; Dobrynin, V. N.; Filippov, S. A.; Frank-Kamenetskii, M. D. *Nature* **1987**, *330*, 495-497.

(135) Lyamichev, V. I.; Mirkin, S. M.; Kumarev, V. P.; Baranova, L. V.; Vologodskii, A. V.; Frank-Kamenetskii, M. D. *Nucl. Acids Res.* **1989**, *17*, 9417-9423.

(136) Belotserkovskii, B. P.; Veselkov, A. G.; Filippov, S. A.; Dobrynin, V. N.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *Nucl. Acids Res.* **1991**, *18*, 6621-6624.

(137) Htun, H.; Dahlberg, J. E. *Science* **1989**, *243*, 1571-1576.

(138) Htun, H.; Dahlberg, J. E. *Science* **1988**, *241*, 1791-1796.

(139) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6292-6296.

(140) Hanvey, J. C.; Klysik, J.; Wells, R. D. *J. Biol. Chem.* **1988**, *263*, 7386-7396.

(141) Polayes, D. A.; Rice, P. W.; Dahlberg, J. E. *J. Bacteriology* **1988**, *170*, 2083-2088.

(142) Glover, J. N. M.; Pulleyblank, D. E. *J. Mol. Biol.* **1990**, *215*, 653-663.

(143) Bernues, J.; Beltran, R.; Casasnovas, J. M.; Azorin, F. *Nucl. Acids Res.* **1990**, *18*, 4067-4073.

(144) Glaser, R. L.; Thomas, G. H.; Siegfried, E.; Elgin, S. C. R.; Lis, J. T. *J. Mol. Biol.* **1990**, *211*, 751-761.

(145) Pei, D.; Corey, D. R.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9858-9862.

(146) Sun, J. S.; Francois, J. C.; Montenay-Garestier, T.; Saison-Behmoaras, T.; Roig, V.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9198-9202.

(147) Collier, D. A.; Thuong, N. T.; Helene, C. *J. Am. Chem. Soc.* **1991**, *113*, 1457-1458.

(148) Praseuth, D.; Perrouault, L.; Le Doan, T.; Chassignol, M.; Thuong, N.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1349-1353.

(149) Perrouault, L.; Asseline, U.; Rivalle, C.; Thuong, N. T.; Bisagni, E.; Giovannangeli, C.; Le Doan, T.; Helene, C. *Nature* **1990**, *344*, 358-360.

(150) Francois, J.-C.; Saison-Behmoaras, T.; Chassignol, M.; Thuong, N. T.; Helene, C. *J. Biol. Chem.* **1989**, *264*, 5891-5898.

(151) Praseuth, D.; Chassignol, M.; Takasugi, M.; Le Doan, T.; Thuong, N. T.; Helene, C. *J. Mol. Biol.* **1987**, *196*, 939-942.

(152) Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J. L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucl. Acids. Res.* **1987**, *15*, 7749-7760.

(153) Ono, A.; Ts'o, P. O. P.; Kan, L.-S. *J. Am. Chem. Soc.* **1991**, *113*, 4032-4033.

(154) Callahan, D. E.; Trapane, T. L.; Miller, P. S.; Ts'o, P. O. P.; Kan, L.-S. *Biochemistry* **1991**, *30*, 1650-1655.

(155) Latimer, L. J. P.; Hampel, K.; Lee, J. S. *Nucl. Acids Res.* **1989**, *17*, 1549-1561.

(156) Haner, R.; Dervan, P. B. *J. Am. Chem. Soc.* in preparation.

(157) Madshus, I. H. *Biochem. J.* **1988**, *250*, 1-8.

(158) Bright, G. R.; Fisher, G. W.; Rogowska, J.; Taylor, D. L. *J. Cell Biol.* **1987**, *104*, 1019-1033.

(159) Boron, W. F. *J. Membr. Biol.* **1983**, *72*, 1-16.

(160) Busa, W. B. *Ann. Rev. Physiol.* **1986**, *48*, 389-402.

(161) Bugg, C. E.; Thomas, J. M.; Sundaralingam, M.; Rao, S. T. *Biopolymers* **1971**, *10*, 175-219.

(162) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag Inc.: New York, 1984.

(163) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059-3061.

(164) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 9428-9430.

Chapter II

Footprinting of Oligonucleotides on Double Helical DNA Using MPE•Fe(II), DNase I, and Dimethyl Sulfate

Introduction

The work of Moser and Dervan indicated that pyrimidine oligonucleotide-directed triple helix formation offers a promising method for the recognition of extended (> 15 bp) sequences of DNA.[1] Pyrimidine oligonucleotides can, when equipped with a thymidine analogue tethered to EDTA•Fe(II) (T^*)[2] (Figure 1.1), bind and cleave a single 15 base pair site within 4 kilobases of double helical DNA.[1] The power of affinity cleaving to elucidate the structural principles for DNA recognition has been exemplified in the discovery of novel base triplets,[3] the synthesis of novel oligonucleotide binding heterocycles,[4] the study of abasic sites within triple helices,[5] and the analysis of novel triple helical motifs.[6] The determination of binding affinities using this technique requires the incorporation of the modified nucleoside thymidine-EDTA into the oligonucleotide, a modification which can result in changes in oligonucleotide binding affinity.[2,7] Analysis additionally requires the addition of reductants and is subject to interference by metals ubiquitous in biology.[8] If binding affinities of unmodified oligonucleotides are to be determined under conditions relevant to those *in vivo*, alternate methods of analysis are required.

In light of this, the footprinting of short (up to 15 base pairs) triple helical regions on restriction fragment size DNA has been undertaken. Footprinting can be accomplished using any reagent capable of non-specific DNA cleavage (Figure 2.1).[9-14] Analysis requires inhibition of cleavage at the binding site by the DNA binding molecule. Since the completion of this study, the footprinting of oligonucleotides by several reagents has been reported.[15,16]

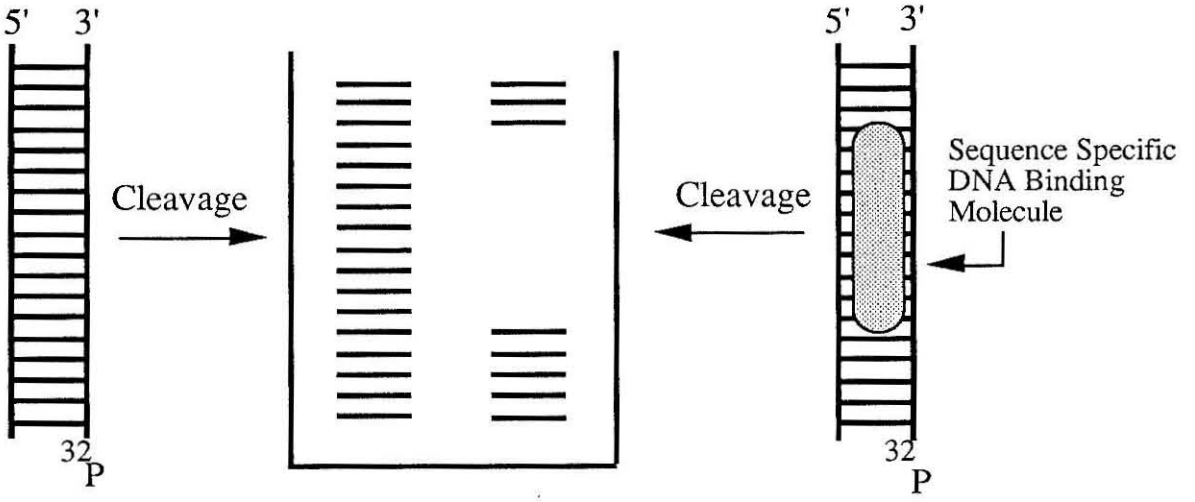


Fig. 2.1. Binding of a sequence specific DNA binding molecule interferes with cleavage by a non-specific DNA cleaving agent (**MPE**•Fe(II), DNase I, dimethyl sulfate) at the bound site.

MPE*•*Fe(II) footprinting. Procedures for the footprinting of oligonucleotides on DNA have been developed involving several reagents. **MPE**•Fe(II) footprinting is advantageous due to the highly non-sequence specific nature of this reagent, allowing high resolution definition of binding site size.[11,13] The ability of ethidium based molecules to bind to triple helical DNAs has been investigated by several groups. Ethidium has been shown to lower the melting temperature of the third strand in the triplex

poly(rA)•2 poly(rU), indicating a stabilization of the duplex poly(rA)•poly(rU) relative to the triplex upon ethidium intercalation.[17] Consistent with this, Le Pecq observed that ethidium preferentially binds to the duplex poly(rA)•poly(rU).[18,19] Lehrman and Crothers reported that RNA•DNA hybrids previously observed only as A-form conformations are induced into a B-form double helix by the presence of ethidium bromide,[20] suggesting the preferential binding of ethidium to B-form DNA. Triplexes, which are thought to exist in an A'-form conformation,[21-23] are therefore destabilized by ethidium bromide.[17] Ethidium is not thought to bind to triple helices containing alternating C+G-C base triplets, due to charge repulsion between the positive charges in the triplex structure and the ethidium cation.[24] In contrast to these studies, it has been reported that ethidium preferentially binds the *duplex* form of poly(dA)•poly(dT).[25] It remained to be determined if a *local* triplex could significantly inhibit intercalation by an ethidium-based reagent (**MPE**•Fe(II)) (binding affinity = $\sim 1.5 \times 10^5\ M^{-1}$)[26] at the triplex site.

DNase I footprinting. Binding of oligonucleotides in the major groove might be expected to result in a change in DNA geometry and groove width at the triplex site, factors which lead to changes in DNase I activity.[10] Polynucleotide triplexes have previously been shown to be resistant to DNase I and RNase degradation.[27] The data here were the first evidence that triple helices can inhibit enzyme activity at the binding site under physiological conditions. Since these studies were undertaken, oligonucleotides have been shown to interfere with DNA binding by a variety of proteins and to prevent subsequent protein activity at the triplex binding site.[16,28-30] Inhibition of enzyme activity (Ava I[28,29] or EcoR I[16] footprinting) can also be used as a method of determining oligonucleotide binding.

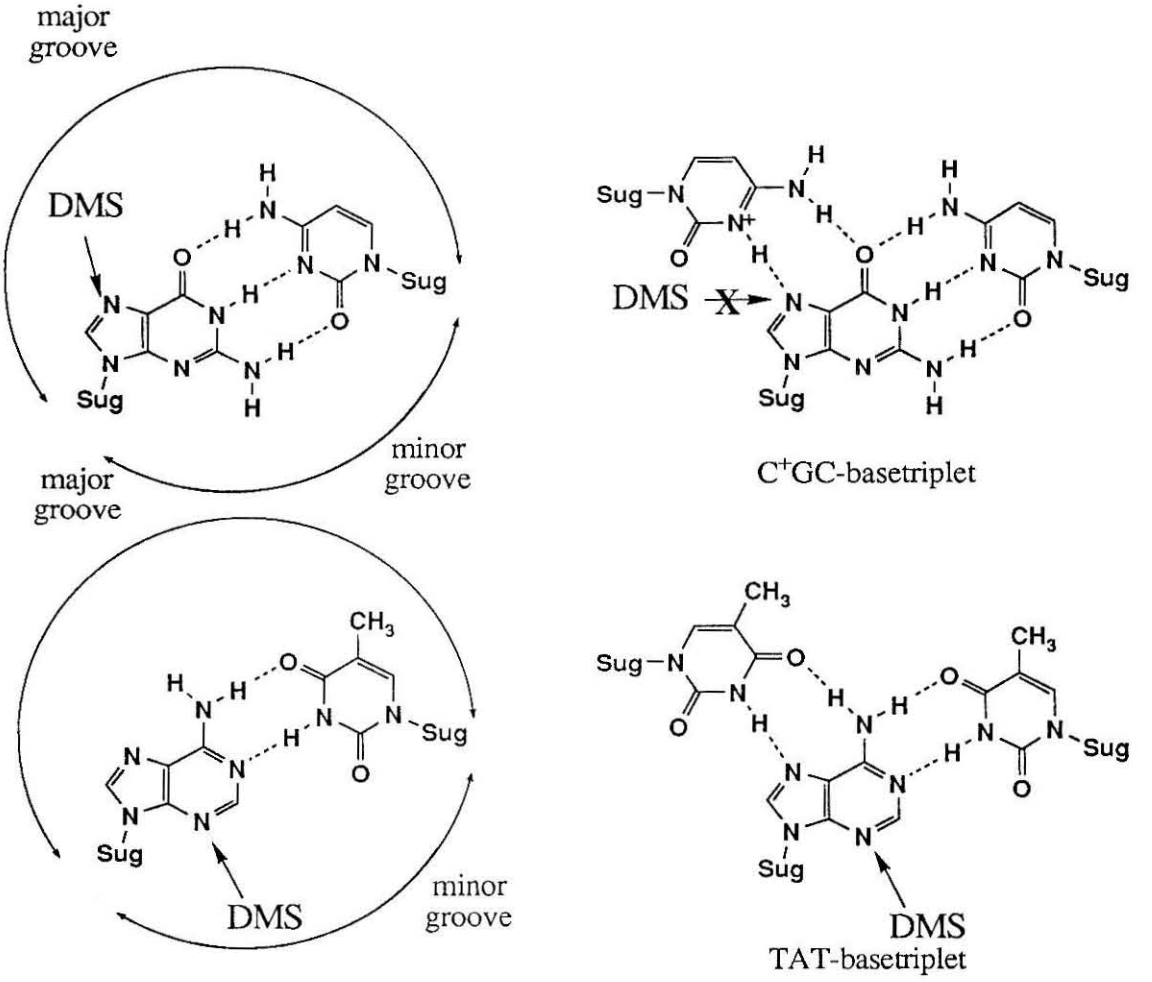


Fig. 2.2. Triple helix formation results in hydrogen bonding at N-7 of guanine bases and inhibition of DMS reactivity. No such decrease in methylation of N-3 of adenine results.

Dimethyl sulfate footprinting. Dimethyl sulfate alkylates DNA primarily at the N-7 position of guanine and N-3 position of adenine (Figure 2.2).[31-33] Base work-up results in cleavage at guanine residues, while acidic work-up results in preferential cleavage at adenine residues.[32,33] Within the pyrimidine motif, guanine bases involved in triple helix formation are hydrogen bonded in both Watson-Crick and Hoogsteen base pairing schemes

(Figure 2.2).[1,34] Hoogsteen hydrogen bonding of the third strand in the major groove is expected to hinder reactivity of guanine bases with alkylating agents such as dimethyl sulfate (DMS) in the major groove (i. e. reaction at N-7 of guanine). Methylation at N-3 of adenine, which occurs from the minor groove, is not expected to be affected. The footprinting of triplexes using dimethyl sulfate has been recently reported.[16]

Results and Discussion

Footprinting of oligonucleotides on double helical DNA is possible using a variety of DNA cleaving reagents, including **MPE**•Fe(II), DNase I, and dimethyl sulfate. Use of these reagents allows analysis of triple helix formation under conditions not amenable to affinity cleaving, enzyme inhibition, or transcriptional inhibition assays. The various footprinting reagents are complementary to each other, as well as to other techniques, for the determination of oligonucleotide affinity for double helical DNA.

MPE•Fe(II) footprinting

A review of the literature indicates that intercalation by ethidium was disfavored in A-form DNA and in triple helical complexes (see discussion above),[17-20,24] although at least one conflicting report has been published.[25] It remained to be determined if local triple helices, formed by binding of an oligonucleotide to a long piece of double helical DNA, are stable to intercalation by **MPE**•Fe(II), which has a binding constant for DNA of approximately $1.5 \times 10^5\ M^{-1}$.[26]

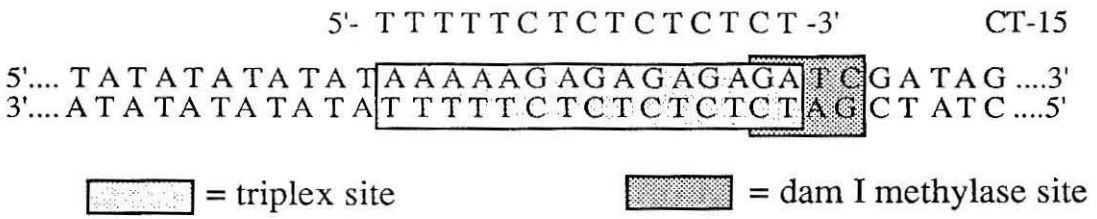
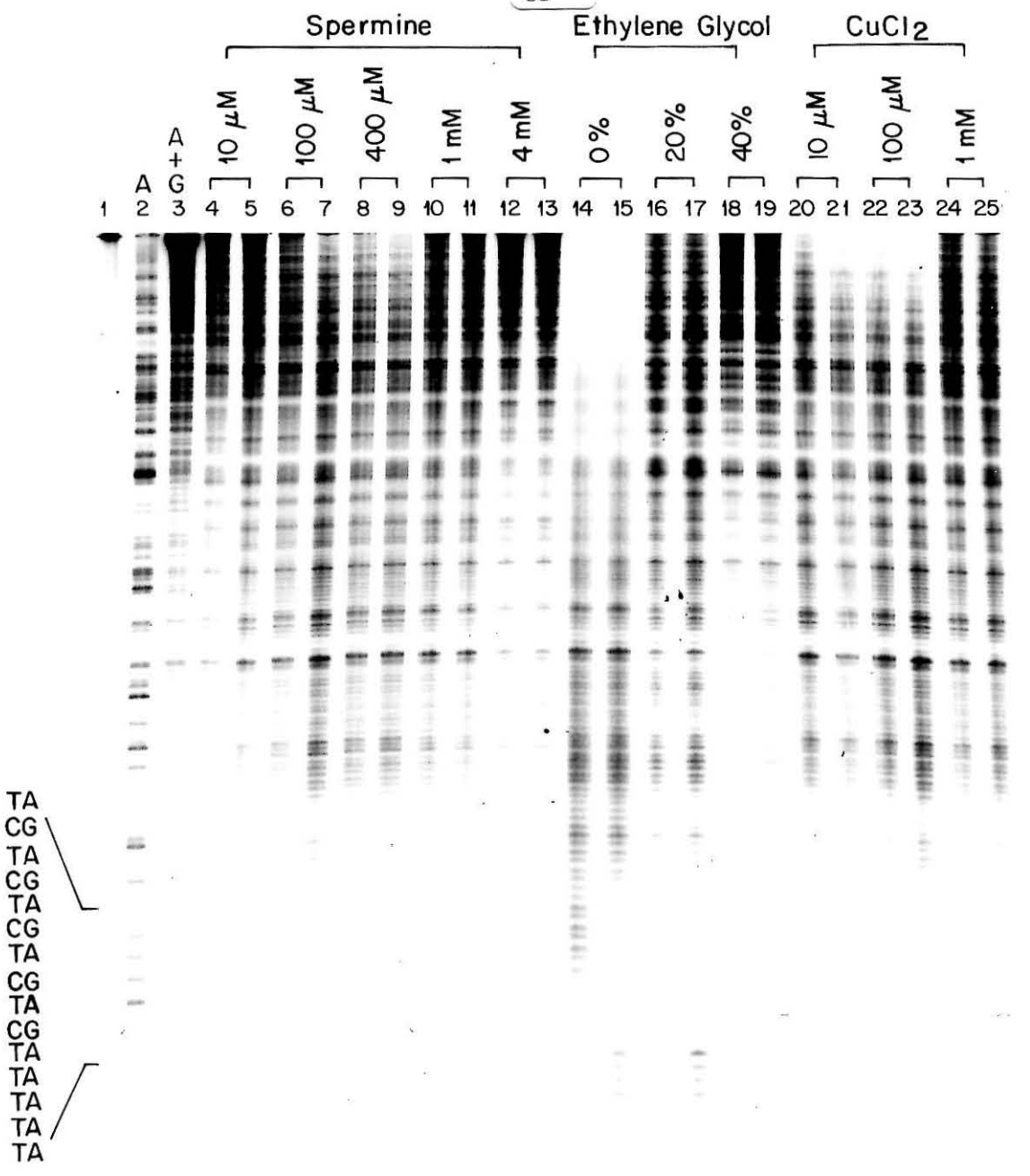


Fig. 2.3. Oligonucleotide CT-15 used in footprinting studies of triple helix formation. Oligonucleotide CT-15 binds a 15 base pair site within plasmid pDMAG10.[1,35] The final two base pairs are part of a dam I methylase site. The 3' terminal A in the triplex site is N-6 methylated in these studies.[36]

Effect of cations and organic co-solvents. Formation of polynucleotide complexes, including triplexes, is supported by the presence of cations.[1,37-40] Polycations such as spermine and $Co(NH_3)_6^{+3}$ support such structures at much lower concentrations than monovalent or divalent cations.[40,41] The effect of various polycations on triplex stability was studied using **MPE**•Fe(II) footprinting.

Binding of the oligonucleotide CT-15 (Figure 2.3) to its duplex DNA target site can be detected using **MPE**•Fe(II) footprinting under conditions in which triplexes are known to form (Figures 2.4 and 2.5).[1,42] In accordance with the work of Moser, who found efficient cleavage of DNA by EDTA•Fe(II) functionalized oligonucleotides in the presence of 1-10 mM $Co(NH_3)_6^{+3}$[42] or 1 mM spermine[1], clear protection against cleavage by **MPE**•Fe(II) is observed at the target sequence under these conditions. High concentrations of cations result in shielding of the DNA and a decrease in DNA cleavage by **MPE**•Fe(II) (Figure 2.4). Optimal footprints are observed at a $Co(NH_3)_6^{+3}$ concentration of 1 mM.

Figure 2.4. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMAG10 (EcoR I/Bgl I) indicating **MPE**•Fe(II) footprinting of oligonucleotide CT-15. Oligonucleotide CT-15 was equilibrated with duplex DNA under conditions stated at 0 °C for at least 60 minutes. **MPE** was equilibrated with 2 equivalents of Fe(II), and added as a 10x solution. After 15 minutes, reactions were activated with the addition of DTT to a final concentration of 2 mM. Reactions were allowed to run for 14 hours at 0 °C, and were stopped by precipitation with NaOAc and ethanol. All lanes: 20% ethylene glycol, 100 mM NaCl, 100 μM calf thymus DNA, 100 mM Tris pH 7.0 buffer unless noted. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G sequencing lane. Lane 3: Maxam-Gilbert A + G sequencing lane. Even lanes: no oligonucleotide. Odd lanes: 2 μM oligonucleotide CT-15. Lanes 4 and 5: 10 μM spermine; Lanes 6 and 7: 100 μM spermine; Lanes 8 and 9: 400 μM spermine; Lanes 10 and 11: 1 mM spermine; Lanes 12 and 13: 4 mM spermine; Lanes 14 - 19: 1 mM $Co(NH_3)_6^{+3}$; Lanes 14 and 15: no ethylene glycol; Lanes 16 and 17: 20% ethylene glycol; Lanes 18 and 19: 40% ethylene glycol; Lanes 20 and 21: 10 μM $CuCl_2$; Lanes 22 and 23: 100 μM $CuCl_2$; Lanes 24 and 25: 1 mM $CuCl_2$.


Spermine
Ethylene Glycol
CuCl2
10 μM
100 μM
400 μM
1 mM
4 mM
0 %
20 %
40 %
10 μM
100 μM
1 mM
A
A + G
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
TA
CG
TA
CG
TA
CG
TA
CG
TA
CG
TA
TA
TA
TA
TA

Among the divalent cations, Cu^{+2} is known to bind most tightly to DNA.[43] For this reason, oligonucleotide binding in the presence of Cu^{+2} was studied by **MPE**•Fe(II) and DMS footprinting techniques (*vide infra*). Micromolar concentrations of Cu^{+2} can be used to stabilize triple helix formation, and **MPE**•Fe(II) can be used under these conditions to footprint triple helical regions (Fig. 2.4). DMS footprinting indicates that binding of shorter oligonucleotides can be promoted by Cu^{+2} (*vide infra*). It has been postulated that Cu^{+2} binds tightly to N-3 of cytosine bases.[44-46] At this position, Cu^{+2} might replace the proton required for Hoogsteen hydrogen bonding. No studies on the effect of pH on oligonucleotide binding in the presence of Cu^{+2} have been undertaken.

The addition of organic solvents increases the cleavage efficiency of oligonucleotides-EDTA•Fe(II).[1] The addition of co-solvents may result in dehydration of the DNA[47] and stabilization of the B to A' transition which is thought to accompany oligonucleotide binding.[21-23] Ethylene glycol is necessary for footprinting of the triple helix by **MPE**•Fe(II) (Figure 2.4). High concentrations of ethylene glycol result in a noticeable decrease in the **MPE**•Fe(II) activity; under otherwise identical conditions, **MPE**•Fe(II) digestion is reduced in the presence of 40% ethylene glycol, due to the radical scavenging properties of this solvent. Approximately 20% ethylene glycol is optimal for **MPE**•Fe(II) footprinting of triple helical complexes.

Since footprinting involves competition for a binding site on DNA between the DNA cleaving reagent and the oligonucleotide, it is not surprising that robust triple helix formation conditions are required for clear **MPE**•Fe(II) footprints to be observed. The binding affinity of **MPE**•Fe(II) is estimated to be in the range of $1.5 \times 10^5\ M^{-1}$,[26] and several **MPE**•Fe(II) binding sites are located within each oligonucleotide binding site. **MPE**•Fe(II)

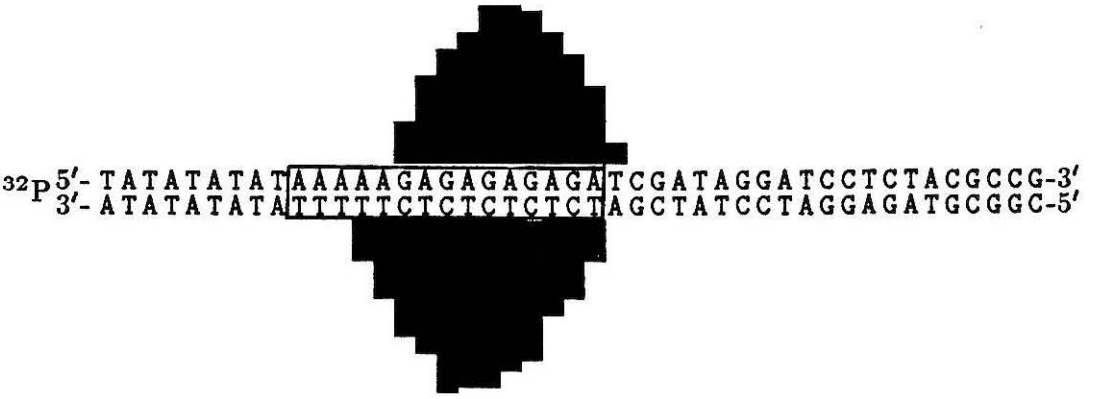
footprinting is the least sensitive of the three footprinting reagents used in this study.

Analysis of footprint size and location. A 3' shift in the cleavage pattern is observed, as is expected for cleavage originating from a minor groove DNA binder (Figure 2.5).[48] The footprint is transposed from the center of the binding site and extends to one side (the alternating G-A, or 3' end of the purine strand) of the binding site. This phenomenon may be due to the inherent stability associated with the B-form of poly(A):poly(T) DNA or a higher stability of C+G-C base triplets. The difficulty with which poly(A)•(T) tracts form triplexes has recently been documented.[49] Gas phase calculations support the greater stability of C+G-C base triplets.[50] The inherent lack of reactivity of **MPE**•Fe(II) at poly d(A):poly d(T) regions renders footprinting within this part of the binding site difficult. The use of spermine•EDTA, which displays avid activity at such sites, as a footprinting agent may help in elucidating this ambiguity.[51]

It has recently been reported that ethidium preferentially intercalates into the *triple helix* poly(dA)•2 poly(dT) compared to the poly(dT)•poly(dA) duplex.[25] No increase in **MPE**•Fe(II) activity is observed at the T_5•A_5-T_5 triplex.

MPE•Fe(II) footprints are known to extend two base pairs to the 3' side of the binding sites of various antibiotics; however, reduction in **MPE**•Fe(II) activity is observed only one base pair beyond the target sequence to the 3' end of the oligonucleotide (Figure 2.5). This suggests that the oligonucleotide 3' end is not fully bound to the DNA, and that the triplex-duplex transition and any change in DNA form associated with oligonucleotide binding is sharp and limited to a very few base pairs. Further evidence for these assertions is obtained from DMS footprinting studies (*vide infra*).

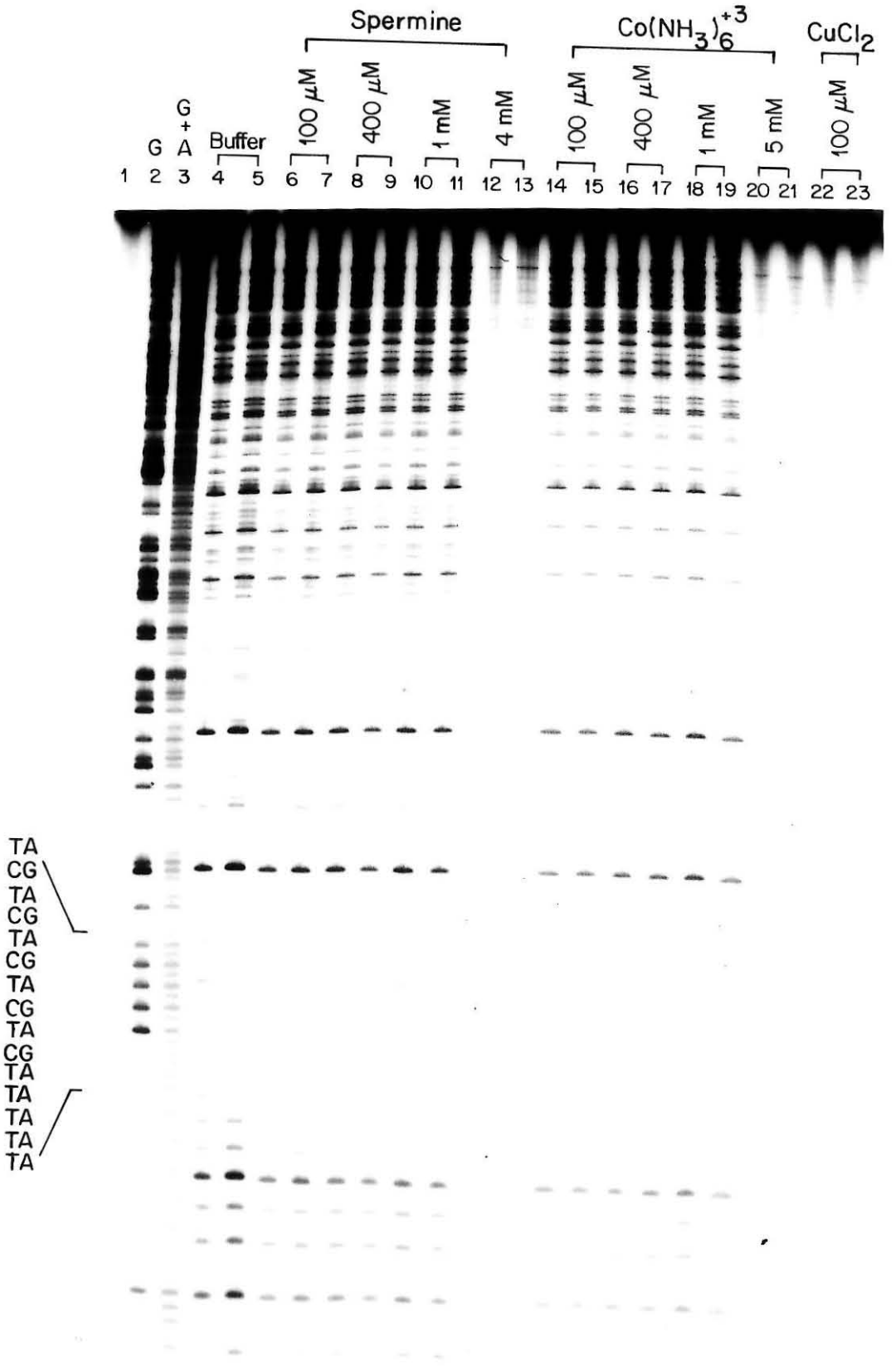
A. MPE •Fe(II) Footprinting



B. DNase I Footprinting

Figure 2.5. Histograms comparing the footprinting of oligonucleoitde CT-15 by **MPE•Fe(II)** and DNase I using data in Figures 2.4 and 2.6. Box indicates the oligonucleotide binding site. Bars represent the extent of inhibition of cleavage by (A) **MPE•Fe(II)** or (B) DNase I. Inhibition of cleavage was determined by densitometry of autoradiograms from Figures 2.4 and 2.6.

Figure 2.6. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMAG10 (EcoR I/Bgl I) indicating DNase I footprinting of oligonucleotide CT-15. Oligonucleotide CT-15 was equilibrated with duplex DNA at 0 °C for at least 60 minutes. DNase I was added to a final concentration of 40 ng/ml, and the reaction allowed to proceed for 12 minutes at 4 °C. Reactions were stopped by precipitation with 0.25 volumes of a 10 M NH_4OAc, 200 mM EDTA solution and 3 volumes of ethanol. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G sequencing lane. Lane 3: Maxam-Gilbert G + A sequencing lane. Even lanes: no oligonucleotide. Odd lanes: 2μM oligonucleotide CT-15 oligonucleotide. All lanes 0% ethylene glycol, 100 μM calf thymus DNA, 100 mM Tris pH 7.0 buffer, 10 mM Mg^{+2}, 10 mM Ca^{+2} and 5 mM K^{+}. Lanes 4 and 5: no additional cations; Lanes 6 and 7: 100 μM spermine; Lanes 8 and 9: 400 μM spermine; Lanes 10 and 11: 1 mM spermine; Lanes 12 and 13: 4 mM spermine; Lanes 14 and 15: 100 μM $Co(NH_3)_6^{+3}$; Lanes 20 and 21: 5 mM $Co(NH_3)_6^{+3}$; Lanes 22 and 23: 100 μM $CuCl_2$.


Spermine
Co(NH3)6+3
CuCl2
G
G + A
Buffer
100 μM
400 μM
1 mM
4 mM
100 μM
400 μM
1 mM
5 mM
100 μM
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
TA
CG
TA
CG
TA
CG
TA
CG
TA
CG
TA
TA
TA
TA
TA

DNase I Footprinting

Polynucleotide triple helices are known to be resistant to both DNase I and RNase degradation.[27] DNase I activity is optimal in the presence of both Mg^{+2} and Ca^{+2} ions; therefore, all DNase I reactions are run in 10 mM of each Mg^{+2} and Ca^{+2}.[52] Due to the high cationic content of the buffer, the addition of spermine, $Co(NH_3)_6^{+3}$ or ethylene glycol to stabilize triple helix formation is not necessary (Figure 2.6). The addition of ethylene glycol results in reduced DNase I activity, as do high concentrations of spermine (> 1 mM) or $Co(NH_3)_6^{+3}$ (> 1 mM). $CuCl_2$, in 10 mM Mg^{+2} and 10 mM Ca^{+2}, destroys DNase I activity even at low concentrations (100 μM) (Fig. 2.6).

DNase I footprint size and location. DNase I footprints extend five base pairs beyond each side of the target site (Figures 2.5 and 2.6). Presumably, the enzyme makes contacts with the DNA several base pairs removed from the site of cleavage.[53] Disruptions of these contacts results in reductions in cleavage intensity. It is interesting that Hanvey and Wells report that protection of duplex DNA from endonuclease digestion requires that the oligonucleotide and restriction endonuclease recognition sites overlap.[16] Binding of the oligonucleotide to a site abutting the restriction enzyme recognition site results in marked decreases in the ability of the oligonucleotide to inhibit DNA cleavage.[16] This is in contrast to the results observed with DNase I, where binding of the oligonucleotide several base pairs away from the site of cleavage impairs DNase I activity.

Footprinting with Dimethyl Sulfate

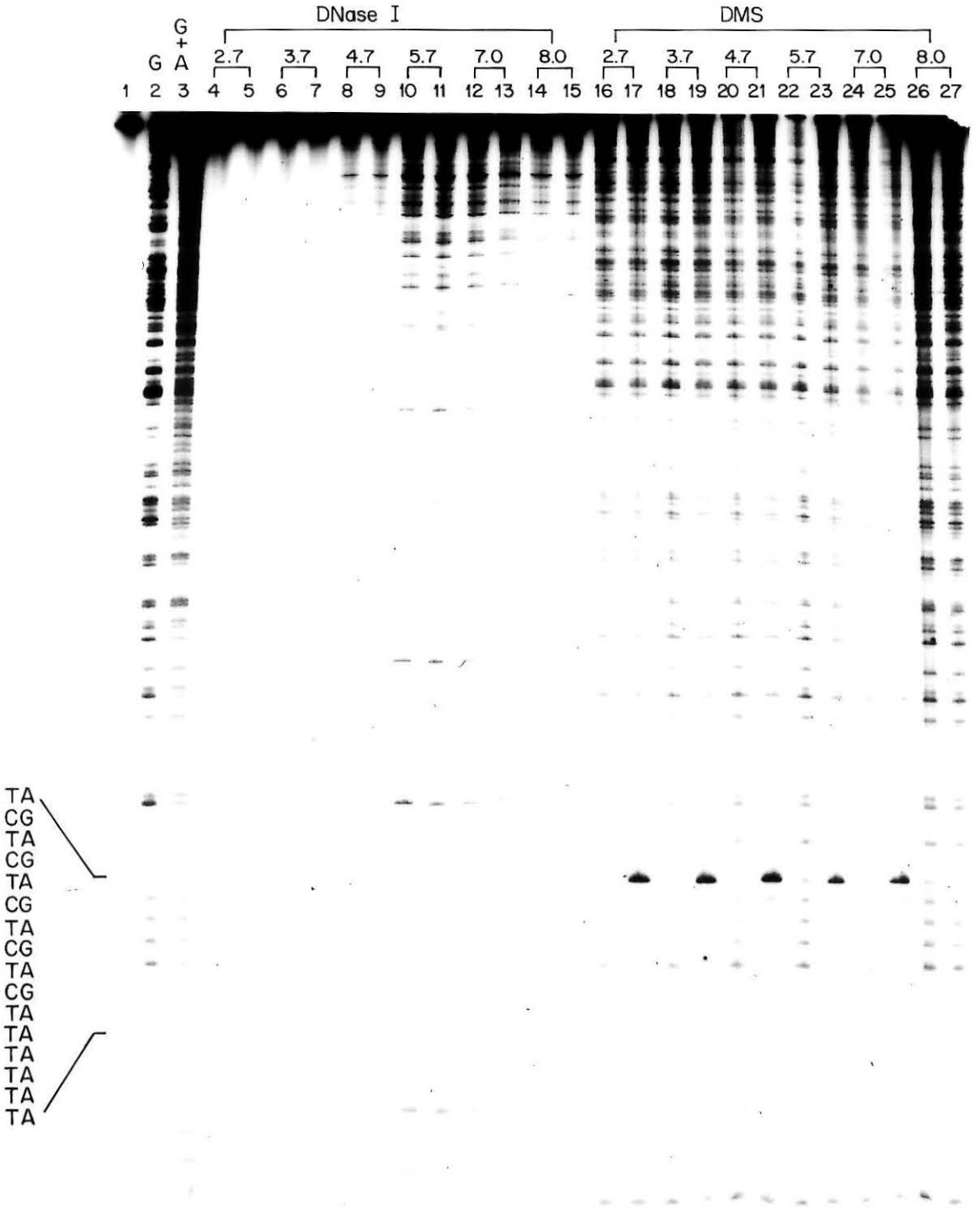
Reaction of DNA with dimethyl sulfate (DMS) results primarily in methylation at N-7 of guanine (reaction from the major groove) and N-3 of adenine (reaction from the minor groove).[31-33] Reaction with N-7 of guanine (major groove) is expected to be inhibited by binding of a third strand in the major groove (Fig. 2.2). No change in reactivity at N-3 of adenine is expected. The effect of oligonucleotide binding on the reactivity of DMS with both guanine and adenine bases within the triple helical complex was investigated. Footprinting of oligonucleotides using DMS and piperidine work-up (G reaction) has been reported since the completion of this work.[16]

The breadth of conditions assayable using DMS footprinting is displayed in Figure 2.7. While DNase I footprinting is possible only at pHs ≥ 5.7, DMS protection studies indicate that the triplexes are stable at low pH (to at least a pH of 2.7), but dissociate at slightly neutral pH (pH = 8).

In DMS footprinting studies, $CuCl_2$ was used to stabilize triple helix formation. $CuCl_2$ was shown by DMS footprinting to be superior to $Co(NH_3)_6^{+3}$ in promoting binding of oligonucleotides to duplex DNA.

Upon formation of the triple strand, a large enhancement in the reactivity of the guanine base at the 3' end of the binding sequence and a decrease in the reactivity of the guanines involved in triple helix formation is observed (for histogram analysis, see Fig. 2.10). A similar increase in the reactivity at the 3' triplex-duplex junction has also been observed with Cu(II)•phenanthroline.[15] No such enhancement in cleavage was observed on the pyrimidine strand at the 5' end.[15]

Figure 2.7. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMAG10 (EcoR I/Bgl I) indicating DNase I and DMS footprinting of oligonucleotide CT-15 as a function of pH. Oligonucleotide CT-15 was incubated with duplex DNA under conditions stated for 60 minutes at 0 °C. DNase I footprinting was effected by addition of a 10x solution of DNase I to a final concentration of 40 ng/ml. Reactions were conducted at 4 °C for 12 minutes, and were quenched by addition of 0.25 volumes of a 10 M NH_4OAc, 200 mM EDTA solution and 3 volumes of ethanol. DMS footprinting was effected by addition of 4 µl of a freshly prepared 2% aqueous solution of dimethyl sulfate. After 90 seconds at 0 °C, reactions were quenched with 5 µl of a DMS stop solution and 785 µl of ethanol.[32,33] After washing with 70% ethanol, strand scission was obtained after incubation at 90 °C for 30 minutes in 10% aqueous piperidine. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G sequencing lane. Lane 3: Maxam-Gilbert A + G sequencing lane. Lanes 4-15: DNase I footprinting, 0% ethylene glycol, 100 µM calf thymus DNA, 100 mM Tris, pH 7.0, 10 mM $MgCl_2$, 10 mM $CaCl_2$, and 5 mM KCl. Even lanes: no oligonucleotide. Odd lanes: 2 µM oligonucleotide CT-15. Lanes 4 and 5: pH 2.7; Lanes 6 and 7: pH 3.7; Lanes 8 and 9: pH 4.7; Lanes 10 and 11: pH 5.7; Lanes 12 and 13: pH 7.0; Lanes 14 and 15: pH 8.0. Lanes 16-27: DMS footprinting, 20% ethylene glycol, 100 µM calf thymus DNA, 100 mM Tris, pH 7.0, 100 mM NaCl, 1 mM $Co(NH_3)^{+3}$. Lanes 16 and 17: pH 2.7; Lanes 18 and 19: pH 3.7; Lanes 20 and 21: pH 4.7; Lanes 22 and 23: pH 5.7; Lanes 24 and 25: pH 7.0; Lanes 26 and 27: pH 8.0.


DNase I
DMS
G
G + A
2.7
3.7
4.7
5.7
7.0
8.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27
TA
CG
TA
CG
TA
CG
TA
CG
TA
CG
TA
TA
TA
TA
TA
TA

Characterization of 3' triplex-duplex junction. Protection of the internal guanine residues is a direct result of formation of Hoogsteen hydrogen bonds at N-7 of guanine. One explanation for the hyperreactivity observed at the 3' end of the binding site is an alteration in the DNA structure at the triplex to duplex junction. The fact that the hyperreactivity is limited to a single guanine indicates that this transition is confined to the bases directly adjacent to the triple helix binding site. Such junctions have been modelled and structures proposed. Selsing and Wells have stated that a B to A transition could be accommodated within a single base pair in which the sugar ring on one strand is C2' endo puckered (as in B DNA) while the complementary sugar on the opposite strand exhibits C3' endo puckering (as is observed for A DNA).[54] The transition does not result in interruption in base stacking, introduce unpaired bases, or result in unwinding of the DNA. Experimental evidence for this model is limited to an NMR study of the oligonucleotide $rC_{11}\bullet dC_{16}\text{-}dG_n$,[54] which indicates that the A to B transition does not occur over an extended stretch of bases.[54]

To characterize the nature of the triplex to duplex junction, the junction was treated with several DNA modification reagents. S1 nuclease efficiently cleaves single stranded DNA regions, as well as duplex DNA containing denatured or single stranded regions.[55-57] To determine if single stranded or denatured regions are introduced upon triplex formation, the DNA restriction fragment containing the triplex binding site was treated with S1 nuclease after incubation with CT-15. Under conditions in which both formation of the triplex and S1 nuclease activity had been independently verified, no sites reactive to S1 nuclease were discovered, indicating that the hyperreactivity observed with dimethyl sulfate is not due to the presence of

single stranded or denatured regions. These experiments were conducted under conditions conducive to triple helix formation (pH 4.7); nonetheless, it cannot be ruled out that the nuclease primarily digests the excess oligonucleotide present in solution. Dissociation of the bound oligonucleotide and loss of the local triple helix structure may subsequently result.

To determine the effect of triple helix formation on DNA reactivity from the minor groove, DNA with a local triplex (CT-15 complex) was methylated with DMS and treated with acid under conditions known to result in strand cleavage at methylated adenine residues.[32,33] Treatment of the triple helical complex in this manner results in strand scission at all adenines in the binding site, and at the guanine which displayed hyperreactivity with piperidine work-up. Methylation from the minor groove is not enhanced or reduced by the presence of the oligonucleotide and formation of the triple helical complex. Strand scission at the hyperreactive guanine occurs because this becomes the overwhelming site of DMS reaction once triple helix formation has occurred.

Triplex formation similarly does not affect the reactivity of the DNA to diethyl pyrocarbonate, a minor groove modification reagent. Thus, the triplex to duplex junction is characterized by an increase in reactivity in the major groove at the 3' end of the triplex site, but little change in reactivity in the minor groove.

Rate of Oligonucleotide Hybridization. The rapid rate of DNA methylation by DMS allows study of the kinetics of oligonucleotide binding. The oligonucleotide CT-15 (2 μM) confers hyperreactivity to the guanine base in less than one minute, indicating that when micromolar concentrations of oligonucleotide are used, binding is rapid (data not shown). A more

Figure 2.8. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMAG10 (EcoR I/Bgl I) indicating DNase I and DMS footprinting of oligonucleotide CT-15 as a function of oligonucleotide concentration. Oligonucleotide CT-15 was equilibrated with duplex DNA under the conditions and concentration specified. DNase I footprinting was effected by addition of a 10x solution of DNase I to a final concentration of 40 ng/ml. Reactions were conducted at 4 °C for 12 minutes and were quenched by addition of 0.25 volumes of a 10 M NH_4OAc, 200 mM EDTA solution and 3 volumes of ethanol. DMS footprinting was effected by addition of 4 µl of a freshly prepared 2% aqueous solution of dimethyl sulfate. After 90 seconds at 0 °C, reactions were quenched with 5 µl of a DMS stop solution and 785 µl of ethanol.[32,33] After washing with 70% ethanol, strand scission was obtained after incubation at 90 °C for 30 minutes in 10% aqueous piperidine. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G + A sequencing lane. Lanes 3-9: DMS footprinting, 20% ethylene glycol, 100 µM calf thymus DNA, 100 mM Tris, pH 7.0, 100 mM NaCl, 1 mM $Co(NH_3)^{+3}$. Lane 3: no oligonucleotide; Lanes 2-9: 2 µM, 660 nM, 200 nM, 66 nM, 20 nM, and 6 nM oligonucleotide CT-15. Lanes 10-16:L DNase I footprinting, 0% ethylene glycol, 100 µM calf thymus DNA, 100 mM Tris pH 7.0, 10 mM Mg^{+2}, 10 mM Ca^{+2}, 5 mM K^{+}. Lane 10: no oligonucleotide; Lanes 11-16: 2 µM, 660 nM, 200 nM, 66 nM, 20 nM and 6 nM oligonucleotide CT-15.

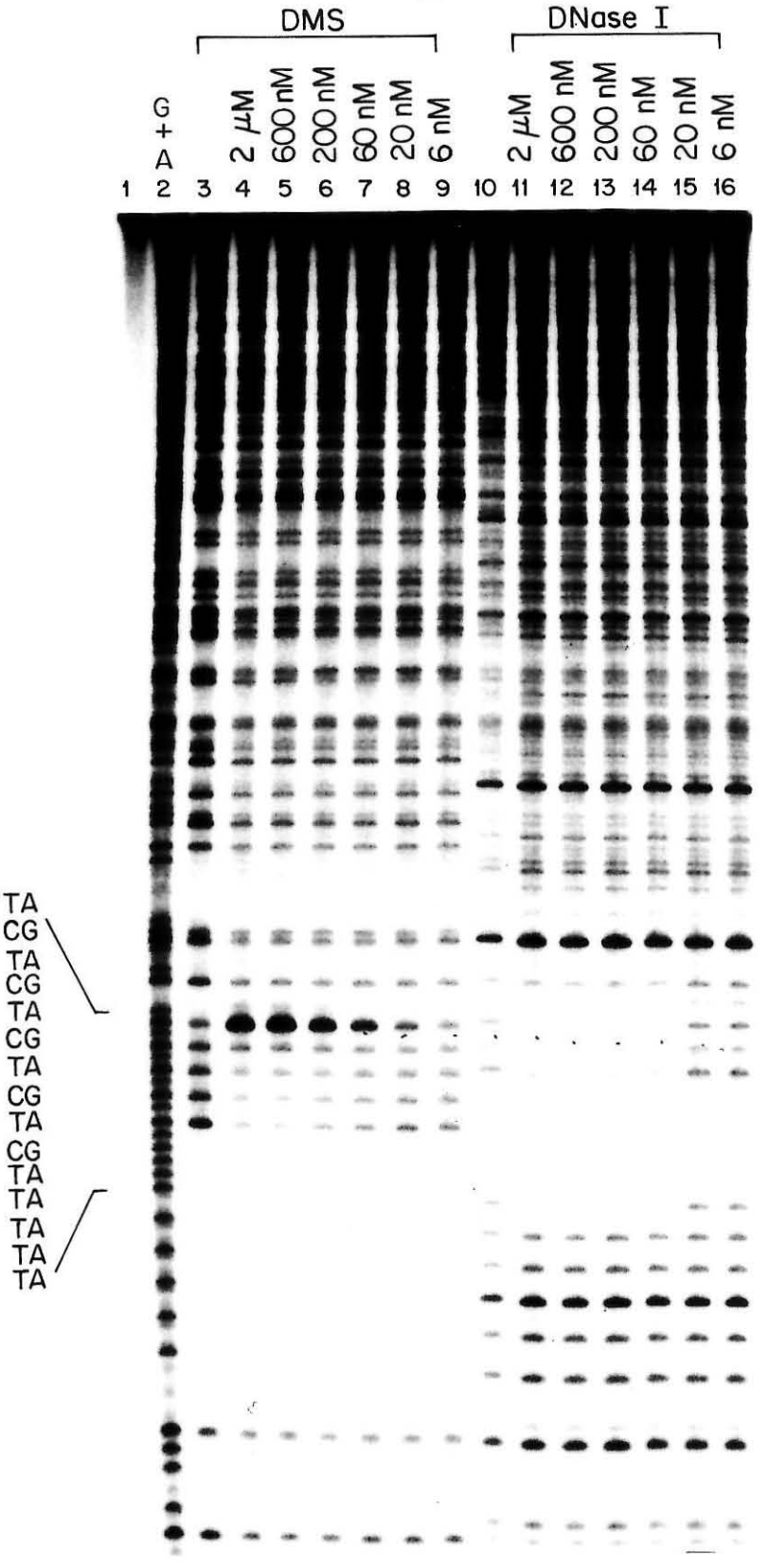
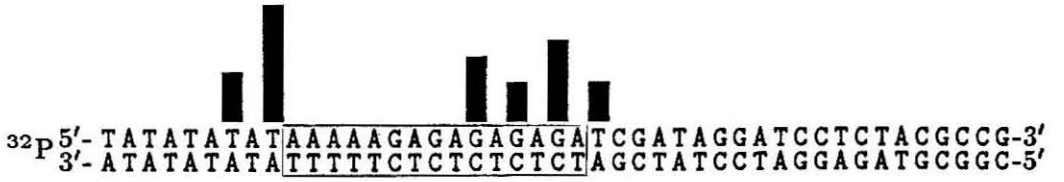
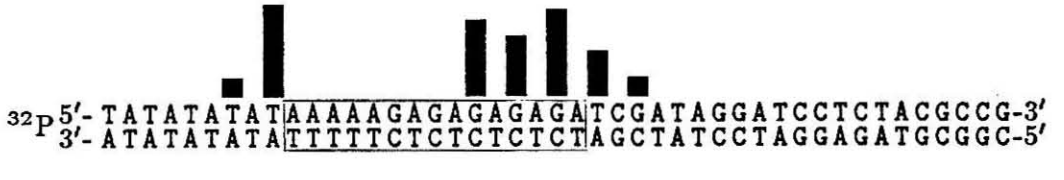

DMS
DNase I
G + A
2 μM
600 nM
200 nM
60 nM
20 nM
6 nM
2 μM
600 nM
200 nM
60 nM
20 nM
6 nM
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
TA
CG
TA
CG
TA
CG
TA
CG
TA
CG
TA
TA
TA
TA
TA

Fig. 2.9. Histograms obtained by densitometry of autoradiogram from Fig. 2.8 indicating the footprinting of oligonucleotide CT-15 using DNase I as a function of oligonucleotide concentration. Bars indicate extent of inhibition of cleavage at base indicated. Box indicates oligonucleotide binding site.

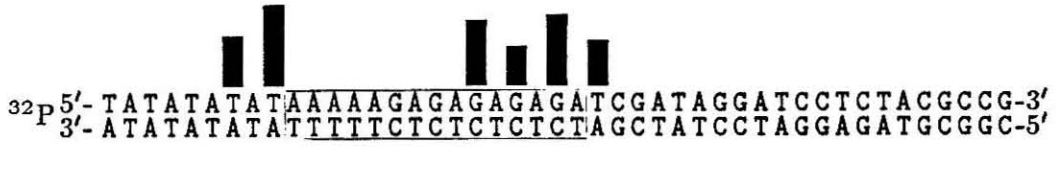
DNase I Footprinting with 2000 nM CT-15 Probe



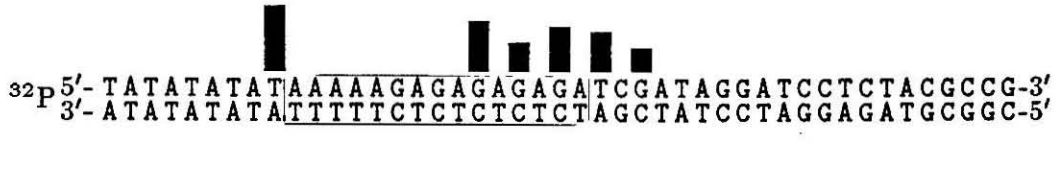
DNase I Footprinting with 600 nM CT-15 Probe



DNase I Footprinting with 200 nM CT-15 Probe



DNase I Footprinting with 60 nM CT-15 Probe



DNase I Footprinting with 20 nM CT-15 Probe

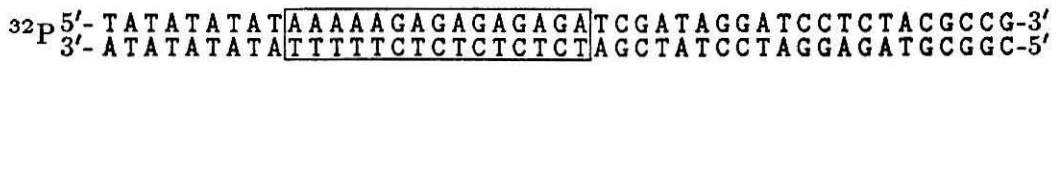
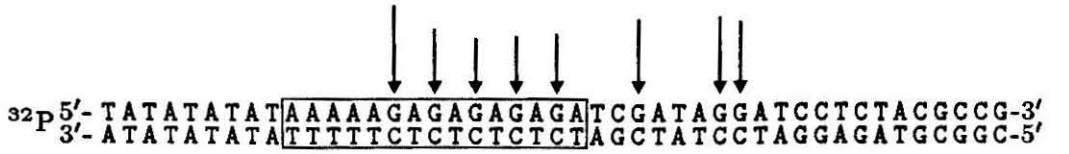
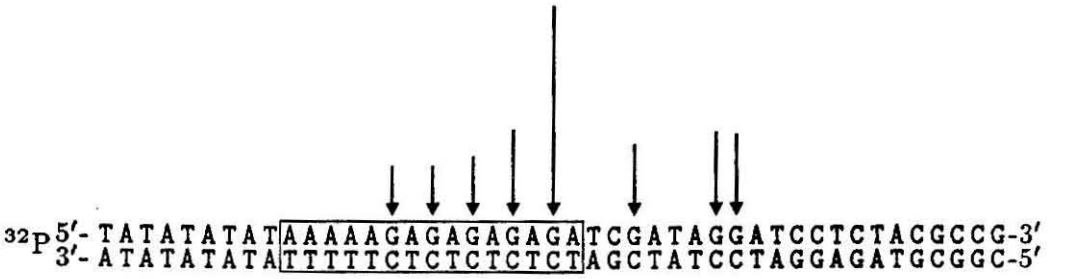


Fig. 2.10. Relative reactivities of guanines to dimethyl sulfate as a function of oligonucleotide concentration as determined by densitometry of autoradiogram in Fig. 2.8. Arrows indicate extent of cleavage at base indicated. Box indicates oligonucleotide binding site.

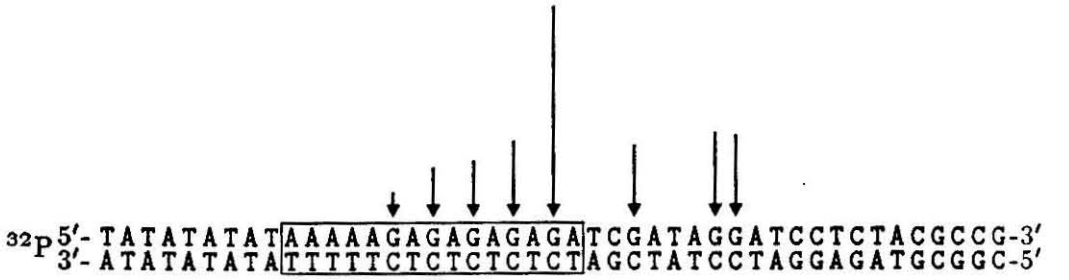
No Probe



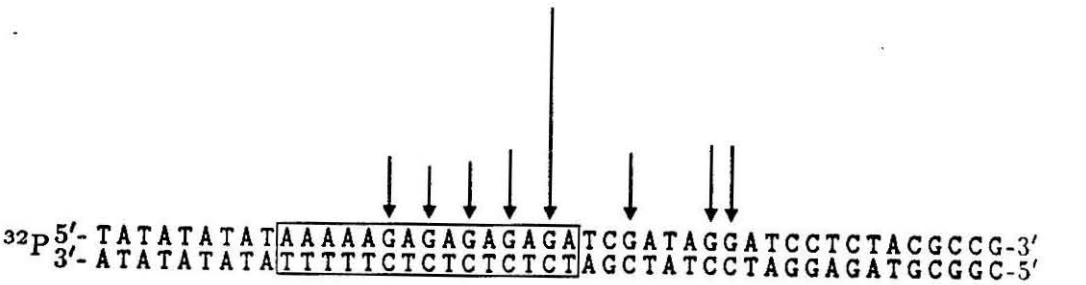
Probe CT-15; 2 μM



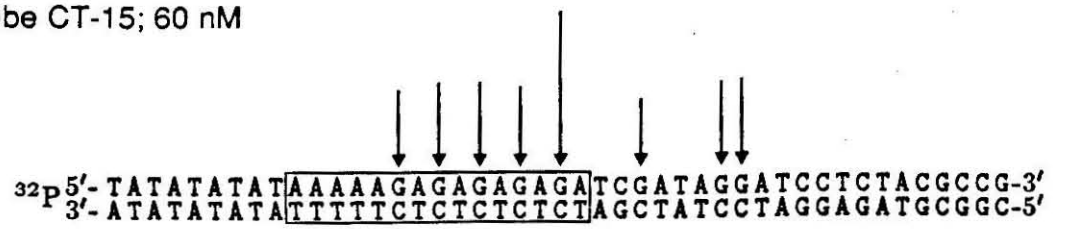
Probe CT-15; 600 nM



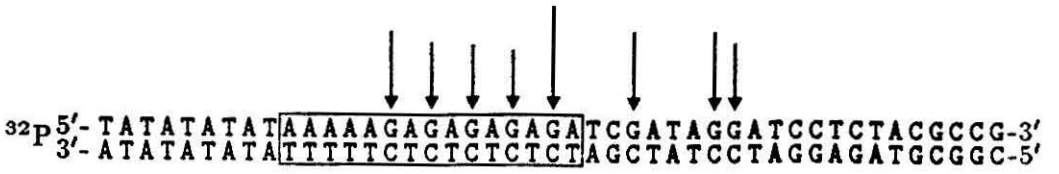
Probe CT-15; 200 nM



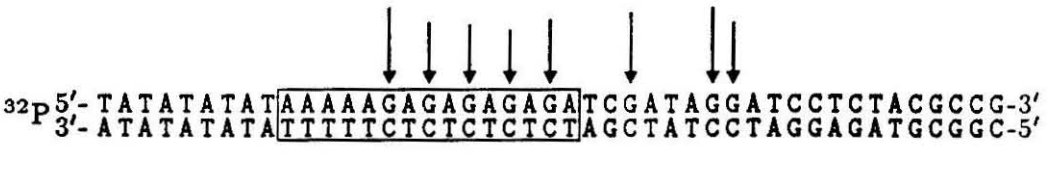
Probe CT-15; 60 nM



Probe CT-15; 20 nM



Probe CT-15; 6 nM



comprehensive study of the kinetics of triple helix formation has recently been reported.[29] The reported results are consistent with these observations.

Binding affinities and oligonucleotide base composition characterized by footprinting.

Determination of binding affinities using DNase I and DMS footprinting. DNase I footprinting as a function of oligonucleotide concentration indicates that binding under conditions of 10 mM Mg^{+2}, 10 mM Ca^{+2}, 5 mM K^{+}, 0% ethylene glycol, pH 7.0 is observed at an oligonucleotide (CT-15) concentration as low as 60 nM (Figures 2.8 and 2.9). Similar results are obtained using DMS footprinting under slightly different conditions (20% ethylene glycol, 25 mM pH 7.0 Tris, 100 mM NaCl, 1 mM $Co(NH_3)_6^{+3}$) (Fig. 2.8 and 2.10). A binding constant of approximately $1 \times 10^7\ M^{-1}$ is inferred. Since the completion of these studies, Maher has shown that the association rates of oligonucleotides for their binding sites are quite slow.[29] At low concentrations, equilibrium may not have been attained during the 60 minute incubation time used in these studies; therefore, the actual binding constant may be higher than that reported here.[29]

Studies of the effect of oligonucleotide composition using DNase I and DMS footprinting. To delineate the role of base composition and length on binding affinity oligonucleotides of various length and base compositions were synthesized and their relative binding affinities determined by DNase I and DMS footprinting (Figure 2.11). Oligonucleotides CT-13, CT-11, CT-09 and CT-07 are sequentially shortened by two bases from the 3' end of the

oligonucleotide CT-15, incrementally removing a C as well as a T residue. Oligonucleotides CT-12, CT-10 and CT-08 are analogous to CT-15, CT-13, and CT-11; however, 3 T bases from the 5' end of the oligonucleotide have been deleted. This results in a series of oligonucleotides with a slightly increased C:T content. Two oligonucleotides 15 bases in length which result in T-G and C-A mismatches upon triplex formation were also constructed. These oligonucleotides are analogous to those studied by Moser using affinity cleaving.

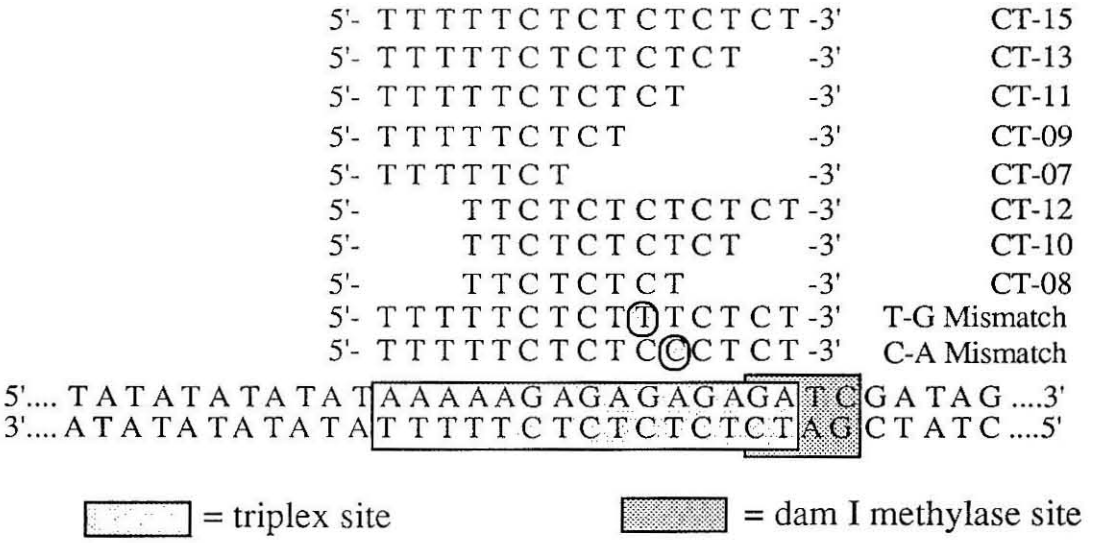


Fig. 2.11. Oligonucleotides used in footprinting studies of triple helix formation. Oligonucleotides bind a site within plasmid pDMAG10.[35,58] The final two base pairs are part of a dam I methylase site. The mismatched base in the T-G and C-A mismatched probes are indicated. The 3' terminal A within the triplex site is N-6 methylated in these studies.

Footprinting using DNase I indicates that the binding affinities of oligonucleotides less than 13 bases in length decrease sharply (Fig 2.12 and 2.13). Under the conditions studied (2 μM oligonucleotide, 10 mM Mg^{+2} and Ca^{+2}, 25 mM Tris, 5 mM K^{+}, pH 7.0), the relative affinities of the

oligonucleotides studied is CT-15 > CT-13 > CT-11 = C-A mismatch > CT-12 > CT-09 = CT-10 > T-G mismatch > CT-08 > CT-07. Binding by the oligonucleotides CT-07 or CT-08 can barely be detected. As oligonucleotide length is decreased, a concomitant decrease in the size of the footprint is observed, the footprint consistently remaining approximately 5 base pairs larger than the oligonucleotide target site.

The binding of these oligonucleotides was also studied by DMS footprinting. Based on the variable temperature data gathered in Figure 2.14, the melting temperature of the oligonucleotide CT-09 (2 mM $CuCl_2$, 2 μM oligonucleotide, 20 % ethylene glycol, 25 mM Tris pH 7.0) is approximately 12 °C. Oligonucleotide CT-11 forms a stable triplex at temperatures > 25 °C.

Oligonucleotide CT-10, which binds DNA via the formation 4 C+G•C base triplets and 6 T•A-T base triplets, does not bind duplex DNA at 25 °C. Oligonucleotides CT-09, which forms 3 C+G•C triplets and 6 T•A-T triplets, melts at 12 °C. By comparison, oligonucleotide CT-11, which forms 4 C+G•C triplets and 7 T•A-T triplets, melts at >25 °C. The series of oligonucleotides lacking the T_5 run at the 5' end appears to form triplexes of lower stability than do the oligonucleotides containing T_5 tails. The higher C content of these oligonucleotides may result in lower triplex stability at the pH used in this study. Frank-Kamenetskii and co-workers have suggested that at low pH, C+G-C base triplets are more stable than T•A-T triplets.[59,60] The stability of the triplexes formed by these two sets of oligonucleotides as a function of pH has not been determined.

The oligonucleotides CT-13, CT-11, CT-09, CT-07, CT-10, and CT-08 each induce hyperreactivity to a single guanine base immediately adjacent to and outside of the oligonucleotide binding site (Figure 2.14-2.16). In contrast, oligonucleotides CT-15 and CT-12 confer increased reactivity to a guanine

Figure 2.12. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMAG10 (EcoR I/Bgl I) indicating DNase I footprinting of various oligonucleotides. Oligonucleotides (final concentration 2 μM) were equilibrated with duplex DNA in 100 μM calf thymus DNA, 100 mM Tris buffer, pH 7.0, 10 mM $MgCl_2$, 10 mM $CaCl_2$, 5 mM KCl. DNase I footprinting was effected by addition of a 10x solution of DNase I to a final concentration of 40 ng/ml. Reactions were conducted at 4 °C for 12 minutes, and were quenched by addition of 0.25 volumes of a 10 M NH_4OAc, 200 mM EDTA solution and 3 volumes of ethanol. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G + A sequencing lane. Lane 3: Maxam-Gilbert G sequencing lane. Lane 4: no oligonucleotide; Lane 5: Oligonucleotide CT-15; Lane 6: Oligonucleotide CT-13; Lane 7: Oligonucleotide CT-11; Lane 8: Oligonucleotide CT-09; Lane 9: Oligonucleotide CT-07; Lane 10: no oligonucleotide; Lane 11: Oligonucleotide CT-12; Lane 12: Oligonucleotide CT-10; Lane 13: Oligonucleotide CT-08; Lane 14: TG mismatch oligonucleotide; Lane 15: CA mismatch oligonucleotide.

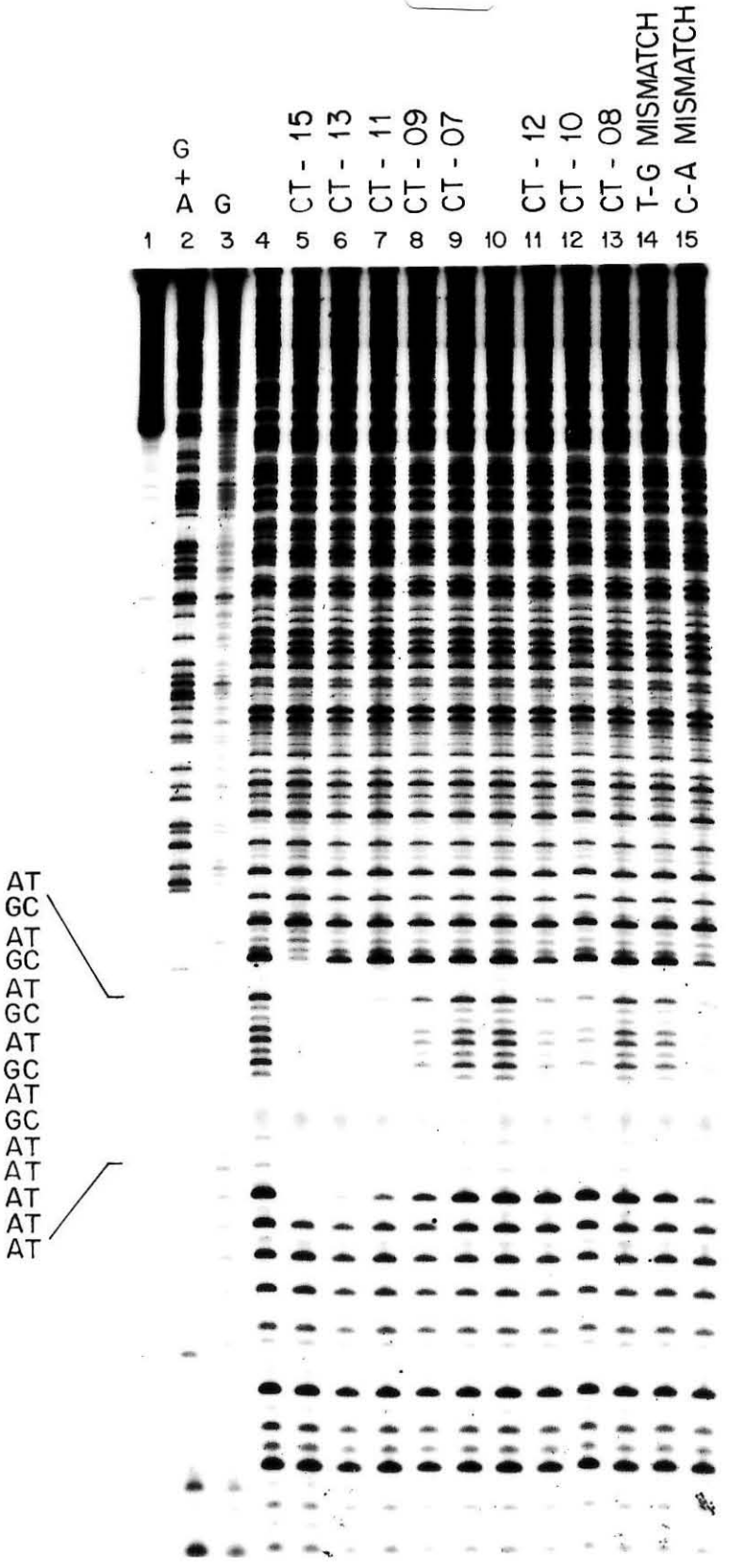
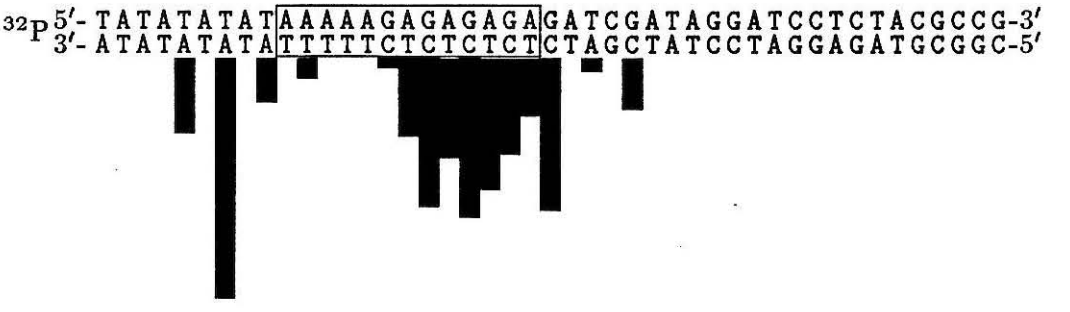

1
2
G
A + G
3
4
CT - 15
5
CT - 13
6
CT - 11
7
CT - 09
8
CT - 07
9
10
CT - 12
11
CT - 10
12
CT - 08
13
T-G MISMATCH
14
C-A MISMATCH
15
AT
GC
AT
GC
AT
GC
AT
GC
AT
GC
AT
AT
AT
AT
AT

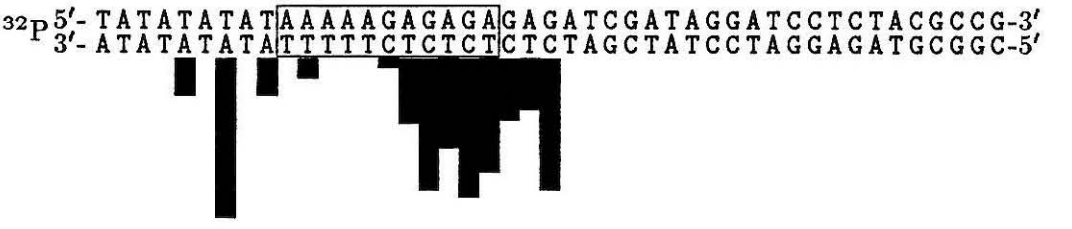
Fig. 2.13. Histograms obtained by densitometry of autoradiogram from Figure 2.12, indicating the binding of oligonucleotides of various lengths and base composition as determined by DNase I footprinting. Bars indicate extent of inhibition of cleavage at base indicated. Boxes indicate oligonucleotide binding site.

DNase I Footprinting of CT-15 Probe

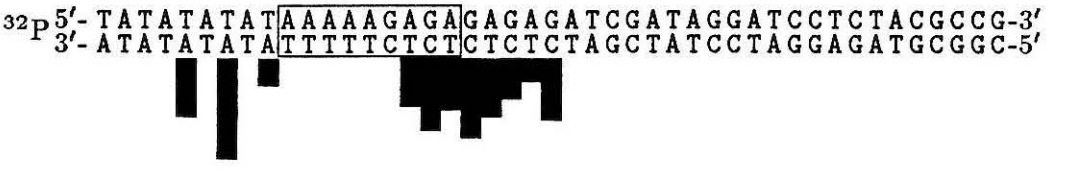
DNase I Footprinting of CT-13 Probe



DNase I Footprinting of CT-11 Probe

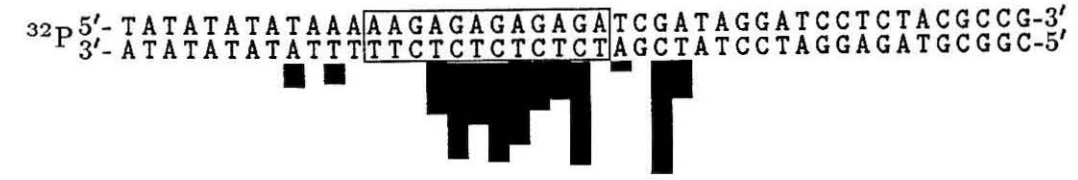


DNase I Footprinting of CT-09 Probe



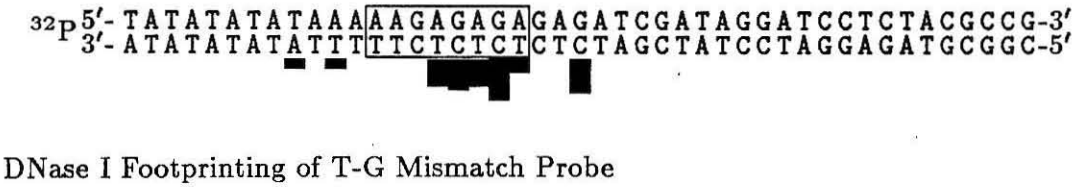
DNase I Footprinting of CT-07 Probe

DNase I Footprinting of CT-12 Probe



DNase I Footprinting of CT-10 Probe

DNase I Footprinting of CT-08 Probe



DNase I Footprinting of T-G Mismatch Probe

DNase I Footprinting of C-A Mismatch Probe

Figure 2.14. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMAG10 (EcoR I/Bgl I) indicating DMS footprinting of oligonucleotides CT-15, CT-13, CT-11, CT-09 and CT-07. Oligonucleotides (final concentration 2 μM) were incubated with the duplex DNA for 60 minutes at 0 °C in 20% ethylene glycol, 100 μM calf thymus DNA, 100 mM Tris pH 7.0, 100 mM NaCl, and 2 mM $CuCl_2$, than equilibrated for 15 minutes at the temperature specified. DMS footprinting was effected by addition of 4 μl of a freshly prepared 2% aqueous solution of dimethyl sulfate. After 90, 45 or 30 seconds at the temperature specified, reactions were quenched with 5 μl of a DMS stop solution and 785 μl of ethanol.[32,33] After washing with 70% ethanol, strand scission was obtained after incubation at 90 °C for 30 minutes in 10% aqueous piperidine. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G + A sequencing lane. Lanes 3-20: 20 % ethylene glycol, 100 μM calf thymus DNA, 100 mM Tris pH 7.0, 100 mM NaCl, 2 mM $CuCl_2$, and 2 μM oligonucleotide. Lanes 3-8: 0°C, Lanes 9-14: 12.5 °C, Lanes 15-20: 25°C. Lanes 3, 9, 15: no oligonucleotide; Lanes 4, 10, 16: Oligonucleotide CT-15. Lanes 5, 11, 17: Oligonucleotide CT-13; Lanes 6,12,18: Oligonucleotide CT-11; Lanes 7, 13, 19: Oligonucleotide CT-09; Lanes 8, 14, 20: Oligonucleotide CT-07.

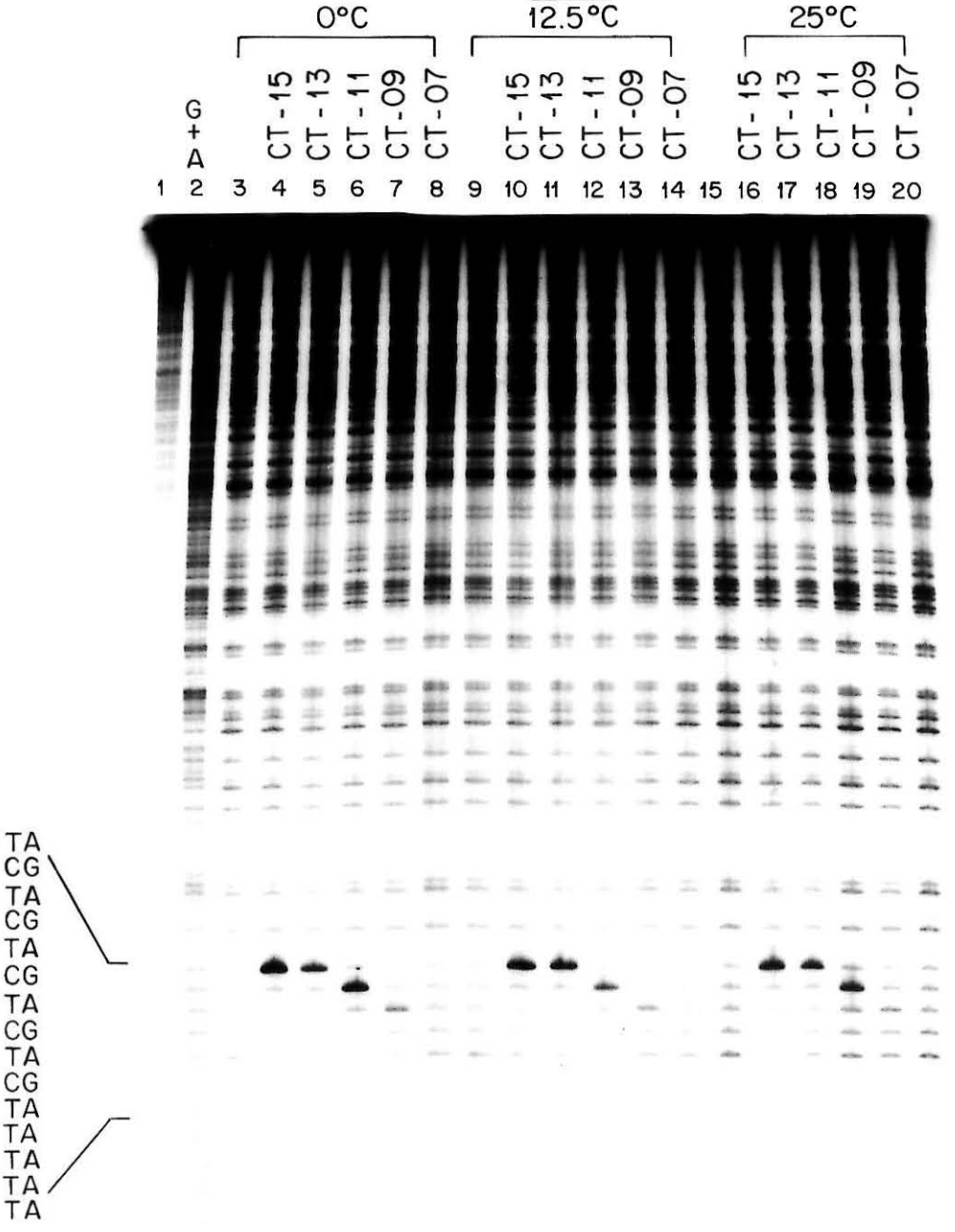

0°C
12.5°C
25°C
G
+
A
CT-15
CT-13
CT-11
CT-09
CT-07
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
TA
CG
TA
CG
TA
CG
TA
CG
TA
CG
TA
TA
TA
TA
TA

Figure 2.15. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMAG10 (EcoR I/Bgl I) indicating DMS footprinting of oligonucleotides CT-12, CT-10, CT-08, T-G mismatch, and C-A mismatch. Oligonucleotides (final concentration 2 μM) were incubated with the duplex DNA for 60 minutes at 0 °C in 20% ethylene glycol, 100 μM calf thymus DNA, 100 mM Tris pH 7.0, 100 mM NaCl, and 2 mM $CuCl_2$, then equilibrated for 15 minutes at the temperature specified. DMS footprinting was effected by addition of 4 μl of a freshly prepared 2% aqueous solution of dimethyl sulfate. After 90 seconds at 0 °C, reactions were quenched with 5 μl of a DMS stop solution and 785 μl of ethanol.[32,33] After washing with 70% ethanol, strand scission was obtained by incubation at 90 °C for 30 minutes in 10% aqueous piperidine. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G + A sequencing lane. Lanes 3-8: 0°C, Lanes 9-14: 12.5 °C, Lanes 15-20: 25°C. Lanes 3, 9, 15: no oligonucleotide; Lanes 4, 10, 16: Oligonucleotide CT-12. Lanes 5, 11, 17: Oligonucleotide CT-10; Lanes 6,12,18: Oligonucleotide CT-08; Lanes 7, 13, 19: TG mismatch oligonucleotide; Lanes 8, 14, 20: CA mismatch oligonucleotide.

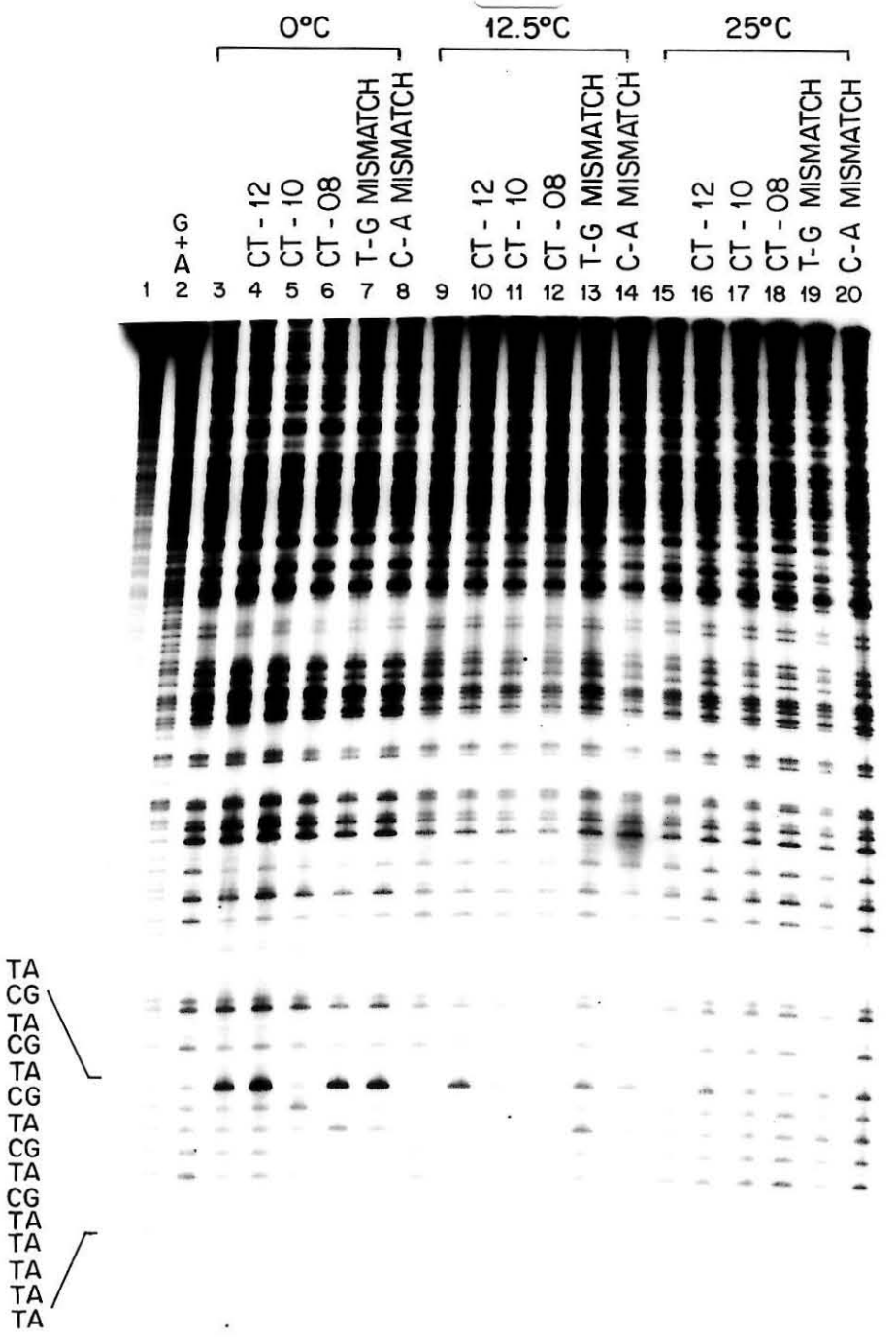
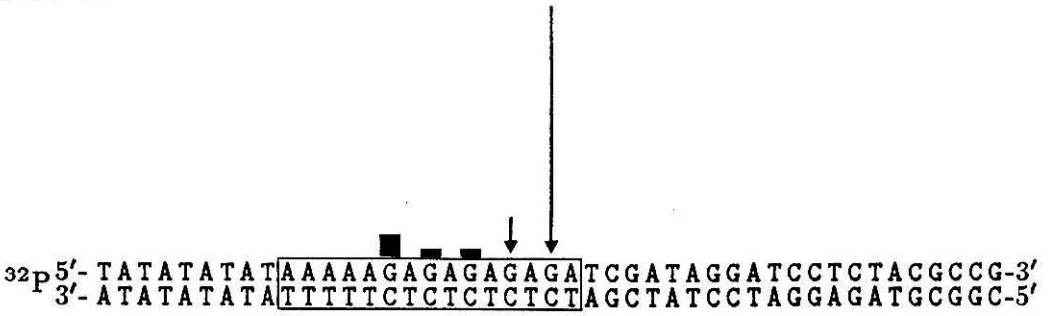
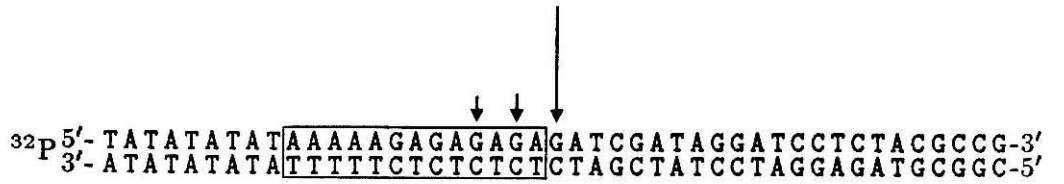

0°C
12.5°C
25°C
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
G
A+G
CT - 12
CT - 10
CT - 08
T-G MISMATCH
C-A MISMATCH
CT - 12
CT - 10
CT - 08
T-G MISMATCH
C-A MISMATCH
CT - 12
CT - 10
CT - 08
T-G MISMATCH
C-A MISMATCH
TA
CG
TA
CG
TA
CG
TA
CG
TA
CG
TA
TA
TA
TA
TA

Fig. 2.16. Histograms obtained from densitometric scanning of autoradiogram from Figures 2.14 and 2.15, indicating the binding of oligonucleotides as determined by DMS footprinting. Bars indicate extent of inhibition of cleavage at base indicated. Boxes indicate oligonucleotide binding site.

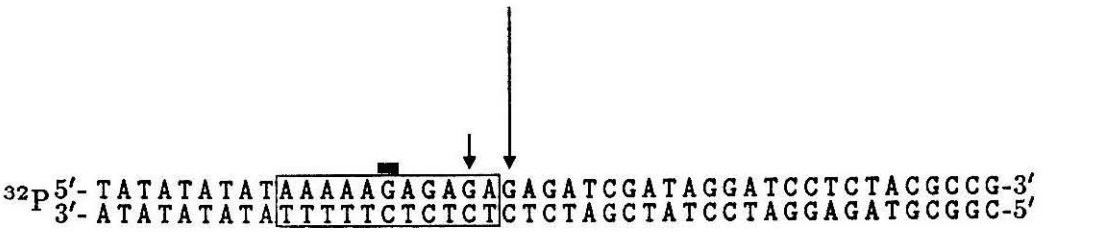
Probe CT-15



Probe CT-13



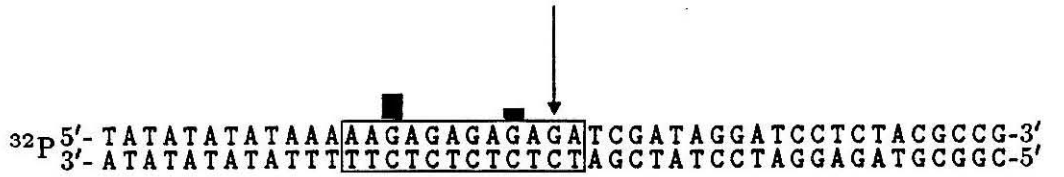
Probe CT-11



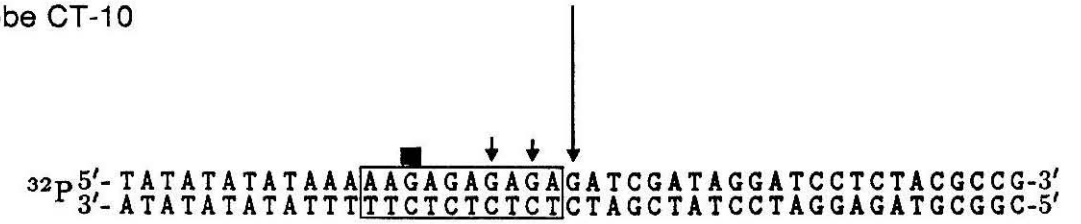
Probe CT-09

Probe CT-07

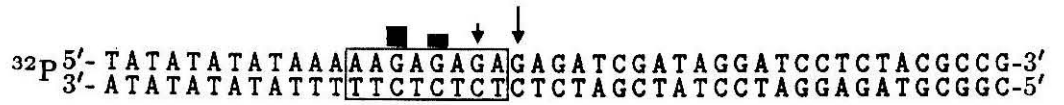
Probe CT-12



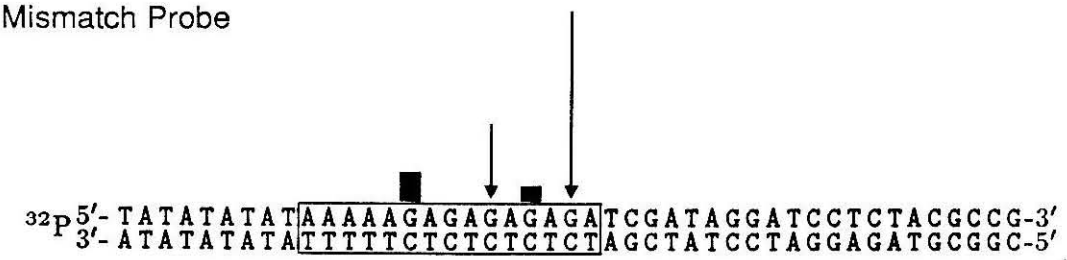
Probe CT-10



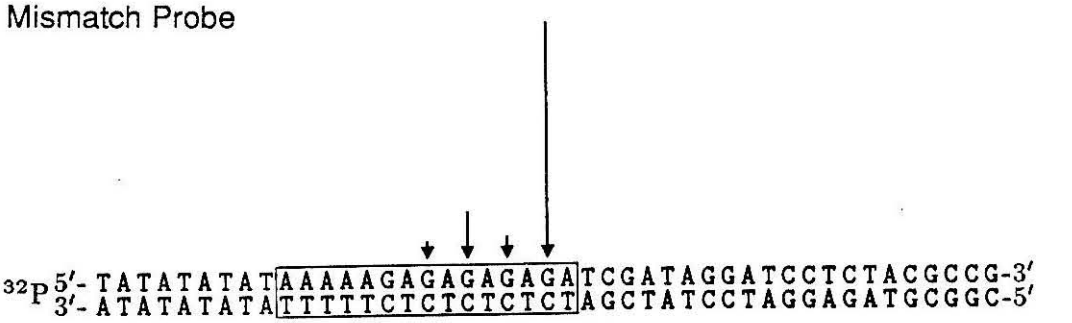
Probe CT-08



T-G Mismatch Probe



C-A Mismatch Probe



located within the binding site, indicating that the two 3' terminal bases of oligonucleotides CT-15 and CT-12 are not bound in the triple helical complex.

Affinity cleaving results indicate that oligonucleotide CT-13-EDTA•Fe(II) cleaves its target site as efficiently as CT-15-EDTA•Fe(II).[1] These data corroborate conclusions drawn from DMS footprinting studies concerning the unbound state of the last two base pairs of the oligonucleotide CT-15.

It is now known that the terminal adenine within the oligonucleotide binding site is a substrate for methylation by the methylase dam I. Under the conditions used to prepare the plasmid $pDMAG_{10}$, 100% methylation at N-6 of this residue is expected. This fact has been confirmed by restriction enzyme inhibition studies.[36] Methylation at this site interferes with Hoogsteen hydrogen bond formation, preventing binding of the final (two) bases of the oligonucleotide to the target site (Fig. 2.18).

DNase I footprinting indicates that the oligonucleotides C-A mismatch and CT-11 bind with approximately equal affinities, while the binding affinity of the T-G mismatch oligonucleotide is intermediate between those of the oligonucleotides CT-09 and CT-07. DMS footprinting indicates that both of these oligonucleotides bind only at 0 °C. The binding affinities of these oligonucleotides as assayed by DMS footprinting are thus lower than that observed for the oligonucleotide CT-09. Binding of these oligonucleotides results in the appearance of two guanines hyperreactive to dimethyl sulfate(Figures 2.15, 2.16), indicating that these oligonucleotides can bind by accomodation of the mismatch via tautomerism of one of the bases, or by forming nine (T-G mismatch oligonucleotide) or ten (C-A mismatched oligonucleotide) matched base triplets at the 5' end of the binding site.

Affinity cleaving data indicate that these mismatched oligonucleotides are thought to bind to the target sequence approximately ten times less tightly

than their fully complementary counterpart, with the C-A mismatch oligonucleotide binding with greater affinity.[1]

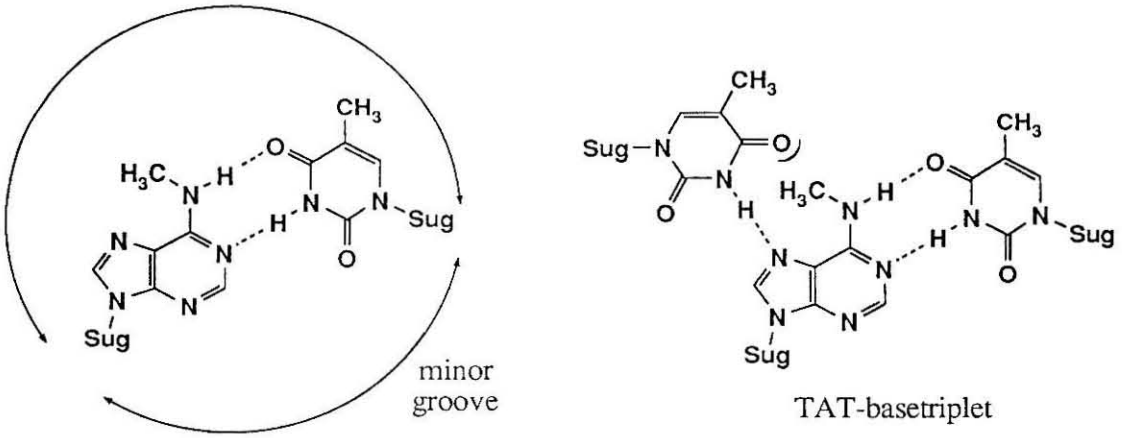


Fig. 2.17. Methylation at N-6 of terminal 3' adenine in triple helix target site interferes with formation of T•AT base triplets.

One explanation for the difference in binding affinities between C-A and T-G mismatched oligonucleotides is that binding of the C-A mismatch allows accomodation of ten matched triplets, while the T-G mismatch oligonucleotide can only bind nine matched triplets. The relative intensities of the two hyperreactive guanines, however, indicates that accomodation of the mismatch is favored over binding of only the first nine or ten base pairs. The higher binding affinity of C-A mismatch may therefore indicate that the C-A mismatch is more easily accomodated than is the T-G mismatch.

Two-dimensional models of the base triplets in these studies indicate that a T-G mismatch can be incorporated within a triplex with the formation of a hydrogen bond (Fig. 2.18). Incorporation of a C-A mismatch requires protonation of cytosine for the formation of a single hydrogen bond and

results in a potential steric clash between the exocyclic amino groups at C-6 of adenine and C-4 of cytosine.

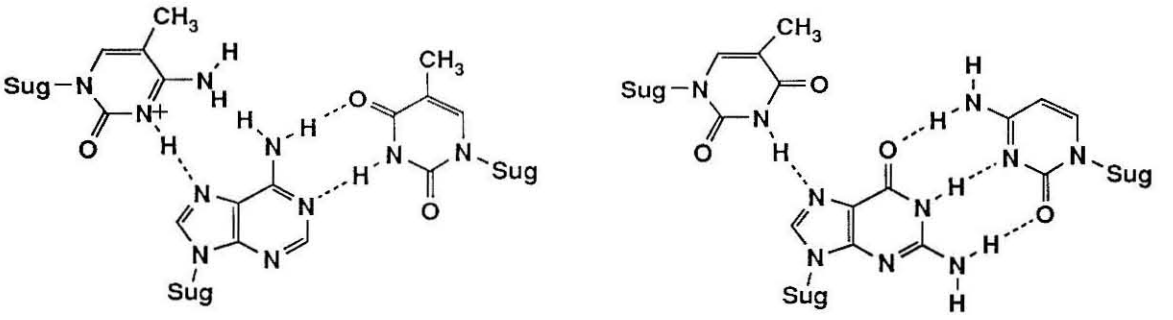


Fig. 2.18. Base triplet mismatches formed by hybridization of the oligonucleotides C-A mismatch and T-G mismatch with triple helical target site.

DMS footprinting of oligonucleotide binding to the sequence A_5G_{10}. DMS footprinting was used to study triple helix formation between the oligonucleotides CC-15, CC-13, CC-11 and the target sequence A_5G_{10} (Fig. 2.19).

```
                    5'- T T T T T C C C C C C C C C -3'      CC-15
                    5'- T T T T T C C C C C C C C   -3'      CC-13
                    5'- T T T T T C C C C C C       -3'      CC-11
5'.... T A T A T A T A T A T [A A A A A G G G G G G G G] T C G A T A G ....3'
3'.... A T A T A T A T A T A [T T T T T C C C C C C C C] A G C T A T C ....5'

                    [ ] = triplex site
```

Fig. 2.19. Oligonucleotides used to study binding to the sequence A_5G_{10} via DMS footprinting. No dam I methylase site exists in this sequence.

Binding of this series of oligonucleotides is extremely pH dependent, with no binding observed at pHs > 3.7 (Fig. 2.20, 2.21). The optimal pH for binding should decrease as the cytosine content of the oligonucleotide

Figure 2.20. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMG10 (EcoR I/Bgl I) indicating DMS footprinting of oligonucleotide CC-13 as a function of pH. Oligonucleotides (final concentration 2 μM) were incubated with the duplex DNA for 60 minutes at 0 °C in 20% ethylene glycol, 100 μM calf thymus DNA, 100 mM Tris, 100 mM NaCl, and 1 mM $Co(NH_3)_6^{+3}$. DMS footprinting was effected by addition of 4 μl of a freshly prepared 2% aqueous solution of dimethyl sulfate. After 90 seconds at 0 °C, reactions were quenched with 5 μl of a DMS stop solution and 785 μl of ethanol.[32,33] After washing with 70% ethanol, strand scission was obtained by incubation at 90 °C for 30 minutes in 10% aqueous piperidine. Lane 1: Intact DNA. Lane 2: Maxam-Gilbert G + A sequencing lane. Lanes 3, 5, 7, 9, 11 and 13: no oligonucleotide. Lane 4, 6, 8, 10, 12, and 14: 2 μM oligonucleotide CC-13. Lanes 3 and 4: pH 2.7; Lanes 5 and 6: pH 3.7; Lanes 7 and 8: pH 4.7; Lanes 9 and 10: pH 5.7; Lanes 11 and 12: pH 7.0; Lanes 13 and 14: pH 8.0.

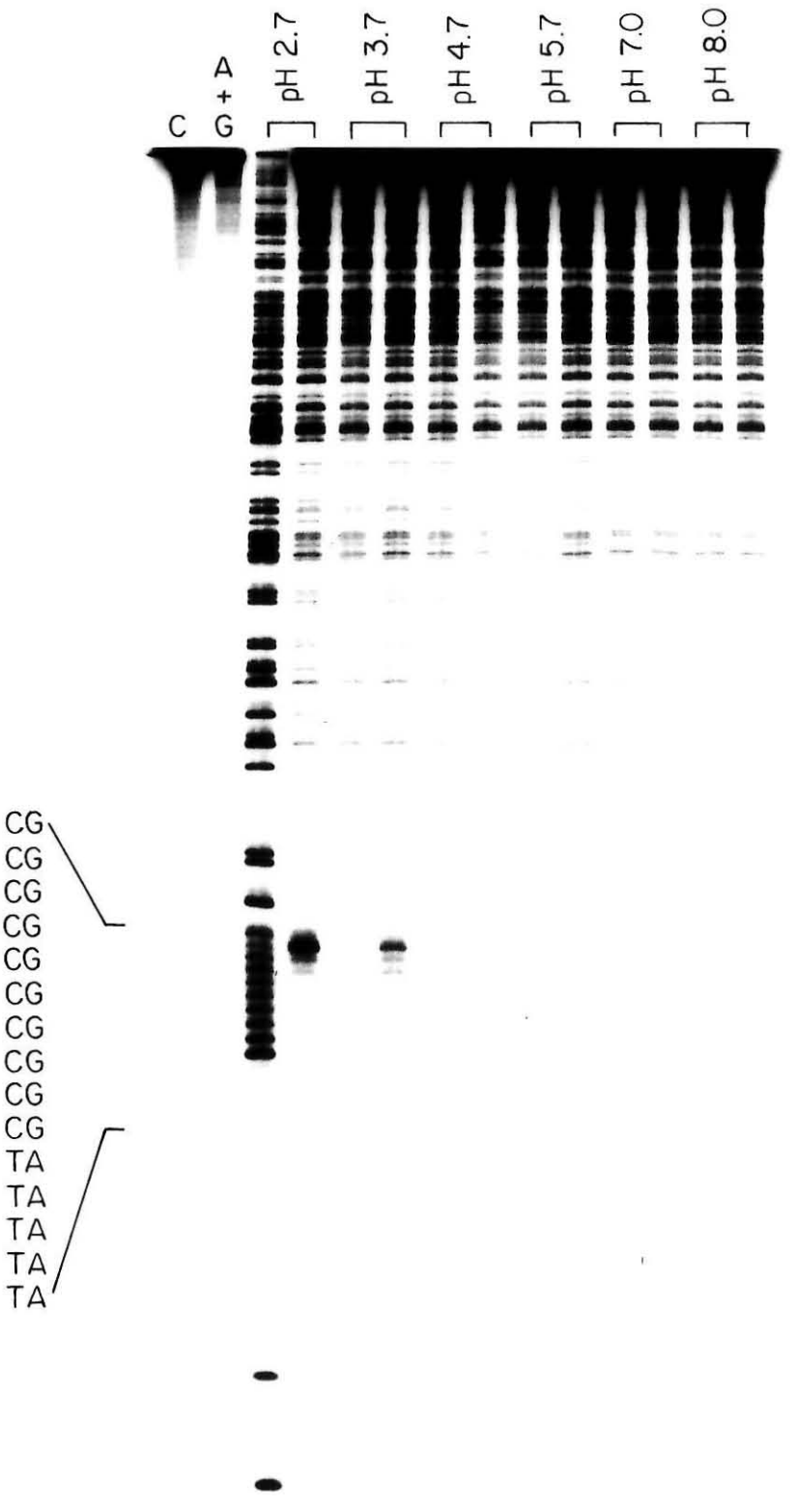
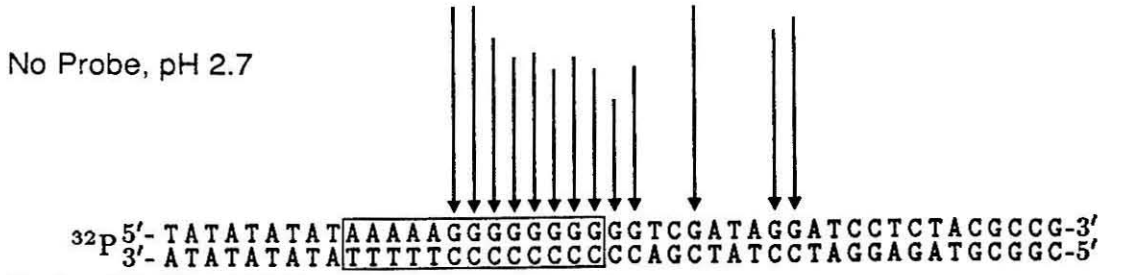
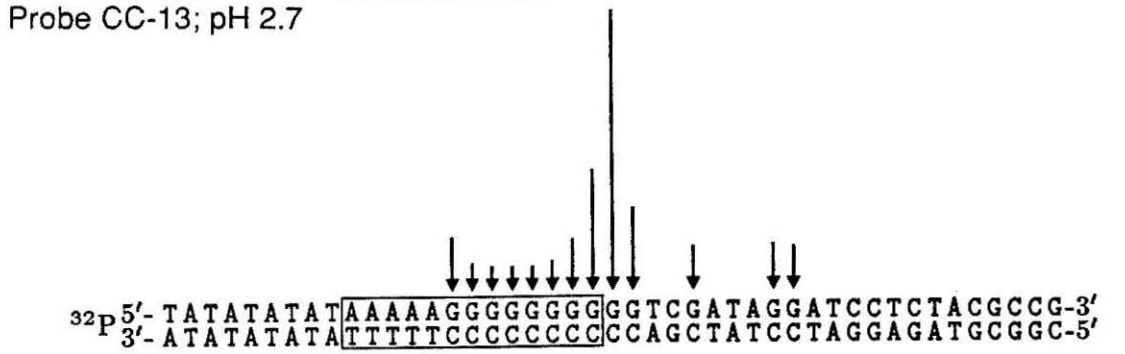

C
A + G
pH 2.7
pH 3.7
pH 4.7
pH 5.7
pH 7.0
pH 8.0
CG
CG
CG
CG
CG
CG
CG
CG
CG
CG
TA
TA
TA
TA
TA

Fig. 2.21. Relative reactivities of guanines in plasmid pDMG10 to dimethyl sulfate as a function of pH and oligonucleotide CC-13 as determined by densitometry of autoradiogram from Fig. 2.20. Arrows indicate extent of cleavage at base indicated. Boxes indicate oligonucleotide binding site.

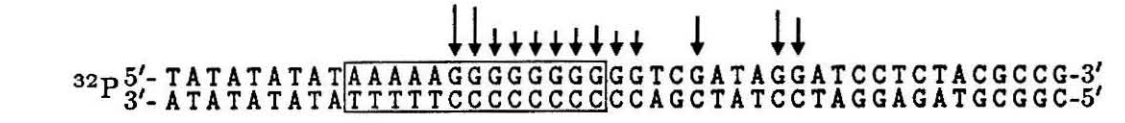
No Probe, pH 2.7



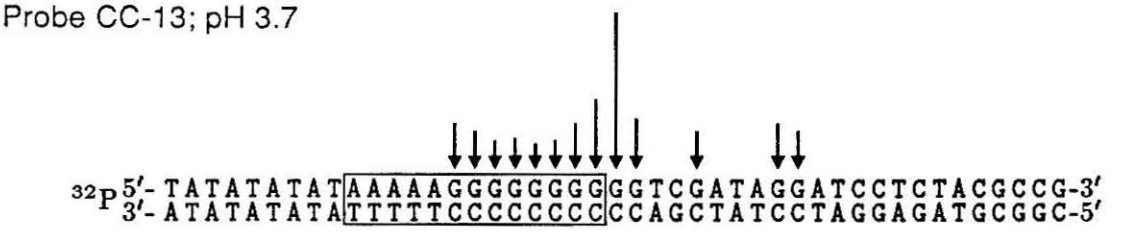
Probe CC-13; pH 2.7



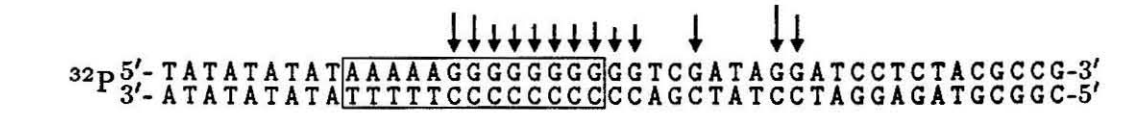
No Probe; pH 3.7



Probe CC-13; pH 3.7



No Probe; pH 4.7



Probe CC-13; pH 4.7

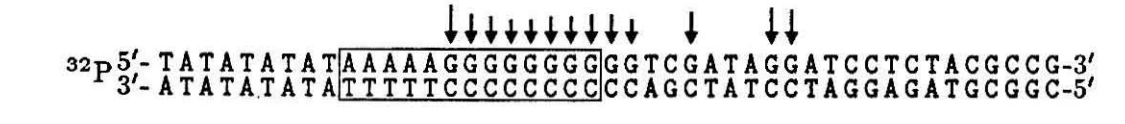
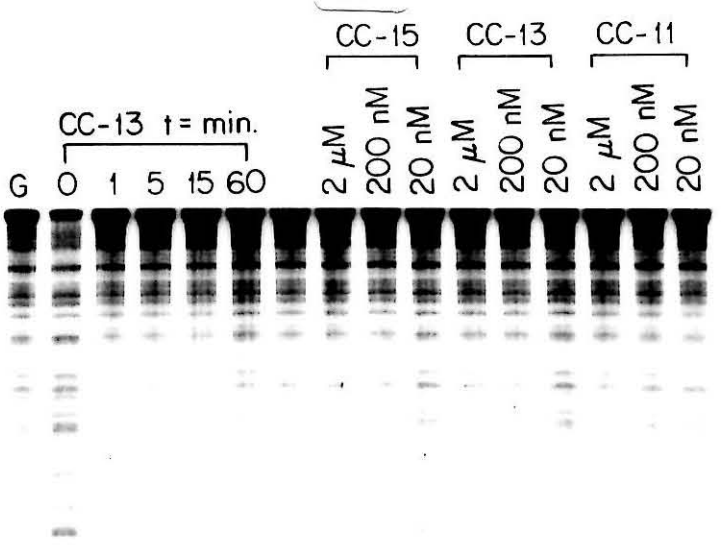


Figure 2.22. Autoradiogram of 8% polyacrylamide high resolution denaturing gel electrophoresis of ^{32}P-endlabeled 628 bp restriction fragment from pDMG10 (EcoR I/Bgl I) indicating DMS footprinting of oligonucleotides CC-15, CC-13 and CC-11 as a function of time of incubation and oligonucleotide concentration. Oligonucleotides were incubated with the duplex DNA at 0 °C in 20% ethylene glycol, 100 μM calf thymus DNA, 100 mM Tris, 100 mM NaCl, and 1 mM $Co(NH_3)_6^{+3}$. DMS footprinting was effected by addition of 4 μl of a freshly prepared 2% aqueous solution of dimethyl sulfate. After 90 seconds at 0 °C, reactions were quenched with 5 μl of a DMS stop solution and 785 μl of ethanol.[32,33] After washing with 70% ethanol, strand scission was obtained by incubation at 90 °C for 30 minutes in 10% aqueous piperidine.

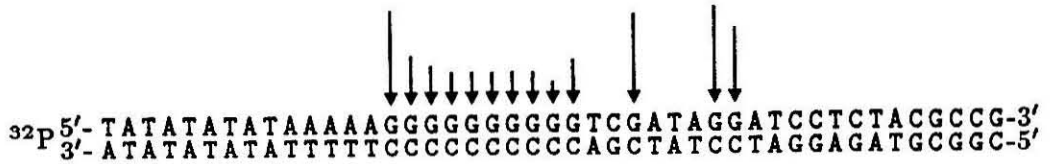
Lane 1: G reaction is absence of oligonucleotide. Lane 2-6: 2 μM oligonucleotide CC-13. Lane 2: no pre-incubation. Lane 3: DMS reaction 1 minute after addition of oligonucleotide. Lane 4: DMS reaction 5 minutes after addition of oligonucleotide. Lane 5: DMS reaction 15 minutes after addition of oligonucleotide. Lane 6: DMS reaction 60 minutes after addition of oligonucleotide. Lanes 7-16: DMS reaction 60 minutes after addition of oligonucleotide. Lane 7: no addition of oligonucleotide. Lane 8: addition of 2 μM oligonucleotide CC-15. Lane 9: addition of 200 nM oligonucleotide CC-15. Lane 10: addition of 20 nM oligonucleotide CC-15. Lane 11: addition of 2 μM oligonucleotide CC-13. Lane 12: addition of 200 nM oligonucleotide CC-13. Lane 13: addition of 20 nM oligonucleotide CC-13. Lane 14: addition of 2 μM oligonucleotide CC-11. Lane 15: addition of 200 nM oligonucleotide CC-11. Lane 16: addition of 20 nM oligonucleotide CC-11.


CC-15
CC-13
CC-11
CC-13 t = min.
G
0
1
5
15
60
2 μM
200 nM
20 nM
2 μM
200 nM
20 nM
2 μM
200 nM
20 nM

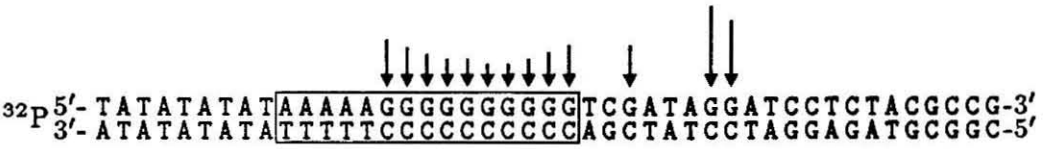

CG
CG
CG
CG
CG
CG
CG
CG
CG
CG
TA
TA
TA
TA
TA

Fig. 2.22. Relative reactivities of guanines in plasmid pDMG10 to dimethyl sulfate as a function of oligonucleotide and oligonucleotide concentration as determined by densitometry of autoradiogram in Fig. 2.21. Bars indicate extent of inhibition of cleavage at base indicated. Boxes indicate oligonucleotide binding site.

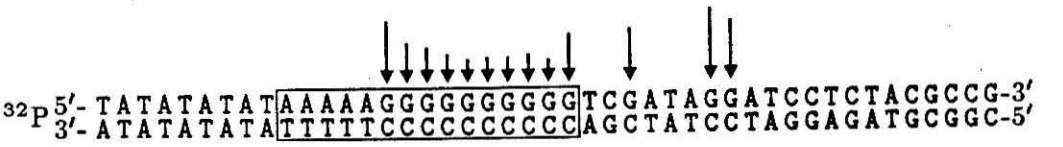
No Probe



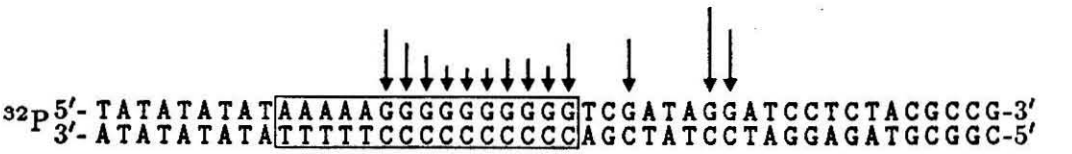
Probe CC-15; 2 μM



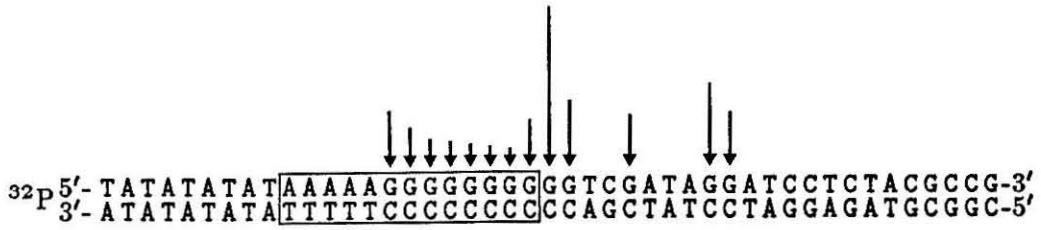
Probe CC-15; 200 nM



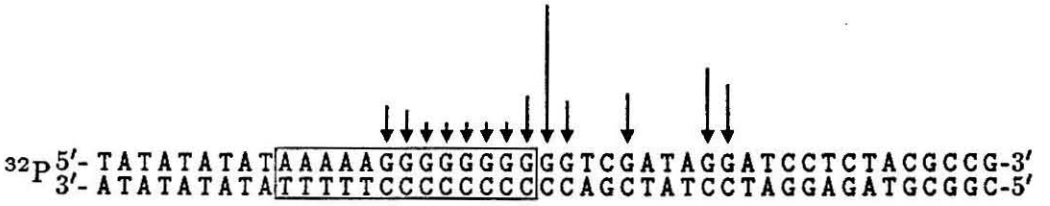
Probe CC-15: 20 nM



Probe CC-13; 2 μM

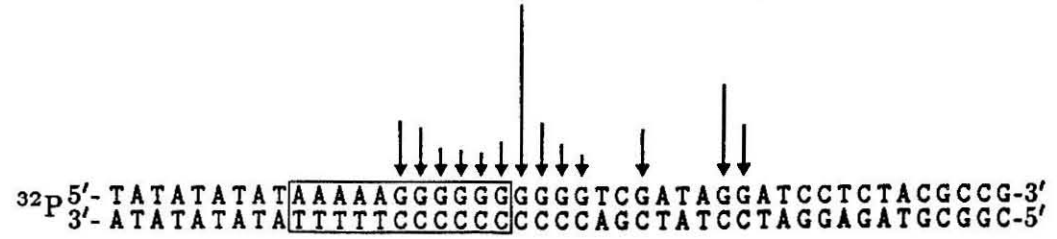


Probe CC-13; 200 nM

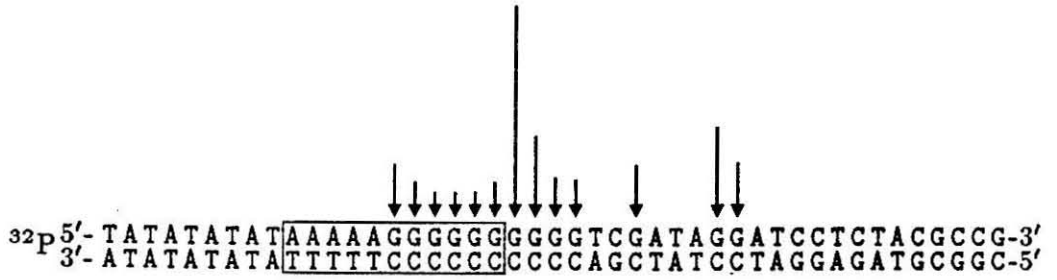


Probe CC-13: 20 nM

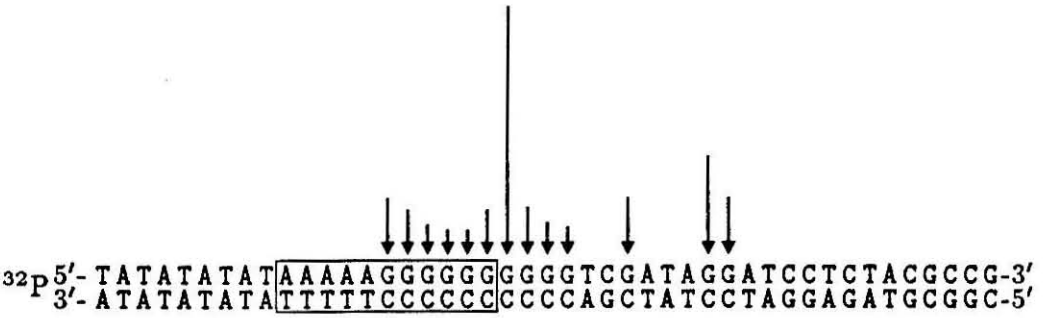
Probe CC-11; 2 μM



Probe CC-11; 200 nM



Probe CC-11; 20 nM



increases, due to the presence on increasing number of protonation sites within the triple helical complex.[59,60] The binding of a series of cytosines may result in further destabilization of the triplex, due to the stacking of the positive charges within the complex.

The kinetics of hybridization between oligonucleotides CC-13 and CT-13 and their target sequences differ. A noticeable increase in hyperreactivity is observed if the oligonucleotide CC-13 is allowed to equilibrate with its target site for at least 60 minutes prior to reaction, however, prior to that time, no change in guanine reactivity to DMS is observed (Fig. 2.22). In contrast, oligonucleotide CT-15 (2 μM) hybridizes to its target sequence within one minute, as determined by a similar assay (*vide supra*). Binding of oligonucleotide CC-13 is monitored at pH 2.7. It is possible that this change in hybridization time is a function of pH, and not of the oligonucleotide under study.

No hyperreactivity was conferred upon hybridization of the 15 base oligonucleotide CC-15 (Fig. 2.22). Hybridization of the oligonucleotides CC-13 and CC-11 results in hyperreactivity at guanines located 1 base 3' to the binding site for that oligonucleotide (Fig. 2.23). Under slightly different conditions (40% ethanol, 1 mM spermine, pH 5.7), the oligonucleotide CC-15-EDTA•Fe(II) was shown to cleave DNA when activated with Fe(II) and DTT; thus, while oligonucleotide CC-15 binds its target sequence, no hyperreactivity to any of the guanines in the target sequence is conferred. This indicates that the oligonucleotide CC-15 binds all 15 base pairs of the target sequence, in contrast to the oligonucleotide CT-15. This is now understood to be a result of methylation by dam I of the plasmid pDMAG10, but not pDMG10.[36]

Hyperreactivity was largely limited to a single guanine. This indicates that the structural distortion associated with a transition from triplex to

duplex DNA is limited to a single base pair. Such a distinction could not be made based on studies with the pDMAG10 plasmid, since guanines are only present at every second base pair.

The binding affinities of the oligonucleotides CC-13 and CC-11 can be analyzed due to the presence of a hyperreactive guanine. The affinities of these oligonucleotides at pH 2.7 are $> 5 \times 10^8\ M^{-1}$, indicating the stabilization of C+G•C base triplets at low pH.[59,60]

Conclusions

The binding of oligonucleotides to double helical DNA via triple helix formation can be detected using a variety of footprinting agents. This allows the determination of oligonucleotide affinity without the incorporation of modified nucleosides which may alter binding characteristics under a variety of conditions.

MPE•Fe(II) can be used to determine binding site size, and indicates that oligonucleotide binding is solvent, cation, and oligonucleotide concentration dependent. Triple helices are stable under conditions of optimal DNase I activity (10 mM in each Mg^{+2} and Ca^{+2}). This study indicates that triplexes can be used to inhibit enzyme activity under physiological conditions. Oligonucleotide binding results in a conformational change in the DNA, which can be detected by dimethyl sulfate (DMS). Guanines located 1 base pair beyond the oligonucleotide binding site become hyper-reactive to DMS upon triplex formation. This technique indicates that the two terminal 3' bases in the oligonucleotide CT-15 do not bind their target sequence. DMS footprinting can be used to determine triplex formation under a variety of conditions (low pH, presence of various cations) which are not compatible

with **MPE**•*Fe(II)* or DNase I activity. Due to the short reaction time required by DMS (90 sec.), this technique can be used to study the kinetics of triplex formation. DNase I and DMS indicate that oligonucleotide CT-15 has a binding affinity of approximately $1 \times 10^7\ M^{-1}$. These footprinting techniques are used to compare the binding affinities of oligonucleotides of various lengths and base composition. Oligonucleotides with higher C/T ratios bind with lower affinities at neutral pH. Oligonucleotides shorter than 11 base pairs bind with weak affinities. The binding of oligonucleotides containing long runs C's (T_5G_{10}) bind with slower kinetics and only at very low pH.

Materials and Methods

Reagents. EcoR I, Calf Thymus Alkaline Phosphatase, T4 Polynucleotide Kinase, and Bgl I were purchased from Boehringer Mannheim. Klenow fragment of DNA polymerase I was obtained from New England Biolabs. Ethylene glycol (Mallinckrodt), spermine•4 HCl (Aldrich, 98% pure), $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ (Baker Analyzed reagent), and dithiothreitol (DTT) (Bethesda Research Laboratories, ultra pure), were used as obtained. Phenol was distilled and stored under buffer at -4 °C for short periods of time. Piperidine (Fluka puriss. p. a.) was distilled and stored at -20 °C. A 1 mM solution of deproteinized calf thymus DNA was prepared by ultrasonic treatment of purchased calf thymus DNA (Sigma), followed by repeated phenolic extraction and lengthy dialysis until a A_{260}/A_{280} of at least 1.8 was obtained. Aqueous solutions of DTT and $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ were prepared just prior to use. General experimental details for working with DNA were taken from Sambrook, Frisch, and Maniatis.[61]

Plasmids. Plasmids pDMAG10 and pDMG10 were obtained from Dave Mendel.[35,58] Large scale preparation was accomplished as described[35,58,61] using the *E. Coli* strain HB101, which contains dam I methylase activity. Plasmids were purified by density gradient centrifugation using a CsCl gradient.[61]

Oligonucleotide synthesis and purification. The synthesis of all oligonucleotides was accomplished on a 1 μmole scale using a Beckman System 1 Plus DNA-synthesizer and the automated solid-phase phosphoramidite method. After brief drying in vacuo, the polymer-bound oligonucleotide was treated with a 1:2:2 solution of thiophenol:dioxane:triethylamine (1.5 ml, 25 °C, 24 hrs.) to demethylate the phosphotriester. The polymer was washed with methanol (1.5 ml, 3x) and ether (1.5 ml 3x). The support was air dried for one hour. Removal of the oligonucleotide from the support was effected using concentrated NH_4OH (1.5 ml, 2 hrs.). A second treatment with NH_4OH (1.5 ml, 2 hrs.) insured that removal was complete. The combined supernatants were reduced to dryness with a stream of argon, and deprotection of the bases effected using concentrated NH_4OH (1.5 ml, 55 °C, 24 hrs.).

The oligonucleotides were purified by gel electrophoresis or liquid chromatography. Purification by electrophoresis was accomplished using 20% polyacrylamide (40x20x0.2 cm, 1:20 cross-linked, 40% urea, TBE buffer; electrophoresis for 16 hrs. at 750 V). Bands in the gel were visualized with UV light and excised, crushed, and eluted with 0.2 M NaCl, 1 mM EDTA (10 ml) for 40 hrs. at 37 °C. The polyacrylamide was removed using a cellulose acetate (Centrex) filter (pore-size 0.45 μm) and the resulting filtrate was dialyzed extensively (five days) against water. A Pharmacia Mono Q 5/5 anion exchange column was used to effect chromatographic separation with

the following solvent system: solvent A: 20 mM Tris, pH 8.3; solvent B: 20 mM Tris, pH 8.3, 1 M KCl. A gradient of 10-70% B over a period of 60 minutes was employed. Collections from the FPLC purification were dialyzed as described. The concentrations of the single stranded oligodeoxynucleotides were determined with the use of the extinction values (260 nm) for the bases T (8700) and C (7300). The oligonucleotides were lyophilized and stored dry at -20 °C.

Radio-labeling of restriction fragment. A uniquely 5'-end labeled 628 bp restriction fragment was obtained by linearization of 15 μgs of the plasmid pDMAG10 or pDMG10 with 80 units of EcoR I in 100 μl, dilution of the reaction with 98 μl of 1x CAP buffer, and treatment with Calf Alkaline Phosphatase (2 units, 1 unit/μl). After phenol extraction (2x), ether extraction (3x), and precipitation (NaOAc, ethanol), labeling with T4 Polynucleotide kinase (20 units) and γ-^{32}P-dATP was achieved as directed.[61] Unincorporated nucleotides were removed by precipitation with NH_4OAc (2x). Cleavage with Bgl I afforded two labeled bands, 628 and 934 bp in length, which were separated on a non-denaturing 5% polyacrylamide gel (150 x 140 x 2 mm, 1:20 cross-linked, TBE buffer) for two hours at 240 V. The radio-labeled DNA was visualized by autoradiography (2 min. exposure), the faster running band was excised from the gel and eluted from the gel slice with 0.2 M NaCl, 1 mM EDTA at 37 °C. The gel slices were removed by filtration through a 0.45 μm Centrex filter, and the filtrate reduced in volume by butanol extraction. The DNA was precipitated twice, washed with 70% ethanol, and dissolved in 10 mM Tris, pH 7.4 to make a solution with approximately 10,000 cpm/μl.

Labeling of the 3'-end was accomplished using EcoR I linearized DNA, phenol extraction and precipitation of the DNA, radio-labeling of the DNA using the Klenow fragment of DNA polymerase I and α-^{32}P dATP. After

labeling, the DNA was purified, cleaved with Bgl I, and the restriction fragment isolated as described.

Sequencing was carried out as described by Maxam and Gilbert[32,61] with the following modifications: (i) G reactions with dimethyl sulfate were carried out at 0 °C for 1 min., (ii) treatment with formic acid was done at 0 °C for four minutes to afford a G+A reaction, and (iii) piperidine work-up was carried out using 0.1 M piperidine at 90 °C for 30 minutes.

MPE*•*Fe(II) footprinting. A 4x stock solution containing 400 μM Calf thymus DNA, 400 mM NaCl, 5' or 3' ^{32}P end-labeled DNA (approximately 40,000 cpm per lane), and 100 mM Tris/acetate buffer (pH 7.0) was prepared just prior to use. The DNA oligonucleotide, ethylene glycol and cations were mixed in an Eppendorf tube and the stock solution above added and allowed to equilibrate with the oligonucleotide at 0 °C for at least 60 min. **MPE**•Fe(II) was made as a 10x solution by adding freshly prepared $Fe(NH_4)_2(SO_4)_2 \bullet 6H_2O$ solution to make a solution of 40 μM in Fe(II) and 20 μM in **MPE**. **MPE**•Fe(II) was allowed to equilibrate with the DNA for 15 min. prior to activation with the addition of DTT (2 mM final concentration). Final reaction volume was 20 μl. The reactions were allowed to run for 14 hours at 0 °C and were stopped by precipitation with NaOAc and ethanol and washing with 70% ethanol. The pellet was dissolved in formamide buffer, and loaded directly onto an 8% polyacrylamide sequencing gel for analysis.

Deoxyribonuclease I Footprinting. Footprinting studies using Deoxyribonuclease I (DNase I) were conducted in much the same manner as the **MPE**•Fe(II) footprinting studies. After the DNA was mixed with the proper combination of Tris and cations, the oligonucleotide was equilibrated with the restriction fragment at 0 °C for at least 60 min. prior to addition of DNase I. Since the equilibration step is slow relative to the footprinting

reaction, the final buffer concentrations listed in figure legends represents the content of the equilibration buffer. Reactions were effected with the addition of 2 µl of a 400 ng/ml solution of DNase I (final concentration 40 ng/ml), The reactions were run at 4 °C for 12 minutes, and were stopped by precipitation of the DNA with the addition of 5 µl of a 10 M NH_4OAc, 200 mM EDTA solution and 75 µl of ethanol. The pellet was washed and loaded onto a 8% polyacrylamide sequencing gel using formamide loading buffer for analysis. Studies at higher temperatures were conducted by allowing the oligonucleotide to equilibrate with the target DNA at 0 °C for one hour, then transferring the reactions to an equilibrated temperature bath for 15 min.

DMS Protection Experiments. Solutions were made as described above for DNase I footprinting in a final volume of 20 µl. The oligonucleotide was equilibrated with the restriction fragment for 60 min., after which 4 µl of a freshly prepared 2% aqueous solution of dimethyl sulfate was added. Reactions were conducted at 0 °C for 90 seconds and were terminated by the addition of 5 µl of DMS stop solution[32,61] and 785 µl of ethanol. The DNA was precipitated and washed (70% ethanol). The pellet was dissolved in 100 µl of 0.1 M piperidine, and incubated at 90 °C for 30 min. After lyophilization, the pellet was dissolved in 10 µl of formamide loading buffer.

Studies at higher temperatures were conducted by allowing the oligonucleotide to equilibrate with the target DNA at 0 °C for one hour, then transferring the reactions to an equilibrated temperature bath for 15 min. Reactions were run at the stated temperature. Reaction times were adjusted to allow for the greater rate of reaction at higher temperatures.

Polyacrylamide electrophoresis. Electrophoresis was accomplished using a denaturing 8% polyacrylamide gel (50% urea, 1:20 cross-linked) at 1100 V for approximately four hours, or until the xylene cyanol migrated 20 cm.

Autoradiography was performed using Kodak X-Omat AR film with the aid of an intensifying screen. Densitometric traces were obtained from high resolution autoradiographs obtained without the use of the intensifying screen using a LKB Bromma Ultrascan XL Laser Densitometer.

References

(1) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.

(2) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968-972.

(3) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-970.

(4) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435-2437.

(5) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* submitted.

(6) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.

(7) Griffin, L. C. Ph. D. Thesis, California Institute of Technology, 1990.

(8) Griffin, J. H. Ph. D. Thesis, California Institute of Technology, 1990.

(9) Landolfi, N. F.; Yin, X.-M.; Capra, J. D.; Tucker, P. W. *BioTechniques* **1989**, *7*, 500-504.

(10) Galas, D. J.; Schmitz, A. *Nucl. Acids Res.* **1978**, *5*, 3157-3170.

(11) Van Dyke, M. W.; Dervan, P. B. *Cold Spring Harbor Sympos. on Quant. Biol.* **1983**, *47*, 347-353.

(12) Van Dyke, M. W.; Dervan, P. B. *Biochemistry* **1983**, *22*, 2373-2377.

(13) Van Dyke, M. W.; Dervan, P. B. *Nucl. Acids Res.* **1983**, *11*, 5555-5567.

(14) Van Dyke, M. W.; Dervan, P. B. *Science* **1984**, *225*, 1122-1127.

(15) Francois, J.-C.; Saison-Behmoaras, T.; Helene, C. *Nucl. Acids Res.* **1988**, *16*, 11431-11440.

(16) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Nucl. Acids Res.* **1989**, *18*, 157-161.

(17) Waring, M. J. *Biochem. J.* **1974**, *143*, 483-486.

(18) Le Pecq, J. B. *Methods Biochem. Anal.* **1971**, *20*, 41-86.

(19) Le Pecq, J. B.; Paoletti, C. *C. R. Acad. Sc. Paris* **1965**, *260*, 7033-7036.

(20) Lehrman, E. A.; Crothers, D. M. *Nucl. Acids Res.* **1977**, *4*, 1381-1392.

(21) Arnott, S.; Selsing, E. *J. Mol. Biol.* **1974**, *88*, 509-521.

(22) de los Santos, C.; Rosen, M.; Patel, D. *Biochemistry* **1989**, *28*, 7282-7289.

(23) Rajagopal, P.; Feigon, J. *Biochemistry* **1989**, *28*, 7859-7870.

(24) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucl. Acids Res.* **1979**, *6*, 3073-3091.

(25) Scaria, P. V.; Shafer, R. H. *J. Biol. Chem.* **1991**, *266*, 5417-5423.

(26) Herzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23*, 3934-3945.

(27) Murray, N. L.; Morgan, A. R. *Can. J. Biochem.* **1973**, *51*, 436-449.

(28) Maher, L. J., III; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-730.

(29) Maher, L. J., III; Dervan, P. B.; Wold, B. *Biochemistry* **1990**, *29*, 8820-8826.

(30) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172-174.

(31) Lawley, P. D.; Brookes, P. *Biochem. J.* **1963**, *89*, 127-138.

(32) Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 560-564.

(33) Maxam, A. M.; Gilbert, W. *Methods in Enzymology* **1980**, *65*, 499.

(34) Hoogsteen, K. *Acata Cryst.* **1959**, *12*, 822-823.

(35) Mendel, D.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 910-914.

(36) Singleton, S. A. Ph. D. Thesis, Cal. Inst. of Tech., in preparation.

(37) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

(38) Felsenfeld, G.; Huang, S. *Biochim. Biophys. Acta* **1960**, *37*, 425-433.

(39) Thomas, T. J.; Bloomfield, V. A. *Biopolymers* **1983**, *22*, 1097-1106.

(40) Sullivan, K. M.; Lilley, D. M. J. *J. Mol. Biol.* **1987**, *193*, 397-404.

(41) Wilson, R. W.; Bloomfield, V. A. *Biochemistry* **1979**, *18*, 2192-2196.

(42) Moser, H. "Oligonucleotide-Directed Cleavage of Double Helical DNA by Triple-Strand Formation," California Institute of Technology, 1988.

(43) Shirai, H.; Itoh, Y.; Kurose, A.; Hanabusa, K.; Abe, K.; Hojo, N. *Polymer J.* **1984**, *16*, 207-215.

(44) Banville, D. L.; Marzilli, L. G.; Wilson, W. D. *Biochemistry* **1986**, *25*, 7393-7401.

(45) Deb, K. K. *Spectroscopy Lett.* **1981**, *14*, 385-393.

(46) Forster, W.; Bauer, E.; Schutz, H.; Berg, H. *Biopolymers* **1979**, *18*, 625-661.

(47) Saenger, W. *Principles of Nucleic Acid Structure;* Springer-Verlag Inc.: New York, 1984.

(48) Dervan, P. B. *Science* **1986**, *232*, 464-471.

(49) Hanvey, J. C.; Klysik, J.; Wells, R. D. *J. Biol. Chem.* **1988**, *263*, 7386-7396.

(50) Pullman, B.; Claverie, P.; Caillet, J. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 1663-1669.

(51) Youngquist, R. S. Ph. D. Thesis, California Institute of Technology, 1988.

(52) Campbell, V. W.; Jackson, D. A. *J. Biol. Chem.* **1980**, *255*, 3726-3735.

(53) Suck, D.; Lahm, A.; Oefner, C. *Nature* **1988**, *332*, 464-468.

(54) Selsing, E.; Wells, R. D.; Early, T. A.; Kearns, D. R. *Nature* **1978**, *275*, 249-250.

(55) Lilley, D. M. J. *Proc. Natl. Acad. Sci. U.S.A.* **1980**, *77*, 6468-6472.

(56) Pulleyblank, D. E.; Haniford, D. B.; Morgan, A. R. *Cell* **1985**, *42*, 271-280.

(57) Yu, Y.-T.; Manley, J. L. *Cell* **1986**, *45*, 743-751.

(58) Mendel, D. Ph. D. Thesis, California Institute of Technology, 1990.

(59) Lyamichev, V. I.; Mirkin, S. M.; Kumarev, V. P.; Baranova, L. V.; Vologodskii, A. V.; Frank-Kamenetskii, M. D. *Nucl. Acids Res.* **1989**, *17*, 9417-9423.

(60) Lyamichev, V. I.; Mirkin, S. M.; Frank-Kamenetskii, M. D. *J. Biomolec. Struct. Dyn.* **1985**, *3*, 327-338.

(61) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N. Y., 1989.

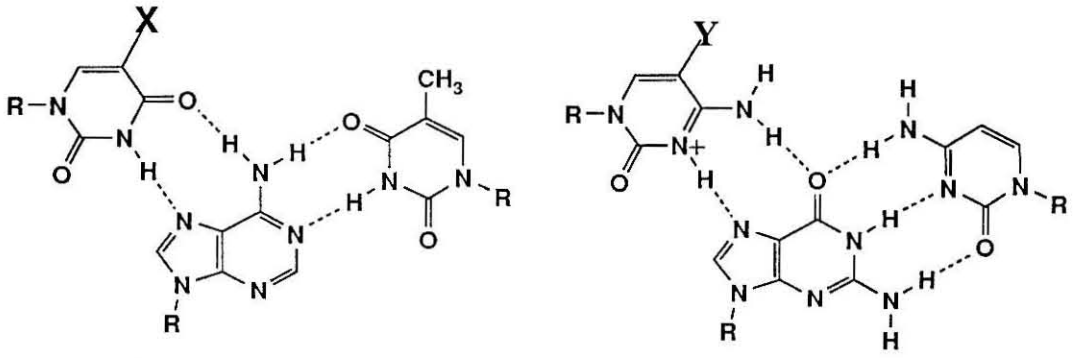
Chapter III

Use of Modified Pyrimidines to Enhance Oligonucleotide-Directed Triple Helix Stabilities

Introduction

In light of the limited specificity (4-8 bp) of restriction enzymes,[1,2] oligonucleotide directed triple helix formation offers a motif for the recognition of extended (≥ 15 base pairs) sequences of DNA.[3-5] Specificity is imparted by the formation of T•A-T and C+G-C Hoogsteen base triplets upon binding of the oligonucleotide in the major groove.[3,6-13] The 16 bp specificity possible with this motif is in principle 10^6 times greater than that attainable by restriction enzymes.

Complexes of triple helical nucleic acids containing cytosine (C) and thymine (T) in the Hoogsteen strand are stable in acidic to neutral solutions but dissociate on increasing pH.[3] Because the specificity offered by this motif could provide a method for the artificial repression of gene expression and viral diseases, it is important to understand structural factors which modulate triple helix stability. Changes in oligonucleotide structure could be used to control oligonucleotide affinity under *in vivo* conditions, where temporal and spatial intracellular pH (7.0 - 7.4) and ionic strength are strictly regulated and cannot be altered.[14-18]



X^5U-A-T basetriplet Y^5C+G-C basetriplet

Fig. 3.1. Isomorphous base triplets formed by incorporation of a third strand in the major groove of double helical DNA parallel to the Watson-Crick purine strand via Hoogsteen base pairing. Substituents at pyrimidine 5-positions protrude from the major groove, affecting affinity, but not complementarity.

Substitution at position 5 of pyrimidines alters the hydrophobic driving force, base stacking, and the electronic complementarity (hydrogen bond donating and/or accepting ability) of the Hoogsteen pyrimidine-purine base pairing for triple strand formation.[19-21] Incorporation of 5-substituted pyrimidines offers a method of modulating binding affinity without changing the hydrogen bonding pattern and sequence specificities of pyrimidine oligonucleotides (Fig. 3.1).

Figure 3.2: (Top) Oligonucleotides **1-6** constructed from deoxyribonucleotide phosphoramidites containing 2'-deoxycytidine (C), 5-methyl-2'-deoxycytidine (Me^5C), 2'-deoxyuridine (U), thymidine (T), 5-bromo-2'-deoxyuridine (Br^5U) and thymidine-EDTA (T^*) nucleotides. (Bottom) Coarse resolution pattern for cleavage of plasmid DNA (4.06 kbp) by oligonucleotide-EDTA **1-6** with simplified model of triple helical complex between the Hoogsteen-bound oligonucleotide (X^5U / Y^5C) and a 15 base pair target site.

1 5'-*T U U U U C U C U C U C U C T-3'

2 5'-*T U U U U C U C U C U C U C T-3' (Me)

3 5'-*T T T T T C T C T C T C T C T -3'

4 5'-*T T T T T C T C T C T C T C T-3' (Me)

5 5'-*T U U U U C U C U C U C U C T-3' (Br)

6 5'-*T U U U U C U C U C U C U C T-3' (Br, Me)

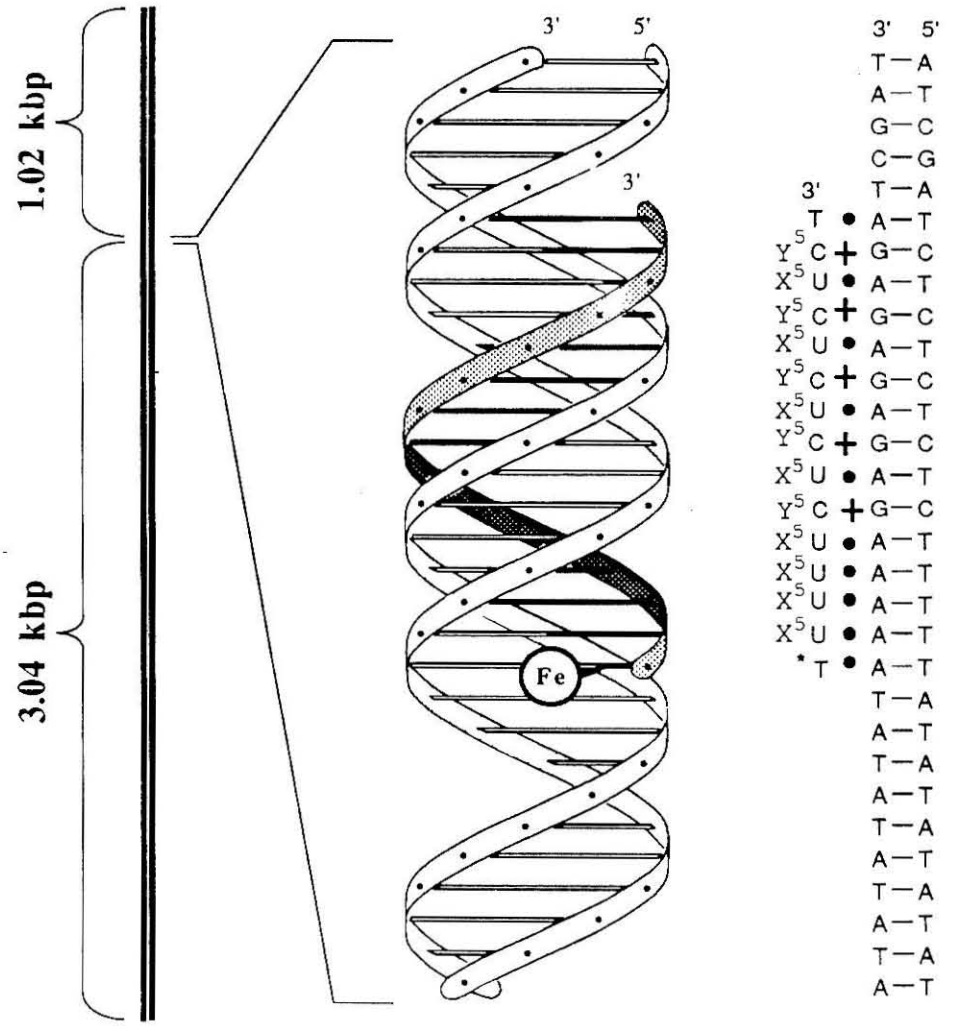
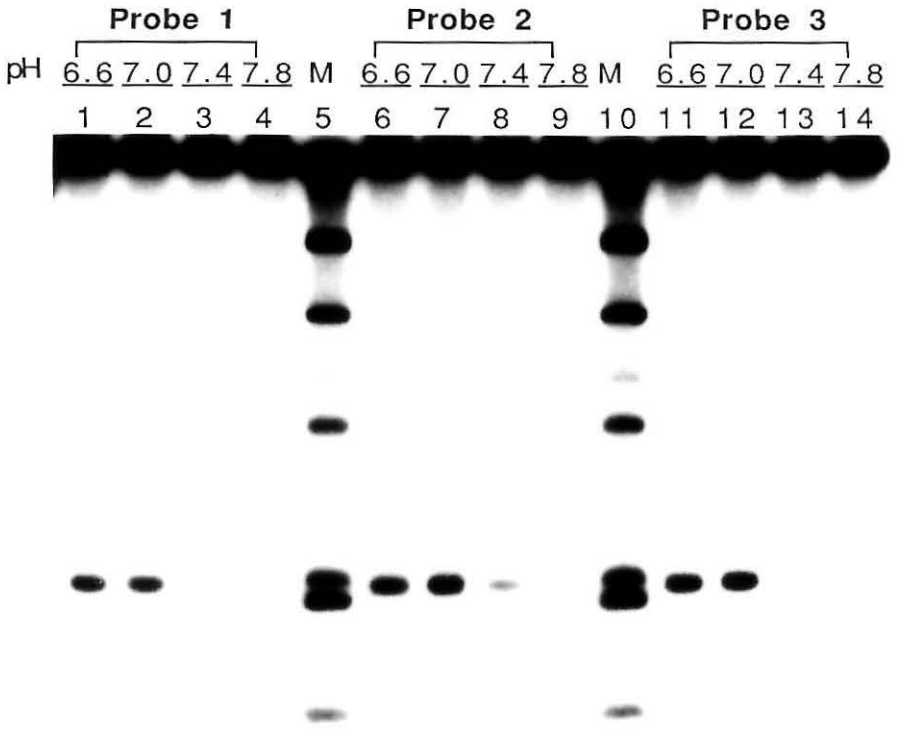
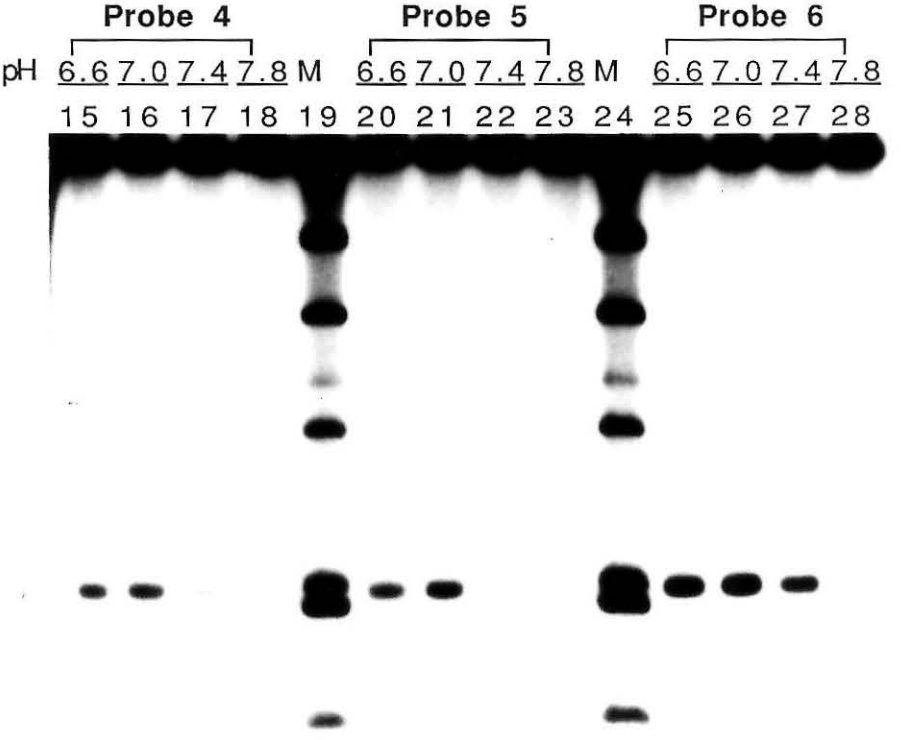


Figure 3.3. Autoradiograms of 0.9% agarose gels indicating double strand cleavage of plasmid DNA by oligonucleotides **1-6**. Plasmid pDMAG10 was linearized with Sty I and labeled with [α-^{32}P]TTP, producing a 4.06 kbp restriction fragment specifically labeled 1.02 kbp from the target sequence. The ^{32}P-end-labeled DNA was dissolved in buffer containing NaCl, Tris, and spermine and was mixed with oligonucleotide-EDTA **1-6** previously equilibrated for 60 sec with 1.5 equiv of Fe(II). After incubation at 0 °C for 60 min, reactions were initiated by the addition of sodium ascorbate (final concentrations: 25 mM Tris-acetate, 1 mM spermine, 100 mM NaCl, 0.8 μM oligonucleotide-EDTA•Fe(II), and 1 mM ascorbate). The cleavage reactions were allowed to proceed for 6 h at 24 °C, and loaded directly onto an 0.9% agarose gel. Electrophoresis at 120 V for 3 hours effected separation of the products. Cleavage conditions: lanes 1-4, **1**•Fe(II); lanes 6-9, **2**•Fe(II); lanes 11-14, **3**•Fe(II); lanes 15-18, **4**•Fe(II); lanes 20-23, **5**•Fe(II); lanes 25-28, **6**•Fe(II); lanes 1, 6, 11, 15, 20, and 25 are at pH 6.6; lanes 2, 7, 12, 16, 21, and 26 are at pH 7.0; lanes 3, 8, 13, 17, 22, and 27 are at pH 7.4; lanes 4, 9, 14, 18, 23, and 28 are at pH 7.8. Lanes 5, 10, 19, and 24 are DNA size markers obtained by digestion of Sty I linearized pDMAG10 with EcoR I, Pvu I, BamH I, and Xmn I, resulting in fragments 4058 (undigested DNA), 3067, 2994, 2371, 1687, 1064, 991, and 664 bp in length.


Probe 1
Probe 2
Probe 3
pH 6.6 7.0 7.4 7.8 M 6.6 7.0 7.4 7.8 M 6.6 7.0 7.4 7.8
1 2 3 4 5 6 7 8 9 10 11 12 13 14


Probe 4
Probe 5
Probe 6
pH 6.6 7.0 7.4 7.8 M 6.6 7.0 7.4 7.8 M 6.6 7.0 7.4 7.8
15 16 17 18 19 20 21 22 23 24 25 26 27 28

Results and Discussion

Oligonucleotides constructed from 2'-deoxyuridine, 5-bromo-2'-deoxyuridine and 5-methyl-2'-deoxycytidine.

Six oligonucleotides containing combinations of uracil/cytosine (U/C, **1**), uracil/5-methylcytosine (U/Me^{5}C, **2**), thymine/cytosine (T/C, **3**), thymine/5-methylcytosine (T/Me^{5}C, **4**), 5-bromouracil/cytosine (Br^{5}U/C, **5**), and 5-bromouracil/5-methylcytosine (Br^{5}U/Me^{5}C, **6**) bases were synthesized by automated methods with thymidine-EDTA (T*)[22] at the 5' end (Fig. 3.2, top). The efficiency of double strand cleavage of DNA by oligonucleotide-EDTA•Fe(II) **1-6** was analyzed over the pH range 6.6 - 7.8 at 25 °C on plasmid DNA (4.06 kilobase pairs) containing the 15 base pair homopurine target sequence, 5'-$A_5(GA)_5$-3' (Fig. 3.2, bottom).[23] The DNA cleavage products were separated by agarose gel electrophoresis and quantitated by scintillation counting of each band.

Oligonucleotide-EDTA•Fe **1-6** cleave double-stranded DNA at a single site corresponding to the target sequence (Fig. 3.3).[3] The cleavage efficiency of **3** containing C and T decreases sharply above pH 7.0.[3] Since the ability of the EDTA•Fe(II) moiety to cleave DNA in Tris buffer increases from pH 6.6 to 7.4, the decrease in cleavage efficiency observed is likely due to a decrease in the binding affinity of the oligonucleotide for its target sequence.[24]

Replacement of C with Me^{5}C (oligonucleotides **2**, **4**, and **6**) increases the oligonucleotide affinity and extends the pH range for binding. Substitution of Br^{5}U for T (**5**) increases binding affinity but does not change the pH profile greatly. Incorporation of both Me^{5}C and Br^{5}U (**6**) results in a large increase in

cleavage efficiency over an extended pH range. Oligonucleotides 1 and 2, constructed with U/C and U/Me^{5}C, show lower binding affinities (Table 3.1).

				cleavage efficiency, pH			
Oligo	(base)	X	Y	6.6	7.0	7.4	7.8
1	(U,C)	H	H	•	•	-	-
2	(U, Me5C)	H	Me	••	••	•	-
3	(T,C)	Me	H	••	••	-	-
4	(T,Me5C)	Me	Me	•••	•••	•	-
5	(Br5U,C)	Br	H	•••	•••	•	-
6	(Br5U,Me5C)	Br	Me	••••	••••	•••	•

Table 3.1. Absolute cleavage efficiencies of oligonucleotides **1-6** as determined by scintillation counting of bands cut from dried gels: • = 2-4%, •• = 5-7%, ••• = 8-10%, •••• = 10-20%.

The relative stabilities of base triplets are Br^{5}U•AT > T•AT > U•AT and Me^{5}C•GC > C$^+$•GC (Fig. 3.4).

Effect of 5-Br substituent on X^{5}U•A Hoogsteen base pairing. The substitution of bromo for methyl might effect the electronic complementarity, hydrophobicity, and polarizability (base stacking) the resulting oligonucleotide.

There are two opposing electronic effects to consider for Hoogsteen hydrogen bonding between X^{5}U and A. The electron-withdrawing bromo substituent increases the acidity at N3H (better hydrogen donor) but decreases the electron-donating properties of the carbonyl lone pair (poorer hydrogen acceptor). Methylation has the opposite effect. The pK_a's of deoxyuridine, thymidine, and 5-bromodeoxyuridine are 9.3,[25] 9.8,[25] and 8.1.[20]

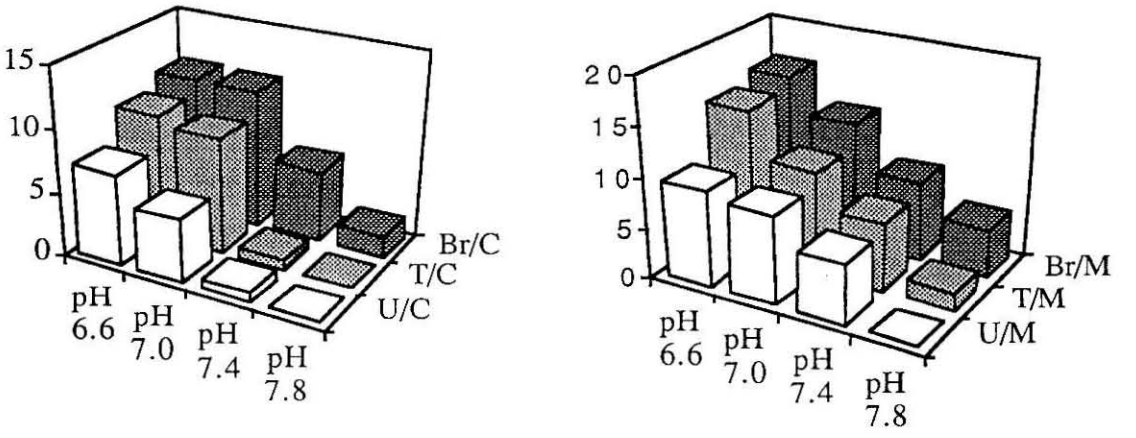


Fig. 3.4. Absolute cleavage efficiencies of oligonucleotides **1-6** quantitated by scintillation counting as a function of pH.

Enhancement of binding by bromo substitution might be due to increased acidity at N3 or enhanced hydrophobicity or both. Polynucleotides containing Br^5U are known to form more stable duplexes than their thymidine analogues.[26-30] NMR data of nucleosides in chloroform indicate that 1-cyclohexyl-5-bromouracil associates with 9-ethyl-adenine about 2.4 times as strongly as does 1-cyclohexyl-thymine[26,27,31]. This NMR data suggest that the increase in complex stability is due to an increase in the hydrogen bond interactions between the bases. In these studies, the Hoogsteen-hydrogen bonded complexes are found to be more stable than Watson-Crick complexes.[32,33] Thus, NMR data concerning the hydrogen-bonding of modified pyrimidines may actually be a better predictor of triplex affinities than duplex stabilities.

The 5-bromo substituent has also been suggested to increase the stacking interactions within poly Br^5U strands upon complex formation.[21] Higuchi and co-workers have found that poly Br^5U forms a 2:1 complex with poly (6-

methyl-A), whereas poly (U) forms only a 1:1 complex.[26,27,34-36] Methylation at N-6 of adenine is expected to disrupt hydrogen bonding with either the Hoogsteen or Crick pyrimidine strand. The stability of the 2:1 Br^5U:6-Me-A complex was interpreted as an indication of the stability imparted by base stacking in poly Br^5U complexes.

Incorporation of Br^5U into DNA is known to result in a higher rate of mutagenesis.[37-39] A variety of mechanisms, including ionization of the base[20,40] and tautomerism to the enol form,[41,42] have been invoked to explain the increase in misincorporation of bases. It has been calculated that at neutral pH, 7.4% of 5-bromo-2'-deoxyuridine residues are ionized, as opposed to 0.16% of residues.[40] The enol tautomer is also known to be more than ten times as favored for Br^5U as it is for T.[41] These considerations may make triple helix forming oligonucleotides containing Br^5U less specific than their T counterparts. No quantitative data concerning the specificity of these oligonucleotides has been gathered. UV melting data suggests that Br^5U is as specific as T,[43] while Br^5U/Me^5C oligonucleotides have been found to bind to a large number of secondary sites on large DNA.[44] The sequence and number of mismatches present at these sites is unknown.

Effect of 5-methyl substituent. Methylation of C also has two electronic effects: electron donation stabilizes protonation at N3 in the triplex, but decreases the acidity of the N4 exocyclic amino group (poorer hydrogen bond donor). In addition, methylation at C causes extension of the pH range (~ 0.4 units) for triple helix stability. This might be explained by an increase in the pK_a of $N3H^+$ in the triplex.[25] Stabilization of protonation at N3 is modest in the uncomplexed C; the reported pK_as of deoxycytidine and 5-methyldeoxycytidine are 4.3 and 4.4.[25]

Methylation of uracil exerts the opposite electronic effect as bromination, decreasing the acidity at N3H (poorer hydrogen donor) but increases the electron-donating properties of the carbonyl lone pair (better hydrogen acceptor). Since both methylation and bromination of uracil bases results in increases in pyrimidine oligonucleotide binding affinity, at least part of this stabilization must arise as a result of effects which are not electronic in nature.

Oligonucleotides **2** and **3**, which involve the triplet changes Me^5C+GC/U•AT to C^+GC/T•AT, show approximately equal binding affinities at or below pH 7.0. This suggests that methylation results in increased binding affinity irrespective of the methylated pyrimidine (Fig. 3.3). Methylation exerts a similar effect on duplex stability. Poly(T)[45-47] and poly(Me^5C)[48-51] form duplexes of greater stabilities with their complementary polymers than do poly(U) and poly(C). Methylation also promotes the stabilities of self-associated complexes and complexes with other purines.[31,52] Lee and Morgan observed that polynucleotide triplexes containing Me^5C in both pyrimidine strands (Me^5C•G-Me^5C base triplets) are stable to pH 8.0.[53] It was not clear if this effect was observed due to the effect of methylation in the Crick strand, Hoogsteen strand, or both strands. The data presented here, which indicates that methylation of the Hoogsteen strand stabilizes triplexes, in conjunction with previous data, in which one of the duplex strands is methylated, support the contention that <u>methylation promotes complex formation in general, regardless of the type of complex formed</u> (duplex or triplex).

One possible interpretation is that substitution of methyl for hydrogen at C-5 of pyrimidines promotes binding of the oligonucleotide via a hydrophobic effect. The observation that methylated polymers form more

stable duplexes with sequences where the methylated residues are stacked supports the contention that methyl-methyl contacts play a role in stabilizing duplexes.[50,51] For instance, poly(dMe^{5}C)•poly(dG) melts at higher temperature than poly(dMe^{5}C-dG)•poly(dG-dMe^{5}C). The unmethylated polynucleotide poly(dC)•poly(dG) melts at a lower temperature than poly(dC-dG)•poly(dG-dC).[50,51] Data suggestive of similar stabilizing contacts in the Hoogsteen strand of the triple helix has been reported.[43]

In the case of mixed oligonucleotides containing Br^{5}U and Me^{5}C (**6**) it is not inconceivable that a "hydrophobic spine" is created by the Br and methyl moieties aligned on the Hoogsteen strand in the major groove of DNA.

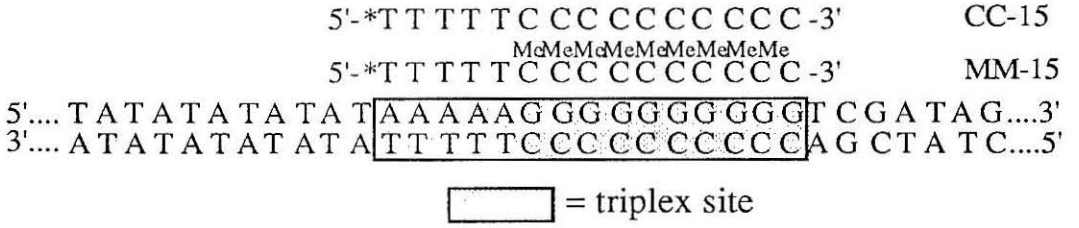


Fig. 3.5. Oligonucleotides CC-15 and MM-15 designed to bind to sequence A_5G_{10} within plasmid pDMG10.

The sequence A_5G_{10}, present in plasmid pDMAG10, has been studied by footprinting (chapter 2) and affinity cleaving.[54] Binding of oligonucleotide CC-15 occurs only below pH 5.8 or in the presence of high ethanol (40%) concentrations at close to neutral pH (pH 6.6), presumably due to the presence of ten protonation sites within the triple helical complex. The oligonucleotide MM-15 was synthesized to determine if the use of 5-methyl-2'-deoxycytidine might allow for the recognition of extended G sequences

under neutral conditions. No cleavage was observed in an aqueous solution at pH 6.6. It is possible that the G_{10} sequence adopts an exceptionally stable duplex conformation, which is difficult to disrupt.[19,55] The use of 5-methylcytosine bases results in a 50-fold increase in the binding affinity of an oligonucleotide 21 nucleotides in length containing 10, and 4 consecutive, cytosine bases.[56]

To better understand those factors which influence triple helix formation, oligonucleotides incorporating a more extensive series of 5-substituted pyrimidines were tested. Two types of modifications were studied. Based on the success of the 5-bromouracil containing oligonucleotides, a series of oligonucleotides containing halogenated uracil and cytosine bases was examined. The observation that methylation of either pyrimidine results in increased binding affinity led to the study of pyrimidines containing larger alkyl and alkynyl substituents at the 5 position.

Oligonucleotides Containing Halo^{5}pyrimidines

The synthesis of 5'-O-DMT-5-fluoro- (F^5U) and 5'-O-DMT-5-iodo-2'-deoxyuridines (I^5U) phosphoramidites followed usual procedures from the 5-iodo-2'-deoxyuridine and 5-fluoro-2'-deoxyuridine nucleosides obtained from Sigma.[57-60] These phosphoramidites, as well as 5-bromo-2'-deoxycytidine (Br^5C) phosphoramidite, are commercially available.

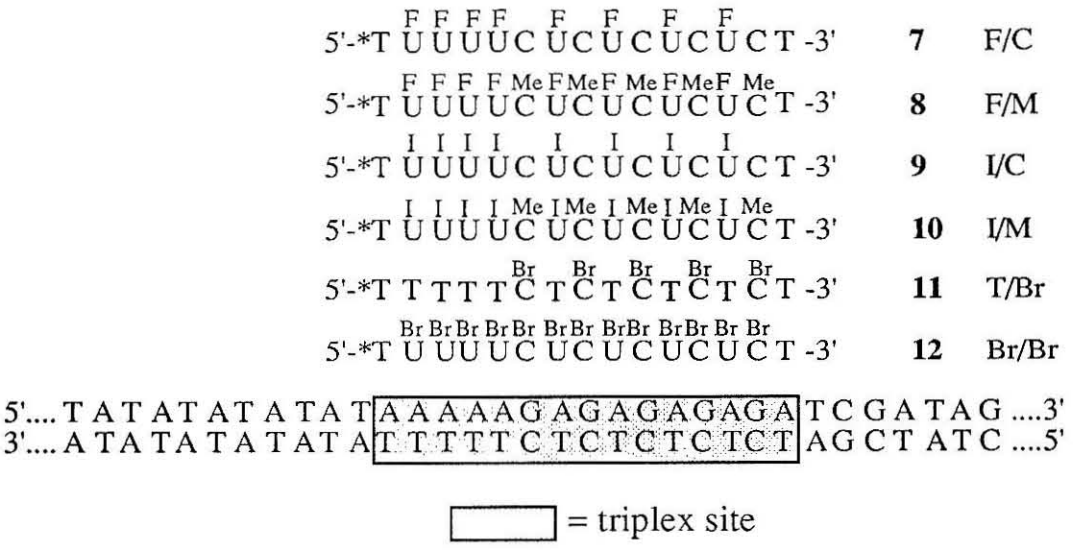


Fig. 3.6. Oligonucleotides **7-12** constructed from deoxyribonucleoside phosphoramidites containing thymidine (T), 5-fluoro-2'-deoxyuridine (F^5U), 5-iodo-2'-deoxyuridine (I^5U), 2'-deoxycytidine (C), 5-methyl-2'-deoxycytidine (Me^5C), 5-bromo-2'-deoxycytidine (Br^5C), and thymidine-EDTA (T*). Oligonucleotides **7-12** bind the target sequence $A_5(GA)_5$ in plasmid pDMAG10.[23]

Oligonucleotides **7-12** containing combinations of F^5U/C (**7**), F^5U/Me^5C (**8**), I^5U/C (**9**), I^5U/Me^5C (**10**), T/Br^5C (**11**) and Br^5U/Br^5C (**12**) were synthesized by automated methods with thymidine-EDTA (T*)[22] at the 5' end. Oligonucleotide concentrations were calculated using extinction coefficients obtained from UV analysis of the individual monomers in water and those found in the literature (Table 3.2).[25,39,61-63]

	260 nm	280 nm		260 nm	280 nm
dC	7300	7000	Br-5-dU	4600	9200
T	8800	6400	F-5-dU	7500	7200
U	9900		I-5-dU	3600	6900
Me-5-dC	5700	8300	Br-5-dC	3300	5800

Table 3.2. Extinction coefficients (M^{-1} cm^{-1}) at 260 and 280 nm for 5-substituted nucleosides.

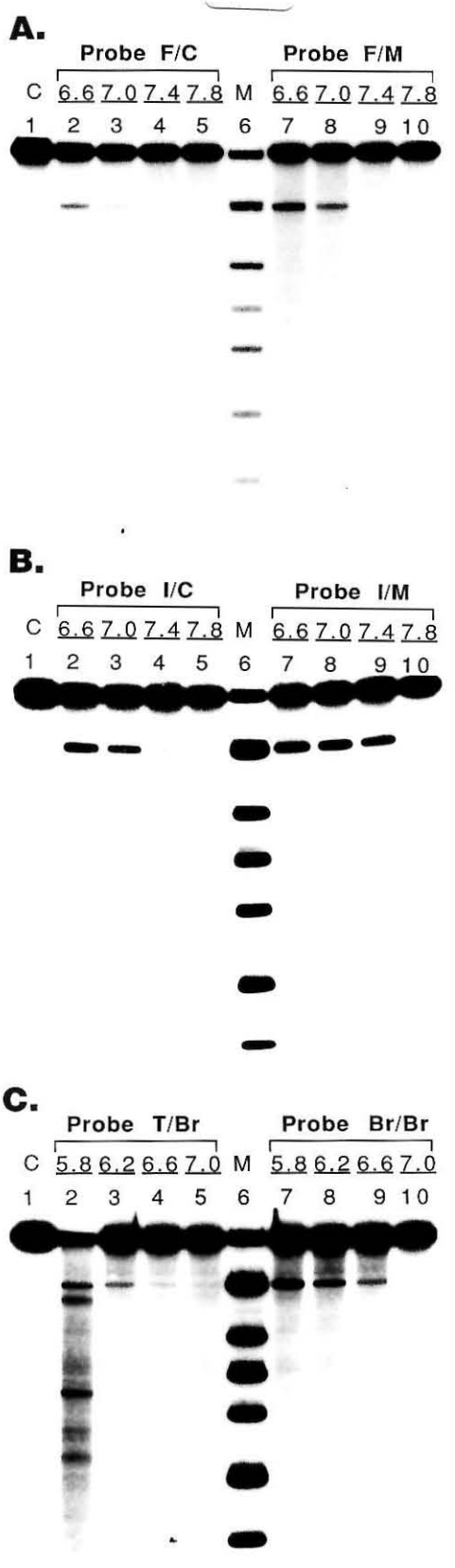
Figure 3.7. Autoradiograms of 0.9% agarose gels indicating double strand cleavage of plasmid DNA by oligonucleotides **7-12**. Plasmid pDMAG10 was linearized with Sty I and labeled with [α-^{32}P] dATP, producing a 4.06 kbp restriction fragment specifically labeled 3.04 kbp from the target sequence. The ^{32}P-end-labeled DNA was dissolved in buffer containing NaCl, Tris, and spermine and was mixed with oligonucleotide-EDTA **7-12** previously equilibrated for 60 sec with 1.5 equiv of Fe(II). After incubation at 0 °C for 60 min, reactions were initiated by the addition of sodium ascorbate (final concentrations: 25 mM Tris-acetate, 1 mM spermine, 100 mM NaCl, 0.8 μM oligonucleotide-EDTA•Fe(II), and 1 mM ascorbate). The cleavage reactions were allowed to proceed for 6 h at 24 °C. The reactions were loaded directly onto an 0.9% agarose gel. Electrophoresis at 120 V for 3 hours effected separation of the products.

Cleavage conditions:

Gel A. lanes 1-4, **7**•Fe(II); lanes 7-10, **8**•Fe(II); lanes 2 and 7 are at pH 6.6; lanes 3 and 8 are at pH 7.0; lanes 4 and 9 are at pH 7.4; lanes 5 and 10 are at pH 7.8; lane 1 is untreated DNA; lane 6 contains DNA size markers obtained by digestion of Sty I linearized pDMAG10 with EcoR I, Pvu I, BamH I, and Xmn I, resulting in fragments 4058 (undigested DNA), 3569, 2992, 2369, 1685, 1460, 1066, 670, and 489 base pairs in length.

Gel B. lanes 1-4, **9**•Fe(II); lanes 7-10, **10**•Fe(II); lanes 2 and 7 are at pH 6.6; lanes 3 and 8 are at pH 7.0; lanes 4 and 9 are at pH 7.4; lanes 5 and 10 are at pH 7.8; lane 1 is untreated DNA; lane 6 contains DNA size markers obtained by digestion of Sty I linearized pDMAG10 with EcoR I, Pvu I, BamH I, and Xmn I, resulting in fragments 4058 (undigested DNA), 3569, 2992, 2369, 1685, 1460, 1066, 670, and 489 base pairs in length.

Gel C. lanes 1-4, **11**•Fe(II); lanes 7-10, **12**•Fe(II); lanes 2 and 7 are at pH 5.8; lanes 3 and 8 are at pH 6.2; lanes 4 and 9 are at pH 6.6; lanes 5 and 10 are at pH 7.0; lane 1 is untreated DNA; lane 6 contains DNA size markers obtained by digestion of Sty I linearized pDMAG10 with EcoR I, Pvu I, BamH I, and Xmn I, resulting in fragments 4058 (undigested DNA), 3569, 2992, 2369, 1685, 1460, 1066, 670, and 489 base pairs in length.


A.
Probe F/C
Probe F/M
C 6.6 7.0 7.4 7.8 M 6.6 7.0 7.4 7.8
1 2 3 4 5 6 7 8 9 10
B.
Probe I/C
Probe I/M
C 6.6 7.0 7.4 7.8 M 6.6 7.0 7.4 7.8
1 2 3 4 5 6 7 8 9 10
C.
Probe T/Br
Probe Br/Br
C 5.8 6.2 6.6 7.0 M 5.8 6.2 6.6 7.0
1 2 3 4 5 6 7 8 9 10

Oligonucleotides **7 - 12** cleave double-stranded pDMAG10 DNA at a single site corresponding to the target sequence (Fig. 3.7). Cleavage efficiencies for these oligonucleotides as determined by scintillation counting of the bands cut from dried gels are indicated in Table 3.3.

				cleavage efficiency, pH			
Oligo	(base)	X	Y	6.6	7	7.4	7.8
7	(F5U,C)	F	H	••	•	-	-
8	(F5U, Me5C)	F	Me	•••	••		-
9	(I5U,C)	I	H	••	••	•	-
10	(I5U,Me5C)	I	Me	•••	•••	••	-
11	(T,Br5C)	Me	Br	-	-	-	-
12	(Br5U,Br5C)	Br	Br	•	-	-	-

Table 3.3. Absolute Cleavage Efficiencies of Oligonucleotides **7-12** as determined by scintillation counting of bands cut from dried gels: • = 2-4%, •• = 5-7%, ••• = 8-10%.

Incorporation of 5-iodo-2'-deoxyuridine. Comparison of halogenated deoxyuridine derivatives indicates that the relative stabilities of the base triplets are Br^5U•AT > I^5U•AT > T (Me^5U)•AT > F^5U•AT (Fig. 3.8). The greater electron withdrawing capability of the bromo substituent results in an increased acidity of the N-3 bond, and formation of stronger Hoogsteen hydrogen bonds with thymidine.[20,21,25,27,31,35] Due to the lower electron affinity of I, the iodo substituent does not increase the strength of this interaction relative to T as greatly as Br. The pK_as of these nucleosides have been determined to be 8.1 for Br^5U[20,63] and 8.3 for I^5U.[63] NMR studies of the bases in chloroform indicate that 5-iodouracil associates with adenine with a relative affinity 2.2 times greater than thymine-adenine, and 0.9 that of 5-

bromouracil-adenine.[31] This result is almost exactly what is observed for the affinity of Hoogsteen oligonucleotides as determined by affinity cleaving.

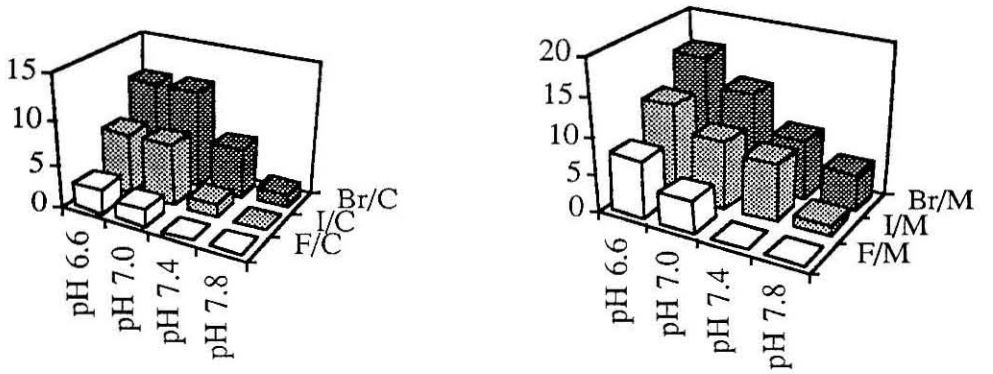


Fig. 3.8. Absolute cleavage efficiencies of oligonucleotide **5-10** as determined by scintillation counting as a function of substitution at 5-uridine positions and pH.

Incorporation of 5-fluoro-2'-deoxyuridine. 5-Fluoro-2'-deoxyuridine containing oligonucleotides bind with lower affinities than thymidine oligonucleotides. Fluorouridine has a pK_a of 7.7[59,63] indicating that, at slightly alkaline pH, a significant fraction of these residues may be ionized.[40,59,61] Lack of a large substituent at the 5-position may result in loss of van der Waals interactions between 5-substituted bases within the triplex. The van der Waals radius of F is 1.35 Å, closer to that of H (1.20 Å) than that of methyl (2.00 Å).[64] The stability of the F^5U•A-T triplet again parallels that observed with F^5U-A base pairs, which are less stable than T-A base pairs.[59]

Incorporation of 5-bromo-2'-deoxycytidine. 5-Bromo-2'-deoxycytidine was incorporated into oligonucleotides to determine if the increased binding observed with 5-bromouridine was due to an electronic effect or to non-base specific van der Waals or hydrophobic interactions in the major groove.

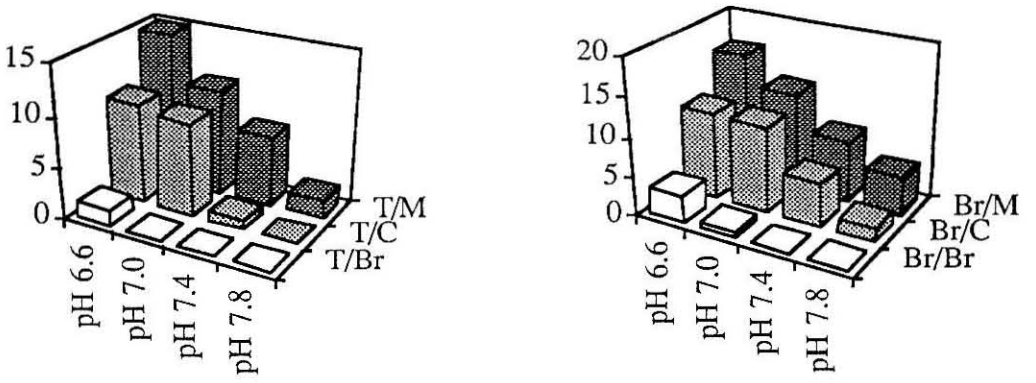


Fig. 3.9 Absolute cleavage efficiency of oligonucleotides **3**, **4**, **5**, **6**, **11**, and **12** as determined by scintillation counting of dried gels as a function of pH and substitution at 5-cytosine positions.

Comparison of 5-substituted cytidines reveals base triplet stability decreases in the order Me^5C+G-C > C+G-C » Br^5C+G-C (Fig. 3.9). This is in contrast to base pair stabilities, in which the Br^5C-G base pair is more stable than C-G.[48] Destabilization of C+G-C base triplet by bromination on the Hoogsteen strand must arise from a predominantly electronic effect. The pK_a of 5-bromocytidine has been reported to be 3.0,[61] indicating that protonation of Br^5C, required for Hoogsteen base pairing, becomes more difficult. This is the only 5-pyrimidine substitution which results different effects on Watson-Crick duplex formation and Hoogsteen triplet formation, and arises from the change in the hydrogen bonding pattern in C recognition of G-C base pairs, which requires protonation, and C recognition of G in duplex DNA.

Effects which are base specific and not general to both triplex and duplex stabilities are most likely due to electronic effects. Substitution at the 5 position of cytosine most likely exerts a large electronic effect, since

protonation and accommodation of a positive charge within the triplex is required for C+G-C base triplets.

Studies of 5-alkyl and 5-alkynyl-2'-deoxyuridines

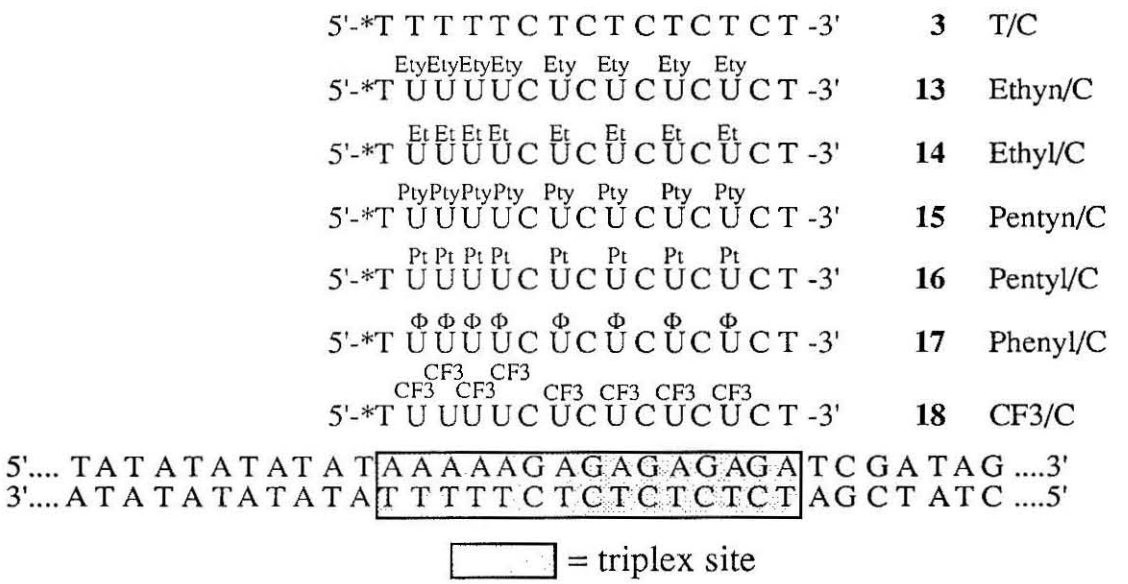


Fig. 3.10. Oligonucleotides **3** and **13 - 18** constructed from deoxyribonucleotide phosphoramidites containing 2'-deoxycytidine (C) and thymidine (T), 5-ethynyl-2'-deoxyuridine (Ety), 5-ethyl-2'-deoxyuridine (Et), 5-pentynyl-2'-deoxyuridine (Pty), 5-pentyl-2'-deoxyuridine (Pt), 5-(2-phenyl-ethynyl)-2'-deoxyuridine (Φ), 5-trifluoromethyl-2'-deoxyuridine (CF3), and thymidine-EDTA (T*).

Methyl group substitution at the 5 positions of cytosine or uracil results in an increase in binding affinity, indicating that the presence of a substituent larger than hydrogen at the 5-position of pyrimidines may promote triple helix formation. This could be due to formation of a hydrophobic spine or disruption of water binding in the major groove.[50] This observation stimulated the synthesis of oligonucleotides containing 5-alkyl and 5-alkynyl substituents of different sizes.

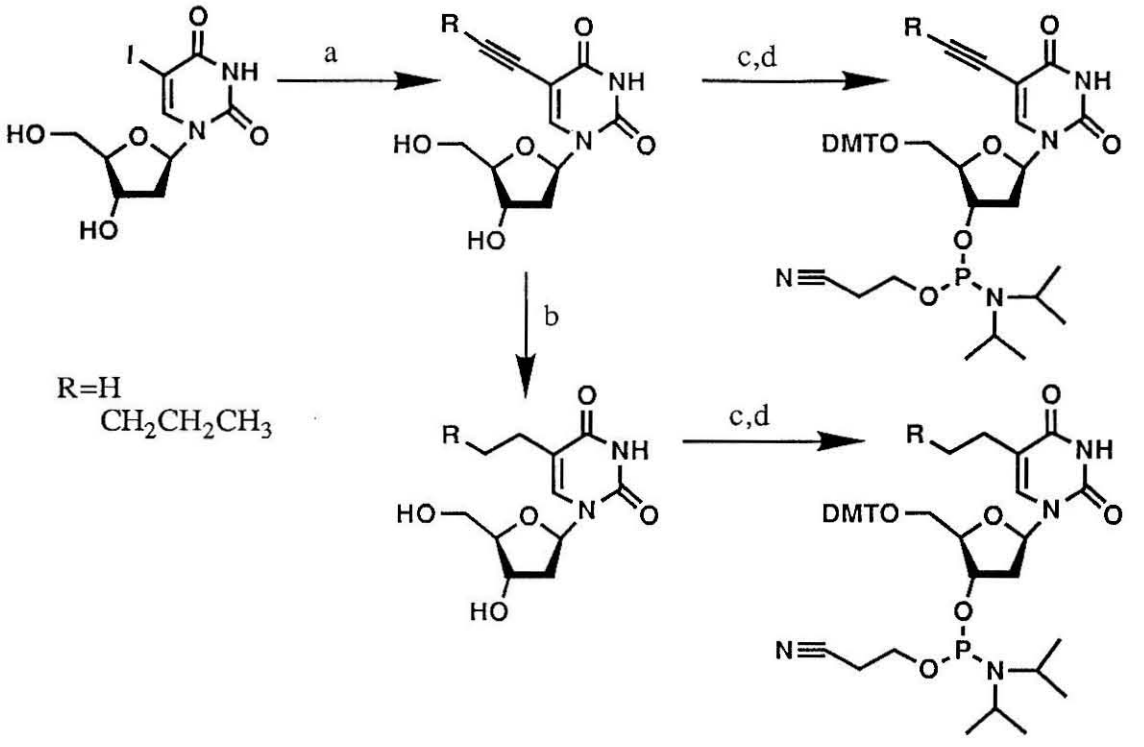


Fig. 3.11. Scheme for the synthesis of 5'-O-DMT-5-alkynyl- and 5'-O-DMT-5-alkyl-2'-deoxyuridine phosphoramidites. a) terminal alkyne in a Pd(0) catalyzed reaction; b) Pd/C, H_2; c) DMTCl, pyr.; d) *N,N*-diisopropyl-2-cyanoethylchloro phosphoramidite, DIPEA, $MeCl_2$.

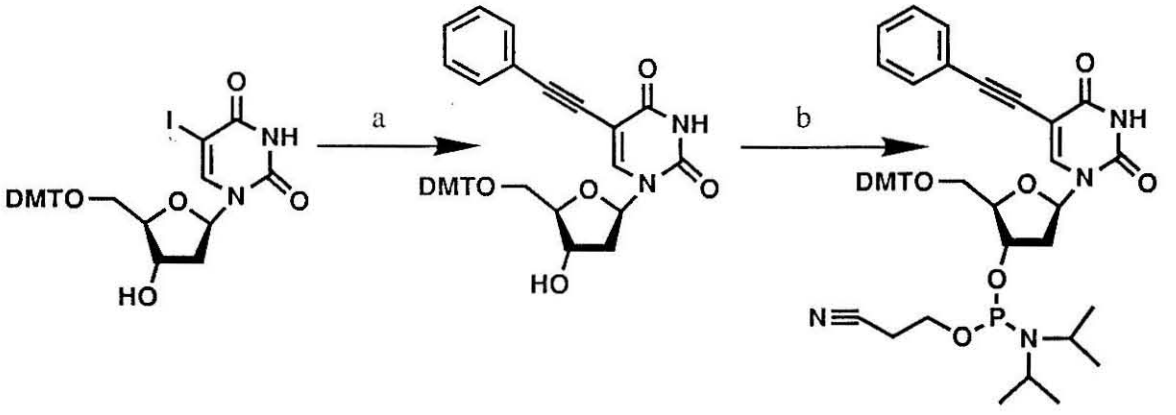


Fig. 3.12. Scheme for the synthesis of 5'-O-DMT-5-(2-phenyl-ethynyl)-2'-deoxyuridine phosphoramidite: a) phenylacetylene, Pd(0), CuI; b) *N,N*-diisopropyl-2-cyanoethylchloro phosphoramidite, DIPEA, $MeCl_2$.

Oligonucleotides containing 5-ethyl-, 5-ethynyl-, 5-(pent-1-ynyl)-, 5-pentyl- 5-(2-phenyl-ethynyl)-2'-deoxyuridines, and 5- trifluoromethyluracil bases were assayed for their ability to bind and cleave pDMAG10 double helical DNA (Fig. 3.9). For ease of discussion, 5-(pent-1-ynyl)-2'-deoxyruidine will be noted as simply 5-pentynyl-2'-deoxyuridine.

Synthesis of nucleosides. 5-Ethynyl-2'-deoxyuridine (Ety^{5}U) and 5-pentynyl-2'-deoxyuridine (Pty^{5}U) were synthesized by the coupling of propargyltrimethylsilane and 1-pentyne to 5-iodo-2'-deoxyuridine using a Pd catalyzed coupling as described by Robins and Barr, and by Hobbs (Fig. 3.11).[65-69] 5-Ethyl-2'-deoxyuridine (Et^{5}U) and 5-pentyl-2'deoxyuridine (Pt^{5}U) were obtained by reduction of 5-ethynyl-2'-deoxyuridine and 5-pentynyl-2'-deoxyuridine respectively with Pd/C and H_2.[67] 5'-O-DMT-5-(2-phenyl-ethynyl)-2'-deoxyuridine (Φ^5U) was obtained by directly coupling phenylacetylene to 5'-O-DMT-5-iodo-2'-deoxyuridine (Fig. 3.12).[67,70]

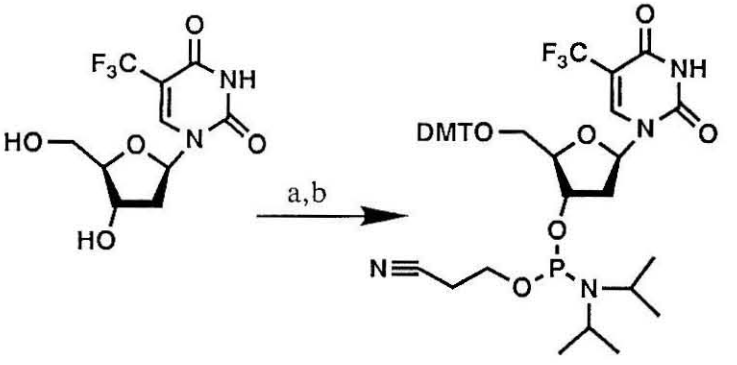


Fig. 3.13. Scheme for the synthesis of 5'-O-DMT-5-trifluoromethyl-2'-deoxyuridine phosphoramidites for the automated synthesis of oligonucleotides containing 5-trifluorouridines. (a) DMTCl, pyr; (b) *N,N*-diisopropyl-2-cyanoethyl-chlorophosphoramidite, $MeCl_2$, DIPEA.

The nucleoside 5-trifluoromethyl-2'-deoxyuridine is commercially available. Protection (5'-O-DMT) and activation (3'-N,N-diisopropylphosphoramidite) affords a nucleoside compatible with automated synthesis (Fig. 3.13).[57]

	260 nm	280 nm
Ethynyl-5-dU	3,800	9700
Ethyl-5-dU	8,700	6400
Pentynyl-5-dU	3,800	8200
Pentyl-5-dU	7,900	7600
2-Ph-Ety-5-dU	12,700	12,200
CF3-5-dU	8,500	2170

Table 3.4. Extinction coefficients (M^{-1} cm^{-1}) determined for 5-substituted-2'-deoxyuridines.

The extinction coefficients of each of the nucleosides were determined at 260 nm and 280 nm, and agreed with those calculated from published values. The extinction coefficients for 5-(2-phenyl-ethynyl)-2'-deoxyuridine were estimated from published values,[67] since the free nucleoside was not synthesized. The results of these determinations are listed in Table 3.4, and were used to quantitate oligonucleotide concentrations.

The cleavage efficiencies of these oligonucleotides as a function of pH was determined using affinity cleaving (Fig. 3.14). The binding affinities of these oligonucleotides as a function of pH are shown in Table 3.5. All of these oligonucleotides contain cytosine bases for the recognition of G-C base pairs. Cleavage efficiency of the oligonucleotide T/C (3) is shown for comparison.

Affinity cleaving indicates that the relative stabilities of base triplets is ethyn^{5}U•AT > T•AT > pentynyl^{5}U•AT > 2-phenyl-ethynyl^{5}U•AT > ethyl^{5}U•AT > pentyl^{5}U•AT (Fig. 3.14). Only the oligonucleotide containing

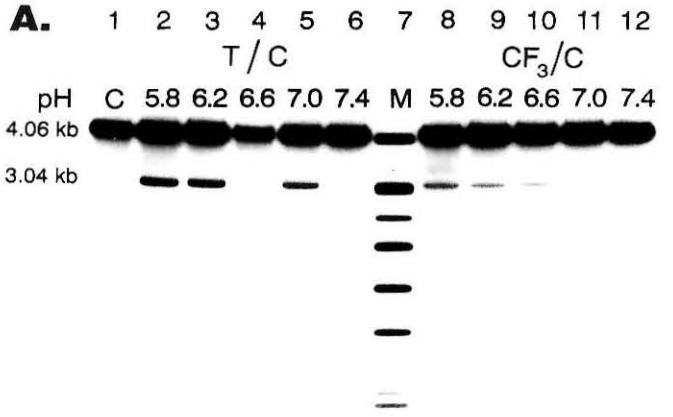
Figure 3.14. Autoradiograms of 0.9% agarose gel indicating double strand cleavage of plasmid pDMAG10 by oligonucleotides **3**, and **13-18**. Plasmid pDMAG10 was linearized with Sty I and labeled with [α-^{32}P] dATP, producing a 4.06 kbp restriction fragment specifically labeled 3.04 kbp from the target sequence. The ^{32}P-end-labeled DNA was dissolved in buffer containing NaCl, Tris, and spermine and was mixed with oligonucleotide-EDTA **3, 13-18** previously equilibrated for 60 sec with 1.5 equiv of Fe(II). After incubation at 0 °C for 60 min, reactions were initiated by the addition of sodium ascorbate (final concentrations: 25 mM Tris-acetate, 1 mM spermine, 100 mM NaCl, 0.8 μM oligonucleotide-EDTA•Fe(II), and 1 mM ascorbate). The cleavage reactions were allowed to proceed for 6 h at 24 °C. The reactions were loaded directly onto an 0.9% agarose gel. Electrophoresis at 120 V for 3 hours effected separation of the products.

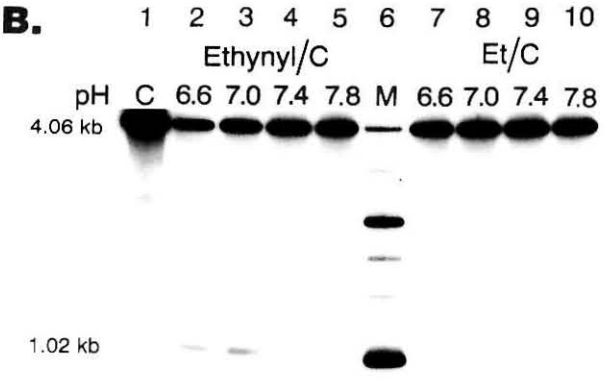
Cleavage conditions:

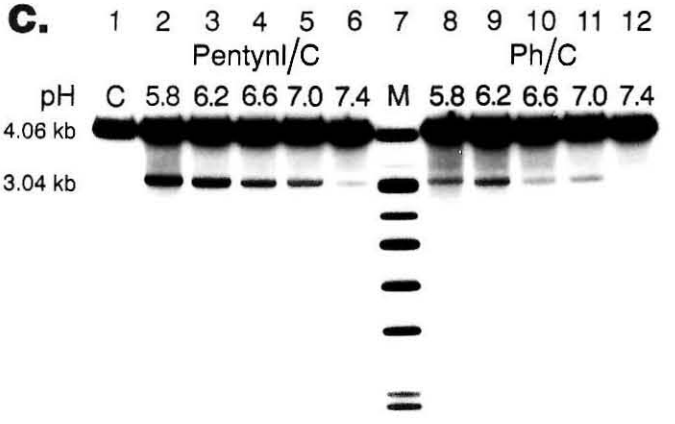
Gel A. lanes 1-5, **3**•Fe(II); lanes 8-12, **18**•Fe(II); lanes 2 and 8 are at pH 5.8; lanes 3 and 9 are at pH 6.2; lanes 4 and 10 are at pH 6.6; lanes 5 and 11 are at pH 7.0; lanes 6 and 11 are at pH 7.4; lane 1 is untreated DNA; lane 7 contains DNA size markers obtained by digestion of Sty I linearized pDMAG10 with EcoR I, Pvu I, BamH I, and Xmn I, resulting in fragments 4058 (undigested DNA), 3569, 2992, 2369, 1685, 1460, 1066, 670, and 489 base pairs in length.

Gel B. lanes 1-4, **13**•Fe(II); lanes 7-10, **14**•Fe(II); lanes 2 and 7 are at pH 6.6; lanes 3 and 8 are at pH 7.0; lanes 4 and 9 are at pH 7.4; lanes 5 and 10 are at pH 7.8; lane 1 is untreated DNA; lane 6 contains DNA size markers obtained by digestion of Sty I linearized pDMAG10 with EcoR I, Pvu I, BamH I, and Xmn I, resulting in fragments 4058 (undigested DNA), 3569, 2992, 2369, 1685, 1460, 1066, 670, and 489 base pairs in length.

Gel C. lanes 1-5, **15**•Fe(II); lanes 8-12, **17**•Fe(II); lanes 2 and 8 are at pH 5.8; lanes 3 and 9 are at pH 6.2; lanes 4 and 10 are at pH 6.6; lanes 5 and 11 are at pH 7.0; lanes 6 and 11 are at pH 7.4; lane 1 is untreated DNA; lane 7 contains DNA size markers obtained by digestion of Sty I linearized pDMAG10 with EcoR I, Pvu I, BamH I, and Xmn I, resulting in fragments 4058 (undigested DNA), 3569, 2992, 2369, 1685, 1460, 1066, 670, and 489 base pairs in length.


A.
1 2 3 4 5 6 7 8 9 10 11 12
T / C
CF3/C
pH C 5.8 6.2 6.6 7.0 7.4 M 5.8 6.2 6.6 7.0 7.4
4.06 kb
3.04 kb


B.
1 2 3 4 5 6 7 8 9 10
Ethynyl/C
Et/C
pH C 6.6 7.0 7.4 7.8 M 6.6 7.0 7.4 7.8
4.06 kb
1.02 kb


C.
1 2 3 4 5 6 7 8 9 10 11 12
Pentynl/C
Ph/C
pH C 5.8 6.2 6.6 7.0 7.4 M 5.8 6.2 6.6 7.0 7.4
4.06 kb
3.04 kb

5-ethynyluracil bases displays a greater binding affinity than oligonucleotide **3** (T/C). Oligonucleotide **17** containing 5-(2-phenyl-ethynyl)-2'-deoxyuridine binds with significantly lower affinities than oligonucleotides containing T/C bases. 5-Trifuoromethyluracil oligonucleotides display weak binding.

				cleavage efficiency, pH				
Oligo	(base)	X	Y	5.8	6.2	6.6	7.0	7.4
3	(T,C)	Me	H			••	••	-
13	(5-ethynylU,C)	Ethyn	H	••••	•••	•••	•••	•
14	(5-ethylU,C)	Ethyl	H	•	•	-	-	-
15	(5-pentynylU,C)	Pentyn	H	•••	•••	••	•	-
16	(5-pentylU,C)	Pentyl	H	-	-	-	-	-
17	(5-phenylU,C)	Phenyl	H	••••	••••	•••	•	-
18	(5-CF3U,C)	CF3	H	••	•	•	•	-

Table 3.5. Absolute double strand DNA cleavage efficiencies of oligonucleotides **13-18**: • = 2-4%, •• = 5-7%, ••• = 8-10%, •••• = 10-20%.

Effect of 5-alkyl substitution. Alkyl substitution results in oligonucleotides with greatly reduced binding affinities. Oligonucleotides containing 5-ethyl-2'-deoxyuridines display very weak binding, even at pH < 6.2. At these pHs, binding to secondary sites is observed, resulting in loss of specificity. No cleavage was observed with oligonucleotides containing 5-pentyl-2'-deoxyuridines.

The effect of substitution at the 5-position of pyrimidines on the stability of X^{5}U•AT triplets parallels the effect on X^{5}U-A base pairs. Incorporation of ethyl^{5}U into DNA results in a decrease in duplex stability.[37,46,71-73] Triplexes of the type poly(Et^{5}U)•poly(A)-poly(Et^{5}U) also display lower melting temperatures than poly(Me^{5}U)•poly(A)-poly(Me^{5}U) triplexes.[46,72] Duplex stability is also decreased by incorporation of larger alkyl groups.[72,74,75] Even

modest changes in the 5-methyl group, such as change to 5-hydroxymethyl group, results in loss of duplex stability.[47] It is unlikely that a large change in the hydrogen bonding interactions occurs upon incorporation of a 5-ethyl substituent.[71] The loss of complex stability has been attributed primarily to a change in the stacking interactions in poly Et^{5}U. [46,71,72]

The incorporation of T*[22] into oligonucleotides has been shown to lower the melting temperature of the duplex formed with a complementary oligonucleotide.[76] Incorporation of more than two T* residues into a triple helix forming oligonucleotide (18 bases long) decreases the amount of cleavage observed.[77] These observations are consistent with and can be understood on the basis of the observation that substitution by large alkyl groups at the 5-position of 2'-deoxyuridines decreases the stabilities of both duplexes and oligonucleotide-directed triplexes.

Effect of 5-alkynyl substitution. Oligonucleotides containing 5-alkynyl-uracil bases bind with much higher affinities that their 5-alkyl counterparts. Oligonucleotides constructed from 5-ethynyl-2'-deoxyuridine not only bind better than oligonucleotides containing 5-ethyl-2'-deoxyuridine, but thymidine as well. Longer alkynyl chains again lower binding affinity (ethynyl » pentynyl). Longer unsaturated chains have not been studied.

The 2-phenyl-ethynyl substituent represents modification with a large planar group. Modelling of the triple helical complexes formed by this oligonucleotide indicates that these groups probably do not stack on one another within a triplex, due to the 30° rotation about the helical axis between each residue. Oligonucleotides containing 5-(2-phenyl-ethynyl)-2'-deoxyuridines bind with lower affinities than thymidine containing oligonucleotides, but with higher affinities than oligonucleotides constructed from 5-alkyl-2'-deoxyuridines.

5-Trifluoromethyl-2'-deoxyuridine. The trifluoromethyl substituent could incorporate both the electron withdrawing and the steric characteristics thought to increase the binding affinity of uridine nucleosides in triple helix formation. Two explanations for the low binding affinity of oligonucleotide **18** are that the pK_a (the pK_a of 5-trifluorouracil is 7.35[64]) of this derivative results in deprotonation at neutral pH, or the substituent may be too large to fit within the major groove (the van der Waals radius is 2.44 , as compared to 2.00 for the methyl group on thymidine[64]). 5-Trifluoromethyl-2'-deoxyuridine is also known to undergo a variety of degradation reactions under basic conditions.[78-81] At the time these experiments were completed, no analysis (HPLC) was conducted on the purified oligonucleotide to determine its nucleoside content, and decomposition of the base during oligonucleotide deprotection may have occurred. Literature precedents suggest that under the basic conditions used to deprotect oligonucleotides, the 5-trifluoromethyl-2'-deoxynucleoside is completely degraded to 5-carboxy or 5-carboxyamido-2'-deoxyuridine.[78-81]

Design of other modified nucleosides. Incorporation of 5-ethylcytosine bases into polymers also lowers the ability of the polymer to form duplex DNAs.[82] This suggests that it is unlikely that 5-alkyl substitution at cytidine will drastically increase oligonucleotide directed triple helical stability. Alkyl substitution at position 8 of purines, however, results in increases in duplex stabilities. This suggests that the use of 8-alkyl purines may increase purine oligonucleotide binding affinity in oligonucleotide-directed triple helices of the purine•purine-pyrimidine type.[83]

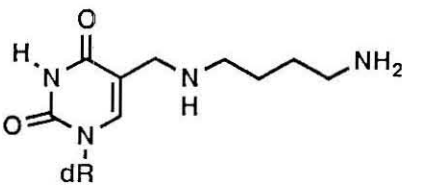


Fig. 3.15. 5-(4-Aminobutylaminomethyl)uracil base found in the DNA of the bacteriophage ϕW-14 .

One organism (the bacteriophage ϕW-14) with 5-modified uridines displays an increase in melting temperature of its DNA. The modified base was identified as α-putrescinyl-thymine.[84,85] DNA containing this modified thymidine displays increased stability and a lower dependence of T_m on salt concentration. Properly substituted amino-thymidines may display similar effects in triple helix formation.

Conclusions

5-substituted pyrimidines can be used to alter the stabilities of oligonucleotide-directed triple helical complexes. Table 3.6 summarizes the cleavage efficiencies observed for the oligonucleotides studied.

Oligonucleotides constructed with 5-methyl substituted pyrimidines (thymidine or 5-methyl-2'-deoxycytidine) form more stable complexes than unmethylated pyrimidines (T•A-T and Me^5C+G-C are more stable than U•A-T or C+G-C base triplets). 5-iodouracil and 5-ethynyluracil bases form better triplets with A-T than T, but poorer than Br^5U. Substitution by larger alkynes (5-(2-phenyl-ethynyl)uracil and 5-pentynyluracil bases) results in decreases in binding affinities. 5-alkyluracil bases (5-ethyluracil and 5-pentyluracil) form very weak triple helical c complexes. These results parallel

Oligo	(Base)	X	Y	Cleavage efficiency, pH 5.8	6.2	6.6	7	7.4	7.8
1	(U,C)	H	H			•	•	-	-
2	(U,Me5C)	H	Me			••	••	•	-
3	(T,C)	Me	H			••	••	-	-
4	(T,Me5C)	Me	Me			•••	•••	•	-
5	(Br5U, C)	Br	H			•••	•••	•	-
6	(Br5U,Me5C)	Br	Me			••••	••••	•••	•
7	(F5U,C)	F	H			••	•	-	-
8	(F5U,Me5C)	F	Me			•••	••	-	-
9	(I5U,C)	I	H			••	••	•	-
10	(I5U,Me5C)	I	Me			•••	•••	••	-
11	(T,Br5C)	Me	Br			-	-	-	-
12	(Br5U,Br5C)	Br	Br			•	-	-	-
13	(Ety5U,C)	Ethyn	H	••••	•••	•••	•••	•	
14	(Et5U,C)	Ethyl	H	•	•	-	-	-	
15	(Pty5U,C)	Pentyn	H	•••	•••	••	•	-	
16	(Pt5U,C)	Pentyl	H	-	-	-	-	-	
17	(PhEty5U,C)	Phenyl	H	••••	••••	•••	•	-	
18	((CF3)5U,C)	CF3	H	••	•	•	•	-	

Table 3.6. Summary of absolute double strand DNA cleavage efficiencies of oligonucleotides **1-18**: - ≤ 2%, • = 2-4%, •• = 5-7%, ••• = 8-10%, •••• = 10-20%. Blank boxes indicate cleavage efficiency was not measured at that pH.

what is observed for duplex formation. 5-bromocytidine is known to stabilize duplex formation, but destabilizes triplex formation.

Substitution results in changes in the electronic complementarity (H-bonding potential), polarizability (base stacking), and hydrophobicity of the bases. While it is difficult to determine which effects predominate, we conclude: (a) substitution on C by electron withdrawing groups dramatically decreases triplex stability at neutral pH, due to the need for protonation at N-3. The opposite is observed in duplex formation; (b) the presence of a group

larger than H at the 5 position stabilizes nucleic acid complexes, due to non-electronic effects; (c) substitution by alkyl chains longer than methyl results in decreased complex stability, an effect most probably due to changes in base stacking interactions.

Materials and Methods

^{1}H nuclear magnetic resonance (NMR) spectra were recorded on a Jeol JNM-GX400 FT 400 MHz spectrometer and are reported in parts per million (ppm) from tetramethylsilane. ^{31}P NMR spectra were recorded on a Varian Associates EM-390 and are reported in parts per million from phosphoric acid. Ultravioltet-visible (UV-Vis) spectra were recorded on a Perkin-Elmer Lambda 4C UV/Vis spectrophotometer. Mass spectral determinations were conducted at the regional mass spectroscopy facility at the University of California, Riverside. Flash chromatography was carried out under positive air pressure using EM Science Kieselgel 60 (230-400 Mesh). Elemental analysis were performed by the analytical laboratory, California Institute of Technology.

Reagents. Dry DMF, methylene chloride and acetonitrile were obtained from Fluka over molecular sieves. 5-iodo-2'-deoxyuridine and 5-trifluoromethyl-2'-deoxyuridine were obtained from Sigma. CuI, triethylamine, 1-pentyne, propargyltrimethylsilane, phenylacetylene, palladium on charcoal (Pd/C), and tetrakis(triphenylphosphine)palladium(0), dimethoxytritylchloride, *N,N*-diisopropyl-2-cyanoethyl-chlorophosphoramidite were purchased from Aldrich and used as received.

Diisopropylethylamine was distilled from CaH_2 and stored under argon. 2'-Deoxyuridine, thymidine, 5-bromo-2'-deoxyuridine, 2'-deoxycytidine, and 5-methyl-2'-deoxycytidine phosphoramidites were purchased from Cruachem. 5-Fluoro-2'-deoxyuridine and 5-iodo-2'-deoxyuridine phsophoramidites were synthesized from 5-fluoro-2'-deoxyuridine and 5-iodo-2'-deoxyuridine by protection of the 5' hydroxyl with dimethoxytrityl chloride and activiation using 2-cyanoethyl-*N*,*N*-diisopropylchloro phosphoramidite via standard procedures.[57] These phosphoramidites are now available from Cruachem. 5-Bromo-2'-deoxycytidine phosphoramidite was purchased from Pharmacia.

Restriction enzymes were obtained from New England BioLabs or Boehringer Mannheim and used with the manufacturer's supplied buffers. Polynucleotide kinase and the Klenow fragment were obtained from Boehringer Mannheim. Radioactive nucleotides (^{32}P-γ-ATP, ^{32}P-α-dATP, and ^{32}P-α-TTP), were purchased from Amersham. Cobalt (III) hexamine trichloride was obtained from Kodak. Water was doubly distilled from a Corning Automatic Collection System.

Nucleoside Synthesis

5-Trimethylsilyl-ethynyl-2'-deoxyuridine. 5-Iodo-2'-deoxyuridine (2.80 g, 7.9 mmol) was dissolved in dry DMF (40 mL) under argon. CuI (327 mg, 1.71 mmol), triethylamine (2.4 mL, 17.2 mmol) and propargyltrimethylsilane (4.1 mL, 29.0 mmol) were added. Tetrakis(triphenylphosphine)palladium (0) (1.0 gms, 0.87 mmol) was added and the solution stirred at room temperature for 12 hours. The solution was diluted with methanol (25 mL) and methylene chloride (25 mL), and $NaHCO_3$ (4 gm) and polymer supported triphenylphosphine (1 gm) were added. After 30 min., the solution was filtered and concentrated under reduced pressure. The DMF was removed by

vacuum distillation (45 °C, < 1 mm Hg). Flash chromatography in methylene chloride (500 mL), followed by 3% methanol in methylene chloride (500 mL) afforded pure product in 84.6 % yield (2.16 gms). ^{1}H NMR and UV spectra were consistent with reported values.[65-67]

5-Ethynyl-2'-deoxyuridine. Sodium metal (462 mg, 20 mmol) was dissolved in MeOH (100 mL) to afford a 2 N solution of sodium methoxide. 5-(Trimethylsilyl)ethynyl-2'-deoxyuridine (1.6 gms. 4.9 mmol) was dissolved in methanol. The NaOMe solution (50 mL) was added to the nucleoside solution and stirred for five hours. TLC showed complete conversion to a single product. Amberlite G-50 resin was added to neutralize the solution. The resin was removed, washed with methanol, and the filtrate concentrated under reduced pressure. The product was dissolved in 95% ethanol and precipitated with diethyl ether to yield 740 mgs (60% yield) of the product as a crystaline solid. NMR and UV were consistent with reported values.[65-67]

5-Ethynyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine. 5-Ethynyl-2'-deoxyuridine (207 mg, 0.82 mmol) was dissolved in dry pyridine (7 mL). Dimethoxytrityl chloride (390 mg, 1.15 mmol) was added, and the solution stirred overnight at 4 °C. Water (12 mL) was added and the product extracted with diethyl ether (3 x 30 mL). After concentration, the product was purified by flash chromatography (64:34:4 ethyl acetate:hexane:methanol) to provide 428 mg of product (95% yield). ^{1}H NMR ($CDCl_3$) δ 8.07 (s, 1H), 7.41 (d, 2H), 7.20-7.33 (m, 8H), 6.83 (d, 4H), 6.27 (t, J = 7 Hz, 1H, $H_{1'}$), 4.55 (m, 1H, $H_{3'}$), 4.12 (m, 1H, $H_{4'}$), 3.79 (s, 6H, OMe), 3.36-3.42 (m, 2H, $H_{5'}$), 2.90 (s, 1H,) 2.47-2.51 (m, 1H, $H_{2'}$), 2.25-2.32 (m, 1H, $H_{2'}$). MS: (positive ion FAB) m/z 554 (M^+, 2), 304 (25), 303 (100), 154 (12), 137 (7), 136 (10). HRMS calc. for $C_{32}H_{30}O_7N_2$ (M^+): 554.2054. Found 554.2062.

5-Ethynyl-5'-O-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyl-*N,N*-diisopropy-phosphoramidite)-2'-deoxyuridine. 5-Ethyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine (382 mg, 0.68 mmol) was dissolved in dry methylene chloride (5 mL). Diisopropylethylamine (510 μL, 2.93 mmol) and 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (340 μL, 1.5 mmol) were added, and the reaction stirred for three hours. The reaction was quenched with ethanol (0.5 mL) and stirred for five minutes. The reaction was diluted with ethyl acetate (15 mL), washed twice with an aqueous solution of satd. $NaHCO_3$ and twice with an aqueous solution of satd. NaCl. After drying (Na_2SO_4) and concentration under reduced pressure, flash chromatography (64:32:5:1, ethyl acetate:hexane:triethylamine:methanol) afforded a white foam. ^1H NMR ($CDCl_3$) δ 8.14 (d, 1H), 7.42 (d, 2H), 7.21-6.34 (m, 8H), 6.82-8.85 (m, 4H), 6.27 (m, 1H), 4.63 (m, 1H), 4.17, 4.22 (d, 1H), 3.79 (s, 6H), 3.58-3.69 (m, 4H), 3.36 (d, 2H), 2.90 (d, 1H), 2.63 (t, 1H), 2.46 (t, 1H), 2.32 (m, 2H), 1.07-1.28 (m, 12H). ^{31}P NMR ($CDCl_3$) δ 149.42, 149.01.

5-Ethyl-2'-deoxyuridine. 5-Ethynyl-2'-deoxyuridine (339 mg, 1.35 mmol) was dissolved in ethanol and 10% Pd/C (200 mg) added. Hydrogenation was effected under 50 psi of H_2 on a Parr rocker. After 18 hours, TLC indicated complete conversion to a single product. The Pd/C was removed by filtration through Celite. Concentration afforded pure product. NMR and UV were consistent with reported values.[65-67]

5-Ethyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine. 5-Ethyl-2'-deoxyuridine (238 mg, 0.93 mmol) was dissolved in dry pyridine (5 mL). Dimethoxytrityl chloride (440 mg, 1.30 mmol) was added, and the solution stirred overnight at 4 °C. Water (15 mL) was added and the product extracted with diethyl ether (4 x 80 mL). After concentration, the product was purified by flash chromatography with 64:32:5 ethyl acetate:hexane:methanol to yield

405 mg of product in 78 % yield. ^{1}H NMR ($CDCl_3$) δ 8.24 (bs, 1H), 7.42 (s, 1H, H_6), 7.38 (d, 2H), 7.24-7.31 (m, 7H), 6.83 (d, 4H), 6.41 (t, 1H, $H_{1'}$), 4.55 (m, 1H, $H_{3'}$), 4.04 (m, 1H, $H_{4'}$), 3.79 (s, 6H, OMe), 3.46 (m, 1H, $H_{5'}$), 3.37 (m, 1H, $H_{5'}$), 2.29-2.36 (m, 2H, $H_{2'}$), 1.93 (m, 2H, CH_2-CH_3), 0.85 (t, 3 H). MS: (positive ion FAB) m/z 558 (M^+, 3), 304 (25), 303 (100), 154 (7), 136 (6). HRMS calc. for $C_{32}H_{34}O_7N_2$ (M^+): 558.2367. Found 558.2386.

5-Ethyl-5'-O-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyl-*N,N*-diisopropy-phosphoramidite)-2'-deoxyuridine. 5-Ethyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine (380 mg, 0.68 mmol) was dissolved in dry methylene chloride (5 mL) Diisopropylethylamine (510 μL, 2.93 mmol) and 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (330 μL, 1.48 mmol) were added, and the reaction stirred for three hours. The reaction was quenched with ethanol (0.5 mL) and stirred for five minutes. The reaction was diluted with ethyl acetate (15 mL), washed twice with an aqueous solution of satd. $NaHCO_3$ and twice with an aqueous solution of satd. NaCl. After drying (Na_2SO_4) and concentration, flash chromatography (64:32:5:1, ethyl acetate:hexane:triethylamine:methanol) afforded a white foam. ^{1}H NMR ($CDCl_3$) δ 8.30 (bs, 1H), 7.48 (d, 1H), 7.39 (d, 2H), 7.23-7.29 (m, 7H), 6.83 (m, 4H), 6.41 (m, 1H), 4.64 (m, 1H), 4.14 (d, 1H), 3.80 (s, 3H), 3.79 (s, 3H), 3.46-3.79 (m, 4H), 3.29 (m, 2H), 2.62 (t, 1H), 2.42 (t, 1H), 2.30 (m, 2H), 1.80-1.93 (m, 2H), 1.04-1.29 (m, 12H), 0.83 (m, 3H). ^{31}P NMR ($CDCl_3$) δ 149.32, 148.88.

5-(Pent-1-ynyl)-2'-deoxyuridine. 5-Iodo-2'-deoxyuridine (2.16 gms, 6.1 mmol) was dissolved in dry DMF (5 mL) under argon. CuI (233 mg, 1.22 mmol), triethylamine (1.7 mL, 12.15 mmol) and 1-pentyne (1.8 mL, 18.3 mmol) were added. Tetrakis(triphenylphosphine)palladium (0) (703 mg, 0.61 mmol) was added and the solution stirred at r. t. for nine hours. $NaHCO_3$ (4 gms) and polymer supported triphenylphosphine (1 gm) were added. After

30 min. the solution was filtered and concentrated by rotory evaporation, and the DMF removed by vacuum distillation (45 °C, < 1 mm Hg). Flash chromatography in 500 mL of methylene chloride, followed by 500 mL of 10% methanol in methylene chloride afforded pure product in 70.2 % yield (1.26 gms). NMR and UV are consistent with reported values.[66,67]

5-Pentynyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine. 5-Pentyl-2'-deoxyuridine (342 mg, 1.16 mmol) was dissolved in anhydrous pyridine (5 mL). Dimethoxytrityl chloride (474 mg, 1.40 mmol) was added and the solution stirred for 16 hours at 4 °C. Analysis by TLC (10% MeOH/$MeCl_2$) indicated some starting material remaining. Methanol (15 mL) was added, and the solution concentrated by rotoevaporation. The product was dissolved in diethyl ether, and washed with water and brine. The organic layer was dried over Na_2SO_4 and concentrated. Flash chromatography in 64:32:2 ethyl acetate:hexane:methanol afforded 550 mg of pure product (79% yield). ^1H NMR ($CDCl_3$) δ 8.07 (s, 1H), 7.95 (s, 1H), 7.43 (d, 2H), 7.22-7.35 (m, 7H), 6.84 (d, 4H), 6.30 (t, J = 6 Hz, 1H), 4.53 (m, 1H), 4.06 (m, 1H), 3.80 (s, 6H), 3.33-3.42 (m, 2H), 2.45 (m, 1H), 2.29 (m, 1H), 2.08 (m, 2H), 1.84 (m, 1H), 1.28 (q, 2H), 0.80 (t, 3H). MS: (positive ion FAB) m/z 596 (M^+, 1), 304 (25), 303 (100), 154 (6), 137 (4), 136 (5). HRMS calc. for $C_{35}H_{36}O_7N_2$ (M^+): 596.2524. Found 596.2552.

5-Pentynyl-5'-O-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyl-*N*,*N*-diisopropy-phosphoramidite)-2'-deoxyuridine. 5-(1-Pentynyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine (661 mg, 1.11 mmol) was dissolved in dry methylene chloride (5 mL). Diisopropylethylamine (800 μL, 4.48 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (570 μL, 2.56 mmol) were added, and the reaction stirred for 30 minutes. The reaction was quenched with ethanol (0.5 mL) and stirred for five minutes. The reaction was diluted with ethyl acetate (15 mL), washed twice with an aqueous solution of satd. $NaHCO_3$ and twice

with an aqueous solution of satd. NaCl. After drying (Na_2SO_4) and concentration, flash chromatography (1,5% triethylamine in methylene chloride) afforded a white foam. ^1H NMR ($CDCl_3$) δ 8.02 and 7.98 (d, 1H), 7.44 (d, 2H), 7.21-7.36 (m, 7H), 6.83 (m, 4H), 6.31 (m, 1H), 4.59 (m, 1H), 4.15 and 4.20 (dm, 1H), 3.78 (s, 3H), 3.79 (s, 3H), 3.50-3.75 (m, 4H), 3.42 (m, 1H), 3.29 (m, 1H), 2.42-2.65 (m, 3H),2.62 (t, 1H), 2.43 (t, 1H), 2.28 (m, 1H), 2.05 (m, 2H), 1.06-1.28 (m, 14H), 0.76 (q, 3H). ^{31}P NMR ($CDCl_3$) δ 149.25, 148.81.

5-Pentyl-2'-deoxyuridine. 5-(1-Pentynyl)-2'-deoxyuridine (380 mg, 1.3 mmol) was dissolved in ethanol and 5% Pd/C (150 mg) was added. Hydrogenation was affected under 50 psi of H_2 on a Parr rocker for one hour. The Pd/C was removed by filtration through Celite. Concentration afforded 360 mg of pure product (95% yield). NMR and UV were consistent with reported values.[66,67]

5-Pentyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine. 5-Pentyl-2'-deoxyuridine (300 mg, 1.0 mmol) was dissolved in anhydrous pyridine (5 mL) at 0 °C. Dimethoxytrityl chloride (481 mg, 1.42 mmol) was added and the solution stirred for 16 hours at 4 °C. Water (15 mL) was added, and the product extracted into 3 x 50 ml aliquots of diethyl ether. After concentration under reduced pressure, the product was purified by flash chromatography (64:32:2 ethyl acetate:hexane:methanol) to afford 500 mg of pure product (90% yield). ^1H NMR ($CDCl_3$) δ 8.03 (s, 1H), 7.44 (s, 1H, H_6), 7.39 (d, 2H), 7.23-7.31 (m, 7H), 6.83 (d, 4H), 6.38 (t, J = 7 Hz, 1H, $H_{1'}$), 4.56 (m, 1H, $H_{3'}$), 4.02 (m, 1H, $H_{4'}$), 3.79 (s, 6H), 3.45-3.49 (m, 2H, $H_{5'}$), 2.31-2.38 (m, 2H, $H_{2'}$), 1.90-2.02 (m, 1H), 1.89 (m, 1H), 1.75 (m, 1H), 1.25 (m, 2H), 1.11 (m, 2H), 1.03 (m, 2H), 0.77 (t, 3H). MS: (positive ion FAB) m/z 600 (M^+, 3), 304 (25), 303 (100), 183 (3), 154 (4), 136 (4). HRMS calc. for $C_{35}H_{40}O_7N_2$ (M^+): 600.2837. Found 600.2861.

5-Pentyl-5'-O-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyl-*N,N*-diisopropy-phosphoramidite)-2'-deoxyuridine. 5-Pentyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine (498 mg, 829 μmoles) was dried overnight under vacuum, and dissolved in dry methylene chloride (5 mL). Diisopropylethylamine (590 μL, 3.39 mmol) and 2-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite (380 μL, 1.7 mmol) were added, and the reaction stirred for one hour. The reaction was quenched with ethanol (1 mL) and stirred for five minutes. The reaction was diluted with ethyl acetate (15 mL), washed twice with an aqueous solution of satd. $NaHCO_3$ and twice with an aqueous solution of satd. NaCl. After drying (Na_2SO_4) and concentration, flash chromatography (64:32:5, ethyl acetate:hexane:triethylamine) afforded a white foam. ^1H NMR ($CDCl_3$) δ 7.95 (bs, 1H), 7.50 (d, 1H), 7.39 (d, 2H), 7.26-7.29 (m, 7H), 6.82 (m, 4H), 6.38 (m, 1H), 4.64 (m, 1H), 4.14 (dm, 1H), 3.78 (s, 6H), 3.46-3.63 (m, 5H), 3.29 (m, 1H), 2.62 (t, 1H), 2.42 (t, 1H), 2.25-2.50 (m, 2H), 1.96 (m, 1H), 1.68 (m, 1H), 0.84-1.25 (m, 18H), 0.74 (t, 3H). ^{31}P NMR ($CDCl_3$) δ 149.25, 148.78.

5-Trifluoromethyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyruridine. 5-Trifluoromethyl-2'-deoxyruridine (1.0 gms, 3.38 mmol) was dissolved in pyridine (7 mL) and chilled to 0 °C. Dimethoxytrityl chloride (1.15 gms, 3.38 mmol) was added, and the solution stirred for 14 hours at 4 °C. TLC (64:32:5, ethyl acetate:hexane:methanol) indicated the reaction was complete. The reaction was diluted with methanol (10 mL), concentrated and the pyridine removed under vacuum. The product was dissolved in methylene chloride and washed twice with water. After concentration, flash chromatography using 60:40:4 ethyl acetate:hexane:methanol afforded 910 mg of product (45 % yield). ^1H NMR ($CDCl_3$) δ 9.06 (bs, 1H), 8.11 (s, 1H), 7.36 (d, 2H), 7.20-7.29 (m, 7H), 6.82 (d, 4H), 6.21 (t, J = 7 Hz, 1H, $H_{1'}$), 4.42 (m, 1H. $H_{3'}$), 4.13 (m, 1H, $H_{4'}$), 3.78 (s, 6H), 3.41 (m, 2H, $H_{5'}$), 2.53-2.59 (m, 1H). 2.23 (m, 1H). MS: (positive ion

FAB) m/z 598 (M^+, 4), 319 (3), 305 (5), 304 (25), 303 (100), 289 (4), 273 (4), 213 (3), 165 (4), 154 (14), 138 (6), 136 (12), 135 (8), 109 (5), 107 (5). HRMS calc. for $C_{31}H_{29}O_7N_2F_3$ (M^+): 598.1928. Found 598.1907.

5-Trifluoromethyl--5'-O-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyl-*N*,*N*-diisopropy-phosphoramidite)-2'-deoxyuridine. 5-Trifluoromethyl-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine (296 mg, 0.49 mmol) was dissolved in dry methylene chloride (3 mL). Diisopropylethylamine (700 μL, 4.02 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (310 μL, 1.39 mmol) were added, and the reaction stirred for 15 minutes. TLC (64:32:5, ethyl acetate:hexane:methanol) indicated > 80% conversion of starting material to product. The reaction was quenched with ethanol (0.5 mL) and stirred for five minutes. The reaction was diluted with ethyl acetate (25 mL), washed twice with an aqueous solution of satd. $NaHCO_3$ and once with an aqueous solution of satd. NaCl. After drying (Na_2SO_4) and concentration, flash chromatography (97:2:1 methylene chloride:methanol:triethylamine) afforded 210 mg of a white foam (53% yield), which was used directly in oligonucleotide synthesis.

5-(2-Phenyl-ethynyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyruridine. 5-Iodo-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine (1.82 gms, 2.77 mmol) was dissolved in dry DMF (15 mL) under argon. CuI (353 mg, 1.85 mmol), triethylamine (0.85 mL, 6.10 mmol) and phenyl acetylene (0.95 mL, 8.7 mmol) were added. Tetrakis(triphenylphosphine)palladium (0) (1.07 gms, 0.93 mmol) was added and the solution stirred at r. t. for nine hours. The reaction was diluted with ethyl acetate (400 mL) and washed with water (60 mL, 5x), once with an aqueous solution of satd. $NaHCO_3$, and once with an aqueous solution of satd. NaCl. After drying over Na_2SO_4 and concentration, flash chromatography in 60:40:5 ethyl acetate:hexane:methanol afforded 610 mg of

pure product (35% yield). ^{1}H NMR ($CDCl_3$) δ 9.15 (bs, 1H), 8.16 (s, 1H), 7.43 (d, 2H), 7.04-7.35 (m, 12 H), 6.76 (t, 4H), 6.36 (t, J = 7 Hz, 1H, $H_{1'}$), 4.56 (m, 1H, $H_{3'}$), 4.14 (m, 1H, $H_{4'}$), 3.68 (s, 3H), 3.66 (s, 3H), 3.42 (m, 1H, $H_{5'}$), 3.31 (m, 1H, $H_{5'}$), 2.55 (m, 1H, $H_{2'}$), 2.31 (m, 1H, $H_{2'}$). MS: (positive ion FAB) m/z 630 (M^+, 0.4), 363 (2.7), 339 (2.5), 303 (100), 279 (11.3), 213 (4.9), 195 (3.0), 165 (4.3), 135 (6.7). HRMS calc. for $C_{38}H_{34}O_7N_2$ (M^+): 630.2367. Found 630.2361.

5-(2-Phenyl-ethynyl)-5'-O-(4,4'-dimethoxytrityl)-3'-(2-cyanoethyl-*N*,*N*-diisopropy-phosphoramidite)-2'-deoxyuridine. 5-(2-Phenyl-ethynyl)-5'-O-(4,4'-dimethoxytrityl)-2'-deoxyuridine (390 mg, 0.62 mmol) was dissolved in dry methylene chloride (3 mL). Diisopropylethylamine (800 μL, 4.6 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (240 μL, 1.0 mmol) were added, and the reaction stirred for 20 minutes. TLC (64:32:5, ethyl acetate:hexane:methanol) indicated reaction was complete. (The reaction was quenched with ethanol (0.5 mL) and stirred for five minutes. The reaction was diluted with ethyl acetate (25 mL), washed twice with an aqueous solution of satd. $NaHCO_3$ and once with an aqueous solution of satd. NaCl. After drying (Na_2SO_4) and concentration, flash chromatography (2% triethylamine in methylene chloride) afforded a white foam, which was lyophilized from dioxane, and used directly in oligonucleotide synthesis.

DNA Manipulations

Oligonucleotide Synthesis. Oligonucleotides were synthesized on 1 μM scale using a Beckman Plus 1 DNA synthesizer. Monomers were diluted to 0.1 M concentration in dry acetonitrile and coupled with no modification in the synthesis program. Deprotection was effected using 0.1 N NaOH for 24 hrs. at 55 °C. The solution was neutralized with 6 μL of glacial acetic acid,

desalted with Sephadex, lyophilized to dryness, and the oligonucleotide purified on a 20% polyacrylamide gel. The product was visualized by UV shadowing of the product band, and elution of the DNA with 0.2 M NaCl, 1 mM EDTA at 37 °C for 24 hrs. Filtration of the gel slice through a 0.45 μm cellulose acetate (Centrex) filter, followed by extensive dialysis against water (> five days, changing dialysis water three times daily) yielded pure metal-free oligonucleotides.

DNA preparation. Plasmid pDMAG10 was grown and amplified as described previously.[23,86] pDMAG10 was purified using two CsCl gradients (first gradient, 24 hrs, 42,000 rpm; second gradient, 12 hrs., 55,000 rpm). Plasmid DNA was linearized with Sty I, which produces overhangs which can be uniquely labeled on either end. Labeling with either [α-^{32}P] TTP or [α-^{32}P] dATP using the Klenow fragment[87] produces a 4.06 kb restriction fragment specifically labeled 1.02 or 3.04 kb from the target sequence. Unincorporated nucleotides were removed using a 1 ml Sephadex G-50-80 spin column.[87]

Affinity cleaving reactions. The ^{32}P-end-labeled DNA was dissolved in buffer containing NaCl, Tris, and spermine and was mixed with oligonucleotide-EDTA **1-18** previously equilibrated for 60 sec with 1.5 equiv of Fe(II). After incubation at 0 °C for 60 min, reactions were initiated by the addition of sodium ascorbate (final concentrations: 25 mM Tris-acetate, 1 mM spermine, 100 mM NaCl, 0.8 μM oligonucleotide-EDTA•Fe(II), and 1 mM ascorbate). The cleavage reactions were allowed to proceed for six hours at 24 °C. After addition of ficoll loading buffer, reactions were loaded directly onto an 0.9% agarose gel and electrophoresed at 120 V for three hours.

References

(1) Modrich, P. *Crit. Rev. Biochem.* **1982**, *13*, 287-323.

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

(3) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.

(4) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927-7929.

(5) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73-75.

(6) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **1957**, *79*, 2023-2024.

(7) Michelson, A. M.; Massoulie, J.; Guschlbauer, W. *Prog. Nucl. Acids. Res. Mol. Biol.* **1967**, *6*, 83-141.

(8) Lipsett, M. N. *J. Biol. Chem.* **1964**, *239*, 1256-1260.

(9) Lipsett, M. N. *Biochem. Biophys. Res. Comm.* **1963**, *11*, 224-228.

(10) Felsenfeld, G.; Miles, H. T. *Ann. Rev. Biochem.* **1967**, *36*, 407.

(11) Miller, J. H.; Sobell, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **1966**, *55*, 1202.

(12) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63-80.

(13) Lee, J. S.; Johnson, D. A.; Morgan, A. R. *Nucl. Acids Res.* **1979**, *6*, 3073-3091.

(14) Busa, W. B. *Am. J. Physiol.* **1984**, *246*, R409-R438.

(15) Busa, W. B. *Ann. Rev. Physiol.* **1986**, *48*, 389-402.

(16) Bright, G. R.; Fisher, G. W.; Rogowska, J.; Taylor, D. L. *J. Cell Biol.* **1987**, *104*, 1019-1033.

(17) Boron, W. F. *J. Membr. Biol.* **1983**, *72*, 1-16.

(18) Madshus, I. H. *Biochem. J.* **1988**, *250*, 1-8.

(19) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag Inc.: New York, 1984.

(20) Lawley, P. D.; Brooks, P. *J. Mol. Biol.* **1962**, *4*, 216-219.

(21) Bugg, C. E.; Thomas, J. M.; Sundaralingam, M.; Rao, S. T. *Biopolymers* **1971**, *10*, 175-219.

(22) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968-972.

(23) Mendel, D.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 910-914.

(24) Herzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23*, 3934-3945.

(25) *Handbook of Biochemistry*; Sober, H. A., Ed.; CRC Press: Cleveland, OH, 1970.

(26) Higuchi, S.; Yasui, K. *Biopolymers* **1984**, *23*, 831-841.

(27) Higuchi, S.; Yasui, K. *Nucl. Acids Res. Sym. Ser.* **1982**, *11*, 189-192.

(28) Inman, R. B.; Baldwin, R. *J. Mol. Biol.* **1962**, *5*, 172-184.

(29) Kit, S.; Hsu, T. C. *Biochem. Biophys. Res. Commun.* **1961**, *5*, 120.

(30) Szybalski, W.; Mennigmann, H. D. *Anal. Biochem.* **1962**, *3*, 267.

(31) Kyogoku, Y.; Lord, R. C.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1967**, *57*, 250-257.

(32) Iwahashi, H.; Sugeta, H.; Kyogoku, Y. *Biochemistry* **1982**, *21*, 631-638.

(33) Engel, J. D.; von Hippel, P. H. *Biochemistry* **1974**, *13*, 4143-4149.

(34) Higuchi, S.; Yasui, K. *Nucl. Acids Res. Sym. Ser.* **1983**, *12*, 185-188.

(35) Higuchi, S. *Biopolymers* **1984**, *23*, 493-509.

(36) Higuchi, S. *J. Biomolec. Struct. Dyn.* **1985**, *2*, 675-682.

(37) De Clercq, E.; Shugar, D. *Biochemical Pharm.* **1975**, *24*, 1073-1078.

(38) Swierkowska, K. M.; Jasinska, J. K.; Steffen, J. A. *Biochem. Pharmac.* **1973**, 22, 85-93.

(39) Hutchinson, F. *Quart. Rev. Biophys.* **1973**, *6*, 201-246.

(40) Driggers, P. H.; Beattie, K. L. *Biochemistry* **1988**, *27*, 1729-1735.

(41) Katritzky, A. R.; Waring, A. J. *J. Chem. Soc.* **1962**, *1962*, 1540-1544.

(42) Freese, E. *J. Mol. Biol.* **1959**, *1*, 87-105.

(43) Kool, E.; Dervan, P. B. California Institute of Technology, 1990.

(44) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172-174.

(45) Szer, W.; Swierkowski, M.; Shugar, D. *Acta Biochim. Polon.* **1963**, *10*, 87.

(46) Barszcz, D.; Sugar, D. *Eur. J. Biochem.* **1968**, *5*, 91-100.

(47) Cassidy, P.; Kahan, F.; Alegria, A. A. *Fed. Proc.* **1965**, *24*, 2.

(48) Hoheisel, J. D.; Craig, A. G.; Lahrach, H. *J. Biol. Chem.* **1990**, *265*, 16656-16660.

(49) Zmudzka, B.; Bollum, F. J.; Shugar, D. *Biochemistry* **1969**, *8*, 3049-3059.

(50) Gill, J. E.; Mazrimas, J. A.; Bishop, C. C., Jr. *Biochim. Biophys. Acta* **1974**, *335*, 330-348.

(51) Szer, W.; Shugar, D. *J. Mol. Biol.* **1966**, *17*, 174.

(52) Fikus, M.; Shugar, D. *Acta Biochim. Polon.* **1969**, *16*, 55.

(53) Lee, J. S.; Woodsworth, M. L.; Latimer, L. J. P.; Morgan, A. R. *Nucl. Acids Res.* **1984**, *12*, 6603-6614.

(54) Moser, H. "Oligonucleotide-Directed Cleavage of Double Helical DNA by Triple-Strand Formation," California Institute of Technology, 1988.

(55) McCall, M.; Brown, T.; Kennard, O. *J. Mol. Biol.* **1985**, *183*, 385-396.

(56) Maher, L. J., III; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-730.

(57) *Oligonucleotide Synthesis: A Practical Approach*; Gait, M. J., Ed.; IRL Press: Oxford, 1984.

(58) Sheardy, R. D.; Seeman, N. C. *J. Org. Chem.* **1986**, *51*, 4301-4303.

(59) Kremer, A. B.; Mikita, T.; Beardsley, G. P. *Biochemistry* **1987**, *26*, 391-397.

(60) Habener, J. F.; Vo, C. D.; Le, D. B.; Gryan, G. P.; Ercolani, L.; Wang, A. H.-J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 1735-1739.

(61) Wempen, I.; Fox, J. J. *J. Med. Chem.* **1964**, *7*, 2474-2477.

(62) Lohmann, W. *Z. Naturforsch* **1974**, *29*, 493-495.

(63) Berens, K.; Shugar, D. *Acta Biochim. Pol.* **1963**, *10*, 25-47.

(64) Gottschling, H.; Heidelberger, C. *J. Mol. Biol.* **1963**, *7*, 541-560.

(65) Barr, P. J.; Jones, A. S.; Serafinowski, P.; Walker, R. T. *J. Chem. Soc. Perkin I* **1978**, *1978*, 1263-1267.

(66) Robins, M. J.; Barr, P. J. *Tetrahedron* **1981**, *22*, 421-424.

(67) Robins, M. J.; Barr, P. J. *J. Org. Chem.* **1983**, *48*, 1854-1862.

(68) Robins, M. J.; Vinayak, R. D.; Wood, S. G. *Tet. Lett.* **1990**, *31*, 3731-3734.

(69) Hobbs, F. W. *J. Org. Chem.* **1989**, *54*, 3420-3422.

(70) Farina, V.; Hauck, S. I. *SYNLETT* **1991**, *1991*, 157-159.

(71) Swierkowski, M.; Shugar, D. *Acta Biochim. Pol.* **1969**, *16*, 263-277.

(72) Swierkowski, M.; Shugar, D. *J. Mol. Biol.* **1970**, *47*, 57-67.

(73) Pietrzykowska, I.; Shugar, D. *Acta Biochim. Pol.* **1967**, *14*, 169-181.

(74) Marmur, J.; Brandon, C.; Neubort, S.; Ehrlich, M.; Mandel, M.; Konvicka, J. *Nature New Biology* **1972**, *239*, 68-70.

(75) Brandon, D.; Gallop, P. M.; Marmur, J.; Hayashi, H.; Nakanishi, K. *Nature New Biology* **1972**, *239*, 70-72.

(76) Griffin, L. C. Ph. D. Thesis, California Institute of Technology, 1990.

(77) Strobel, S. A. Ph. D. Thesis, California Institute of Technology, 1991.

(78) Reyes, P.; Heidelberger, C. *Mol. Pharmacol.* **1965**, *1*, 14-30.

(79) Heidelberger, C.; Parsons, D. G.; Remy, D. C. *J. Med. Chem.* **1964**, *7*, 1-5.

(80) Ryan, K. J.; Acton, E. M.; Goodman, L. *J. Org. Chem.* **1966**, *31*, 1181-1184.

(81) Khaja, T. A.; Heidelberger, C. *J. Med. Chem.* **1969**, *12*, 543-545.

(82) Kulikowski, T.; Shugar, D. *Biochim. Biophys. Acta* **1974**, *374*, 164-175.

(83) Helmkamp, C. K.; Kondo, N. S. *Biochim. Biophys. Acta* **1968**, *157*, 242.

(84) Takeda, T.; Ikeda, K. *Nucl. Acids Res. Sym. Ser.* **1983**, *12*, 75-78.

(85) Kropinski, A. M. B.; Bose, R. J.; Warren, R. A. J. *Biochemistry* **1973**, *12*, 151-157.

(86) Mendel, D. Ph. D. Thesis, California Institute of Technology, 1990.

(87) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N. Y., 1989.

Chapter IV

Efficient Base Specific Alkylation of Double Helical DNA by *N*-Bromoacetyloligonucleotides

Introduction

Attachment of a non-specific cleaving moiety capable of generating a diffusible cleaving species to a DNA binding molecule allows for the elucidation of the structural principles for DNA recognition, a technique termed affinity cleaving.[1,2] The determination of the sequence specificities, groove locations, and binding orientations of peptide analogues,[3-12] protein-DNA binding complexes,[13-17] and oligonucleotide-triple helix structures[18-24] has provided reliable models for the sequence specific recognition of double helical DNA.

Pyrimidine oligonucleotide-directed triple helix formation is a motif for the recognition of double helical DNA, which allows the facile construction of molecules (oligonucleotides) capable of binding a wide variety of DNA sequences.[18,25-27] Within this motif, a pyrimidine oligonucleotide binds in the major groove parallel to the Watson-Crick purine strand.[18,25-27] Generality and specificity are imparted by the existence of a Hoogsteen base pairing code, wherein thymidine binds specifically to adenine-thymine base pairs (T•A-T triplet formation) and protonated cytosine recognizes guanosine-cytosine base pairs (C+G-C triplet formation).[18,25-28] The recent discoveries of other base triplets, such as G•TA,[20] the development of 3'-3'

linked oligonucleotides for alternate-strand triple helix formation,[21] and the formation of novel triple helical structures[24,29,30] has greatly extended the number of sites capable of being recognized. Triple helix formation is both generalizable (in principle, the recognition of any purine- or purine-pyrimidine sequence is possible) and, by virtue of its >15 base pair specificity, more than a million fold more specific than conventional 6-base pair restriction enzymes. This specificity is sufficient for the recognition of single sites in genomic DNA.[27,31]

Studies using oligonucleotides equipped with a non-specific cleaving functionality (EDTA•Fe(II)) indicate that recognition of and cleavage at a single site within the 48,000 base pairs of the lambda phage genome is possible.[32] Furthermore, the cleavage of a yeast chromosome (340 kbp) at an engineered site, albeit in low yield (6%), has been accomplished.[25] Cleavage at secondary sites is observed.[25] Presumably, these sites are largely complementary in a Hoogsteen sense to the oligonucleotide-EDTA•Fe(II). The number of such sites found (four secondary sites were observed within the 340 kbp of chromosome III, indicating an approximate ratio of one site per 100,000 base pairs of DNA) may be due to the known high incidence of polypurine runs in genomic DNA.[33,34] This study indicates that there are limits to the specificity of oligonucleotide-EDTA•Fe(II) molecules.

These experiences indicate that *specificity* and *yield* are key issues in the design of molecules capable of and useful for the modification of DNA on the chromosomal level. Recent studies indicate that molecules which have evolved through selection to effect DNA modification, such as the restriction endonuclease EcoR I, derive their specificities not only from their lower binding affinities for mismatched sites, but also due to a loss of hydrolytic activity at such sites.[35] For instance, binding at sites with a single mismatch is

disfavored by 4 to 6 kcal/mol. The specificity of the enzyme couples this decrease in binding affinity with an additional 4 to 7 kcal penalty for transition state destabilization at secondary sites. Thus, restriction enzyme specificity is a composite of recognition and cleavage specificities.[35]

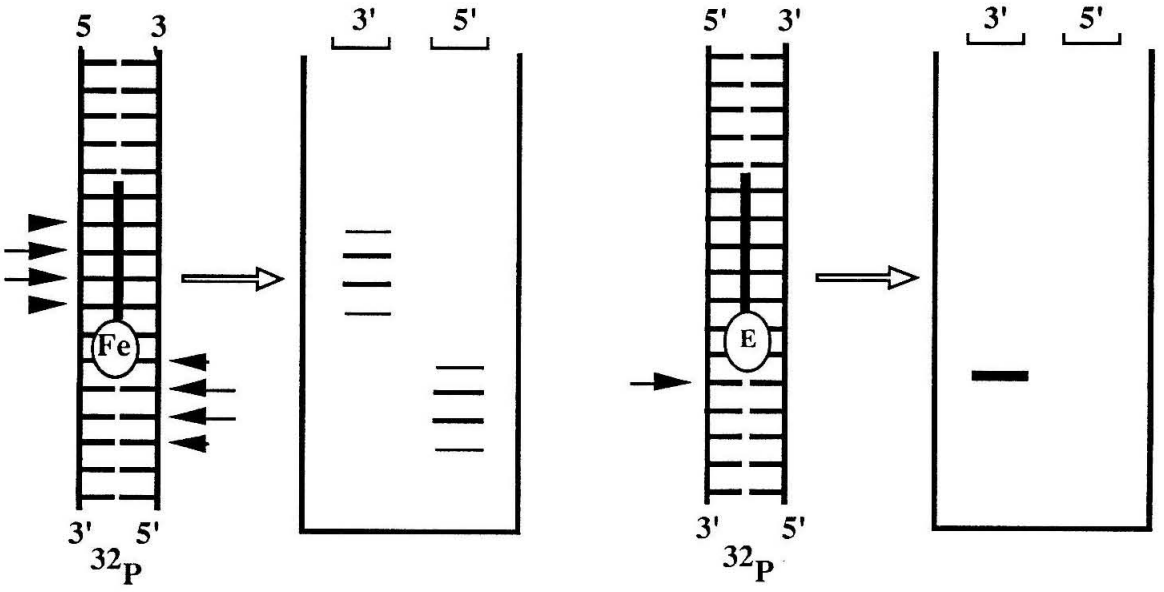


Fig. 4.1. Replacement of a *diffusible nonspecific* DNA cleaving moiety generated by EDTA•Fe(II) (hydroxyl radical), useful for studying DNA recognition (affinity cleaving) to a *non-diffusible base-specific* moiety. This is a key issue with respect to the design of sequence-specific DNA cleaving molecules. Sequence dependent recognition is coupled with sequence dependent cleavage.

Once the structural principles for double strand DNA recognition have been determined, it becomes possible to design and attach structural domains designed to carry out a desired DNA modification (Fig. 4.1). Early examples were the conversion of distamycin-EDTA•Fe(II) to *N*-

bromoacetyldistamycin[36-38] and Fe(II)•EDTA-Hin(139-190) to Cu/Ni•GGH-Hin(139-190).[39,40]

Cleavage of DNA can be achieved by modification of the intact duplex at the sugar (oxidative reaction), phosphate backbone (nucleophilic attack and/or hydrolysis), or at the bases (electrophilic attack) followed by base treatment. Of these, only attack at the bases offers the possibility of *additional sequence discrimination* during the cleavage reaction. Analysis of DNA reactivity data suggests that attack at N-7 of guanine offers the greatest opportunity for *high yield* reaction.[41-45] Methylation by relatively non-specific alkylation agents (methyl methanesulfonate) results in > 89% reaction at N-7 of guanine.[41-43]

The sequence and base specific modification of single stranded DNA has been achieved by complementary oligonucleotides bearing a number of reactive functional groups. Belikova originally proposed the modification of a complementary single stranded DNA using a modified oligonucleotide-electrophile via Watson-Crick duplex formation in 1967.[46] Vlassov and co-workers have targeted single stranded DNA's with platinum derivatives,[47] and an aromatic 2-chloroethylamino electrophile.[48-52] The use of a halo-acetyl moiety was originally proposed by Summerton and Bartlett[53] and has recently been used to modify single stranded DNA in high yield.[54] Modification of a complementary guanine via an cytosine-aziridine moiety[55,56] has been reported. Methylation of DNA by an oligonucleotide-methyl thioether results in exclusive modification at the N-7 position of guanine bases; properly positioned adenines are not alkylated.[57,58] All of these functionalities target the N-7 atom of guanine for electrophilic attack.

More recent work has targeted double helical DNA via triple helix formation using an oligonucleotide linked at the 5' end to an aromatic 2-chloroethylamino moiety.[59,60]

This chapter presents data which indicate that *N*-bromoacetyloligonucleotides are capable of: (i) near quantitative modification of double helical DNA; (ii) at a single base position; (iii) in a manner which produces ends which are "clonable" with compatible ends produced by conventional restriction enzyme digestion. Furthermore, the specificity of *N*-bromoacetyloligonucleotides is sufficient to produce efficient chemical cleavage at a single site within a yeast chromosome. The effective molarity obtained by using organic synthesis to tether a reactive functionality to a DNA binding molecule is measured. Finally, the utility of the *N*-bromoacetyl moiety as a reporter group for studying small changes in oligonucleotide binding affinities is demonstrated in studies concerning the cleavage efficiency of *N*-bromoacetyloligonucleotides as a function of oligonucleotide length, and in studies which quantitate the cooperativity observed by binding of two oligonucleotides to abutting sites.

Part I

Single Strand Alkylation of Double Helical DNA Using *N*-Bromoacetyloligonucleotides

Design of Electrophile-Oligonucleotide Conjugates for the Efficient Base Specific Alkylation of Double Helical DNA.[61]

The judicious attachment of an electrophile allows targeting for alkylation of a specific guanine base adjacent to a triple helix site on double helical DNA. *N*-Bromoacetyloligonucleotides were designed to combine the specificity of triple-helix formation with the requirement for the presence of a properly positioned guanine base. A moderate electrophile, such as a bromoacetyl moiety, was chosen to minimize loss of yield due to competing reactions with solvent, buffers, or other reaction components.

Model building of a triple helical complex indicates that a pyrimidine oligonucleotide bound in the major groove parallel to the purine strand of the duplex could be equipped with a bromoacetyl moiety at the 5'-end such that the electrophile is proximal to a guanine base located two base pairs to the 5'-side of the target sequence (Figures 4.2 and 4.5). Reaction of the electrophile carbon with N-7 of guanine adjacent to the triple helix would result in covalent attachment of the oligonucleotide to the duplex (Fig. 4.2). Upon warming and base treatment, depurination and cleavage of the DNA backbone at the position of alkylation is expected.[62,63]

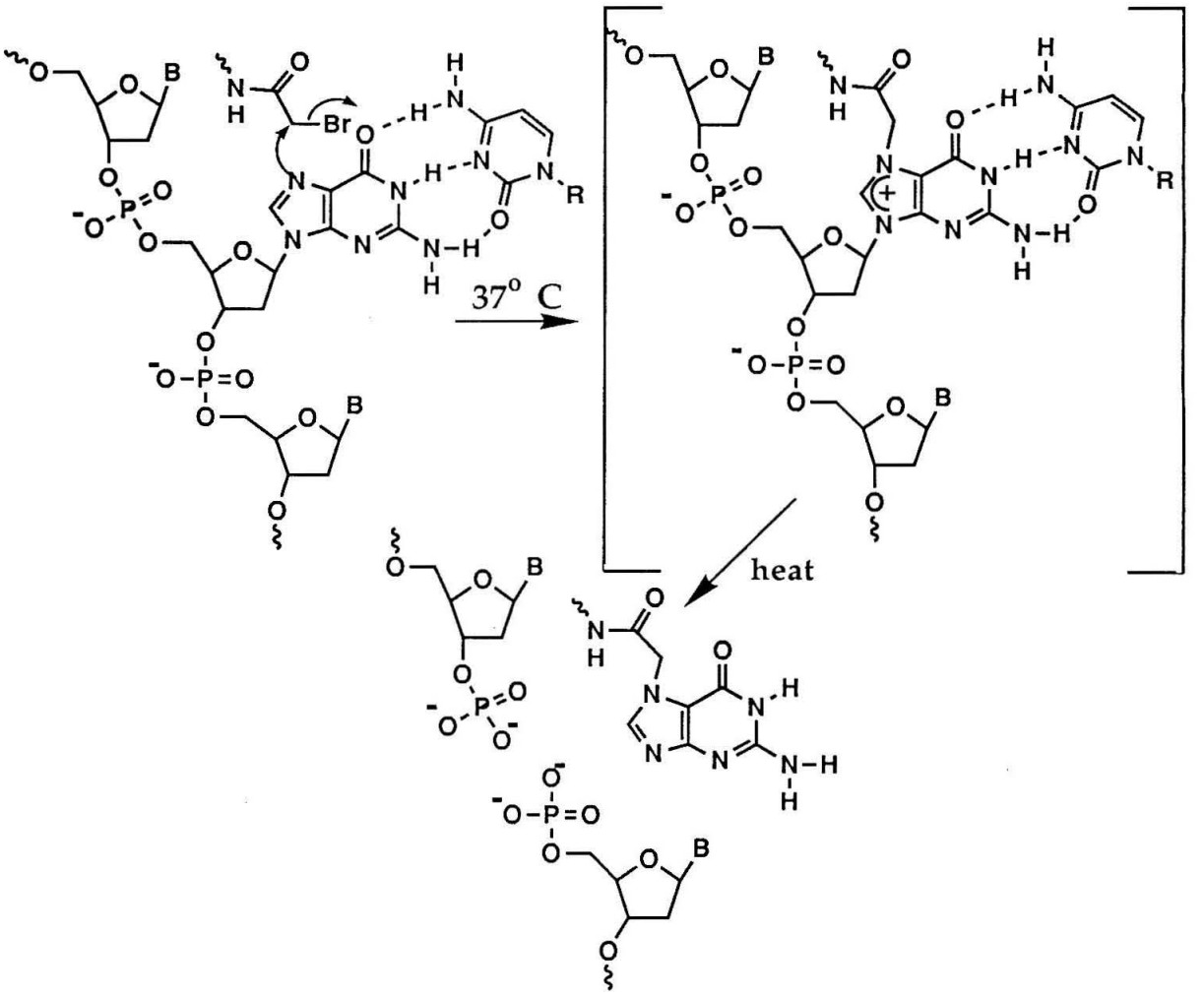


Fig. 4.2. Scheme for alkylation and cleavage of DNA using a moderate electrophile (*N*-bromoacetyl) for the base specific and high yield modification of DNA. *N*-bromoacetyloligonucleotide localizes the electrophilic moiety at the DNA double helix. Reaction results in cross-linking of the oligonucleotide to the duplex. Depurination and loss of the intact modified guanine base from the double helix can be effected by warming. Piperidine treatment effects strand scission. The production of 5' and 3' phosphate ends is expected.

Results and Discussion

Synthesis, derivatization, and purification of N-bromoacetyloligodeoxyribonucleotides. The synthesis of the 5'-O-DMT-3'-phosphoramidite of T^E follows in two steps from the previously described aminonucleoside **1** (Figure 4.3).[64] Protection of the primary amine **1** with fluorenylmethyl chloroformate,[65,66] followed by activation of the 3' alcohol with β-cyanoethyl-*N,N*-diisopropylchlorophosphoramidite,[67] affords a thymidine derivative containing an appropriately protected linker-amine suitable for use in standard automated oligodeoxyribonucleotide synthesis. The modified 5-amino-thymidine was incorporated into the oligonucleotide, 5'-^{E}TT$_3$(Me^{5}C)T$_4$(Me^{5}C)$_2$T$_3$(Me^{5}C)T$_4$-3', as the 5' terminal residue by coupling of phosphoramidite **3**. Deprotection of the amine occurs quantitatively under standard deprotection conditions (concentrated NH_4OH, 55 °C, 20 hrs.),[65,66] as judged by reverse-phase and anion-exchange HPLC and polyacrylamide electrophoresis. The deprotected oligonucleotide **4** was purified by polyacrylamide gel electrophoresis. Alternate amine protecting groups compatible with automated oligonucleotide synthesis (monomethoxytrityl (MMT) and *N*-trifluoroacetyl) can also used to protect the amine in nucleosides **2** and **3**.

Derivatization of the oligonucleotide was effected via reaction with the *N*-hydroxysuccinimidyl bromoacetate.[68-70] Purification of the derivatized oligonucleotide was initially accomplished using reverse-phase HPLC .[71-73] The buffers used for such purification consist of a water/acetonitrile gradient buffered with small concentrations of triethylammonium or ammonium acetate, which can easily be removed by lyophilization.[71-73] Due to the

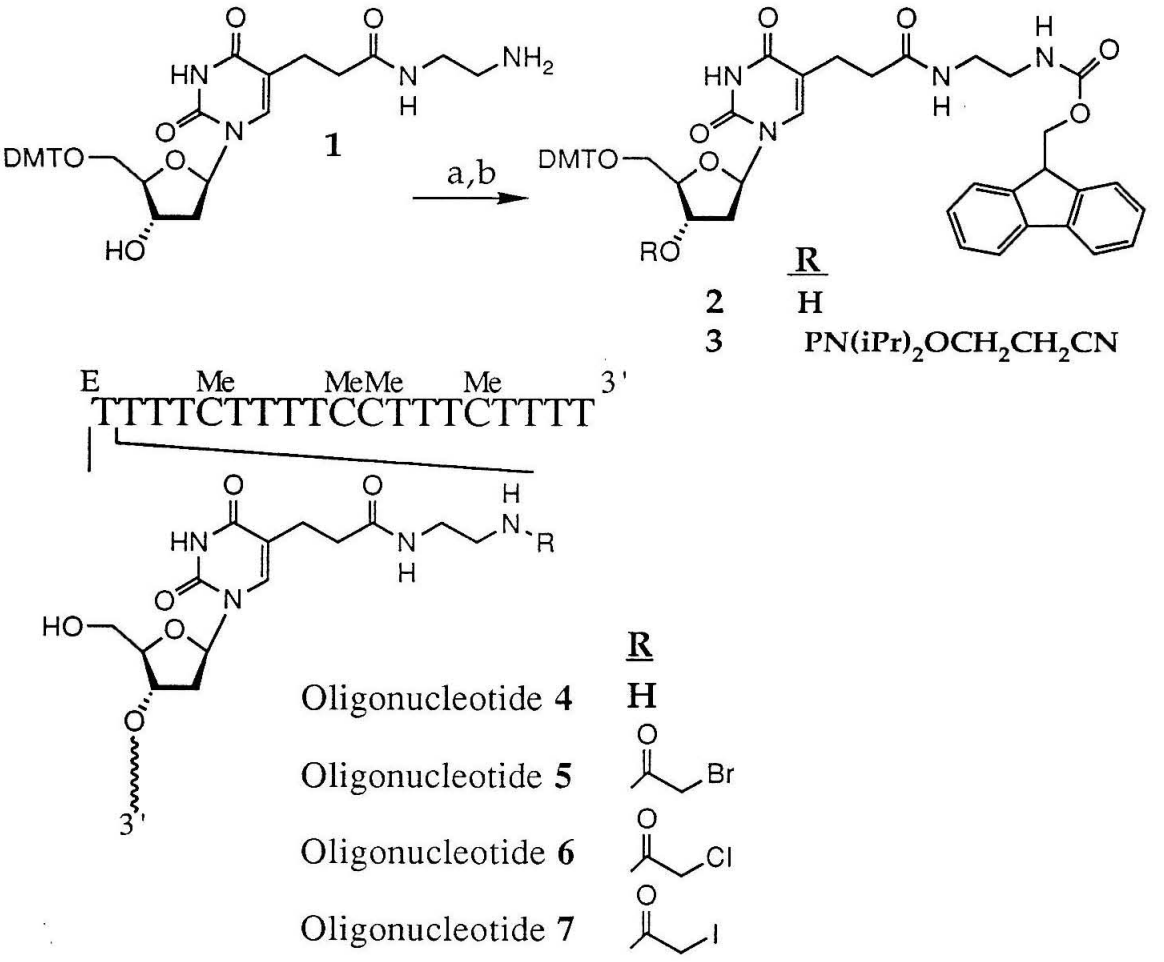


Fig. 4.3. *(Top)* Synthetic scheme for construction of Fmoc-protected thymidine 2-cyanoethyl-*N*,*N*-diisopropylphosphoramidite **3** used for the automated synthesis of oligonucleotides with a primary amine attached to the 5-position of a single thymidine position from the known amine **1** in two steps. *(Bottom)* Oligonucleotides **4** through **7** containing T, Me^5C, and a unique thymidine at the 5'-end modified at the 5-position with a spacer-primary amine (**4**), spacer-*N*-bromoacetamide (**5**), spacer-*N*-chloroacetamide (**6**), and spacer-*N*-iodoacetamide (**7**).

possibility of reaction between the bromoacetyl moiety and ammonia, the buffer was acidified to pH 4.7. Initial studies were conducted with *N*-bromoacetyloligonucleotides purified in this manner. In separate studies, it was observed that the bromoacetyl moiety purified by reverse phase HPLC

undergoes degradation to an unidentified product (*vide infra*). In later studies, *N*-bromoacetyloligonucleotides were purified by anion exchange chromatography. This requires additional treatment after HPLC purification to obtain salt-free DNA; however, after the implementation of this purification procedure, the yields of alkylation increased slightly.

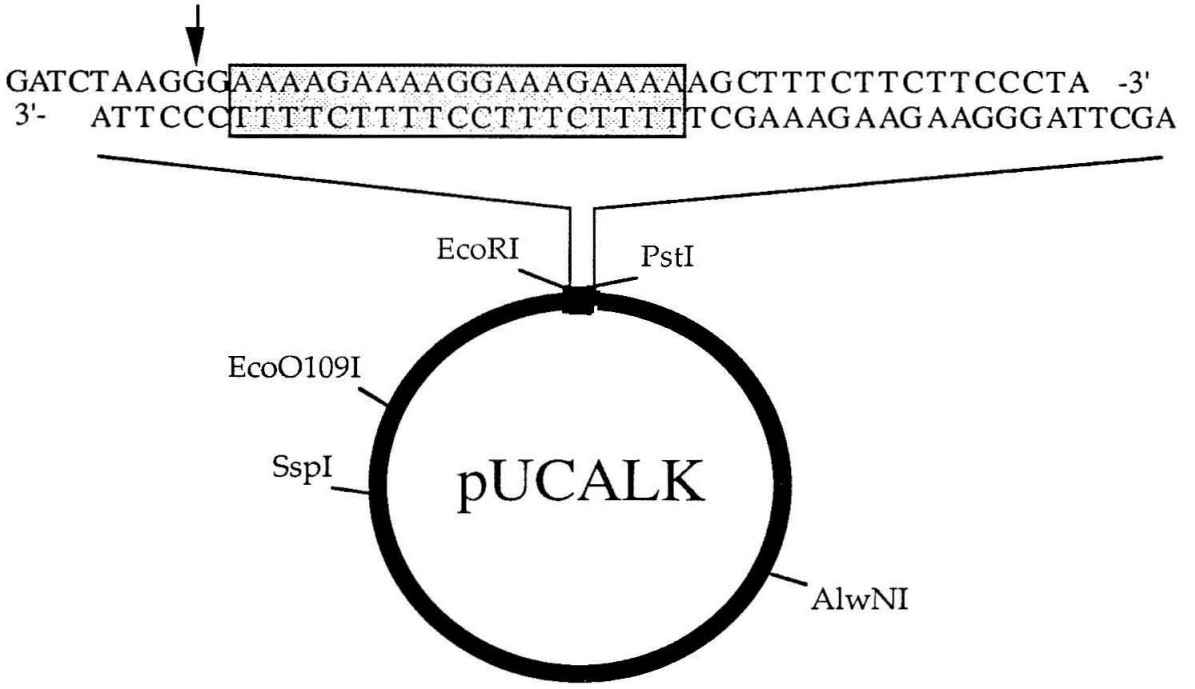
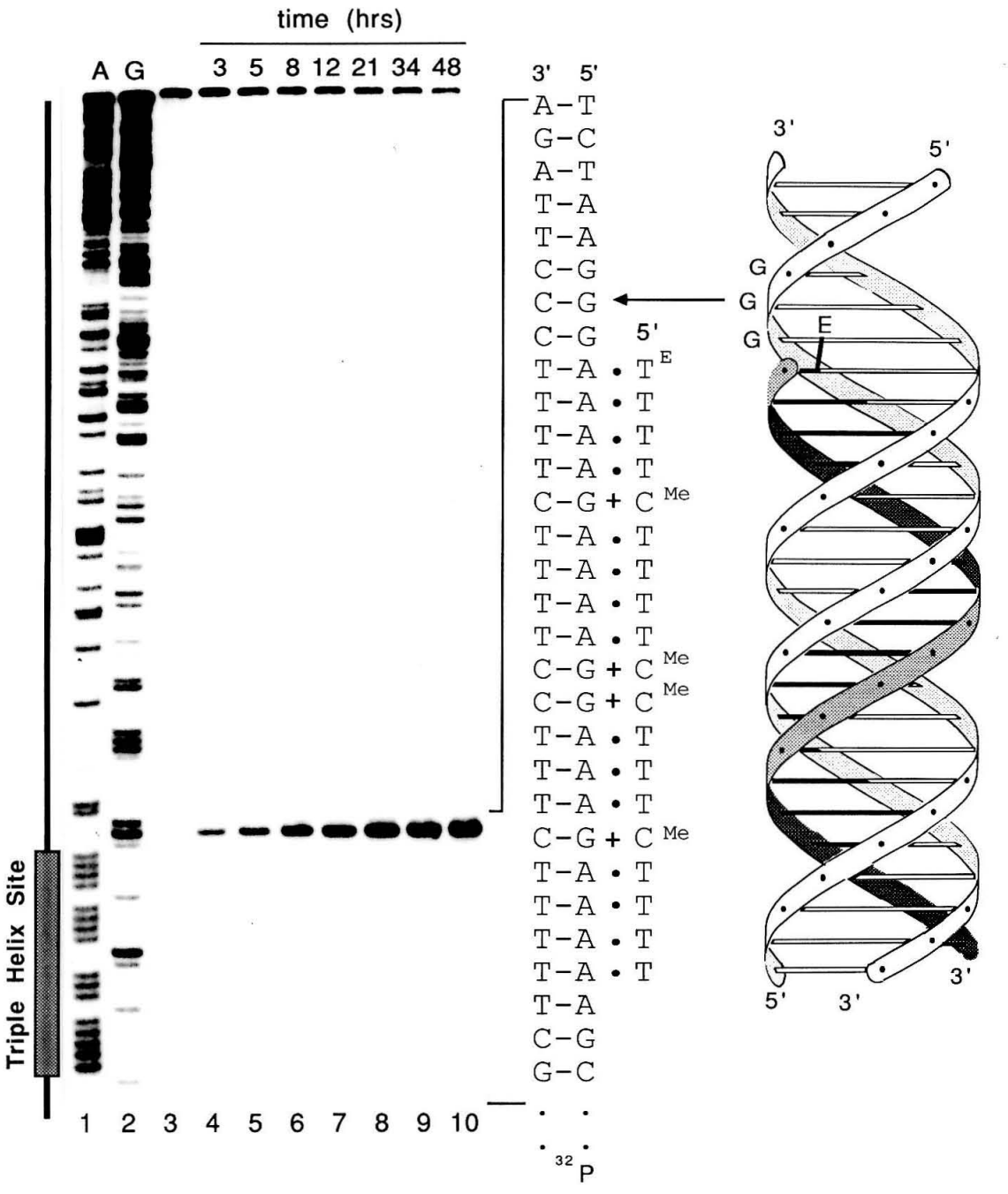


Fig. 4.4. Construction of the plasmid pUCALK from pUC19 via ligation of the oligonucleotides indicated into the BamH I / Sal I sites within the polylinker region. Binding site for the *N*-bromoacetyloligonucleotide **5** is boxed, and the targeted guanine base is in indicated. Restriction endonuclease sites used for the isolation of desired restriction fragments are indicated.

Specificity and yield of reaction. To test both the specificity and yield of the alkylation reaction, three consecutive G•C base pairs were incorporated immediately 5' to the 19 base pair purine target site, 5'-$A_4GA_4GGA_3GA_4$-3', within a plasmid 2.3 kbp in size (pUCALK) (Fig. 4.4).

Fig. 4.5. Autoradiogram of a high-resolution 8% denaturing polyacrylamide gel of cleavage products from the reaction of oligonucleotide **5** and a ^{32}P 3'-end-labeled 659-bp restriction fragment (*Pst I/ Ssp I*). Reaction conditions were 1 μM concentration of oligonucleotide **5**, 20 mM Hepes, pH 7.4, 0.8 mM $Co(NH_3)_6^{+3}$, and 10,000 cpm end-labeled DNA in a total volume of 15 μl. Reactions were incubated at 37 °C, precipitated with NaOAc/EtOH, washed with 70% EtOH, and treated with 0.4 M piperidine (90 °C, 30 min.). Cleavage products were analyzed on an 8% 1:20 cross-linked, 48% urea polyacrylamide gel, 0.4 mm thick. Lanes 1 and 2 are A-^{74} and G-^{62,63} specific chemical sequencing reactions, respectively. Lane 3 contains DNA incubated for 48 hours in the absence of *N*-bromoacetyloligonucleotide **5**. Lanes 4-10: DNA incubated in the presence of *N*-bromoacetyloligonucleotide for 3, 5, 8, 12, 21, 34 and 48 hours. Center: Sequence of the oligonucleotide-DNA triplex within plasmid pUCALK. The major site of modification is indicated. Right: ribbon model of triple-helical complex with the oligonucleotide bound parallel to the purine strand in the major groove of the duplex DNA. The positions of the electrophile and the three guanine bases proximal to the binding site are indicated. Arrow indicates base position of predominant cleavage.


time (hrs)
A G
3 5 8 12 21 34 48
Triple Helix Site
1 2 3 4 5 6 7 8 9 10
3' 5'
A-T
G-C
A-T
T-A
T-A
C-G
C-G
C-G
T-A • T
T-A • T
T-A • T
T-A • T
C-G + C Me
T-A • T
T-A • T
T-A • T
T-A • T
C-G + C Me
C-G + C Me
T-A • T
T-A • T
T-A • T
C-G + C Me
T-A • T
T-A • T
T-A • T
T-A • T
T-A
C-G
G-C
5'
E
32 P
3'
5'
G
G
G
E
5'
3'
3'

N-Bromoacetyloligonucleotide **5** (1 μM) was allowed to react with a 659 base pair Pst I/Ssp I ^{32}P end-labeled restriction fragment at 37 °C in the presence of 0.8 mM $Co(NH_3)_6^{+3}$ and 20 mM Hepes buffer (pH 7.4). The reaction is allowed to proceed for the desired period at which time piperidine treatment is used to effect strand scission at modified bases. The extent of cleavage of the labeled restriction fragment by the *N*-bromoacetyloligonucleotide **5** was followed for 48 hours at 37 °C (Figure 4.5). Yields and specificity were determined by phosphorimaging of high resolution polyacrylamide gels.

Analysis by high resolution gel electrophoresis reveals that reaction occurs predominantly at the guanine base located two base pairs to the 5'-end of the target site in 87% yield. No modification occurs on the pyrimidine strand. Modification occurs at the flanking guanines with relative rates of 0.03 and 0.06 that of the major site of reaction. Apparently, there exists sufficient flexibility in the linker arm and/or the junction of the local triple-helical complex to access all three guanine bases for modification to some extent.

Maximal cleavage conditions. Not unexpectedly, maximal modification/cleavage is observed when the presence and activity of nucleophiles capable of reaction with the bromoacetyl moiety is limited. The presence of spermine and Tris buffers results in a slight decrease in the yield of DNA cleavage; therefore, the polycation $Co(NH_3)_6^{+3}$ and Hepes buffers have been substituted. To study the reactivity of the bromoacetyl group in solution, the dinucleoside compound T-T^E (**8**) was synthesized (Fig. 4.6). After derivatization of **8**, **9** was incubated with $Co(NH_3)_6^{+3}$ and Hepes at 37 °C and 55 °C. HPLC analysis indicates that no degradation of **9** occurs in an

aqueous 1 mM $Co(NH_3)^{+3}$ solution; however, degradation of **9** does occur upon incubation in Hepes buffer at 55 °C. The high yield of DNA modification indicates that this reaction is a minor one at 37 °C, at least with oligonucleotide bound to the double helix.

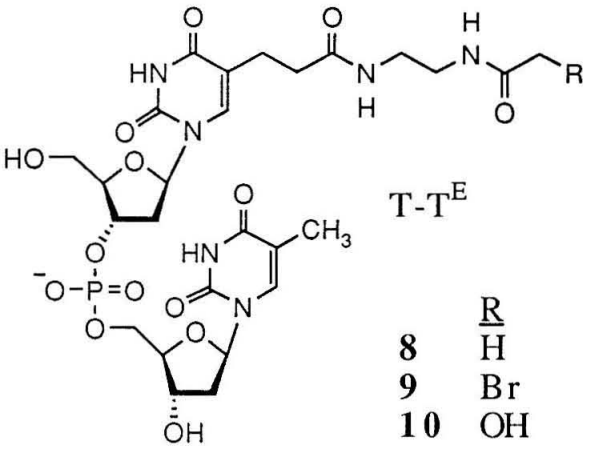


Fig. 4.6. T-T^E **8-10** synthesized to study the reactivity of the bromoacetamide functionality under DNA alkylation conditions.

Independent synthesis of the expected hydrolysis product **10** indicates that resolution of **9** and **10** is achieved under the HPLC conditions used. No hydrolysis of **9** to form **10** is observed upon incubation in aqueous solutions at 37 °C or 55 °C.

One explanation for the decreased yields observed in the presence of spermine is reaction between the tetraamine and the bromoacetyl moiety. Interestingly, spermine inhibits the methylation of DNA by *N*-methyl-*N*-nitrosourea.[75] An alternative explanation, as proposed by Zakrzewska and Pullman, for the decrease in yield of alkylation observed in the presence of spermine, involves binding of spermine to DNA, resulting in either a change in the electrostatic potential of the major groove or steric blockade of

electrophilic sites in the major groove. A similar mechanism could inhibit reaction of *N*-bromoacetyloligonucleotides with DNA.

The T-T^E dinucleoside **9** was also used to study various HPLC purification conditions. Isolation of the purified dinucleoside by reverse phase HPLC, lyophilization of the compound, and re-injection indicates the presence of an impurity. The impurity represents about 10% of the total material. This impurity is postulated to result from attack of ammonia or acetate on the bromoacetyl moiety during the lyophilization period.

Determination of Rate of Reaction of N-haloacetyloligonucleotides with DNA. Reaction of the *N*-bromoacetyloligonucleotide **5** within a triple helical complex follows a first order rate equation given by :

$$[DNA]_{intact} = [DNA]_{initial}e^{-kt}$$

where k is the unimolecular rate constant, and t is time.

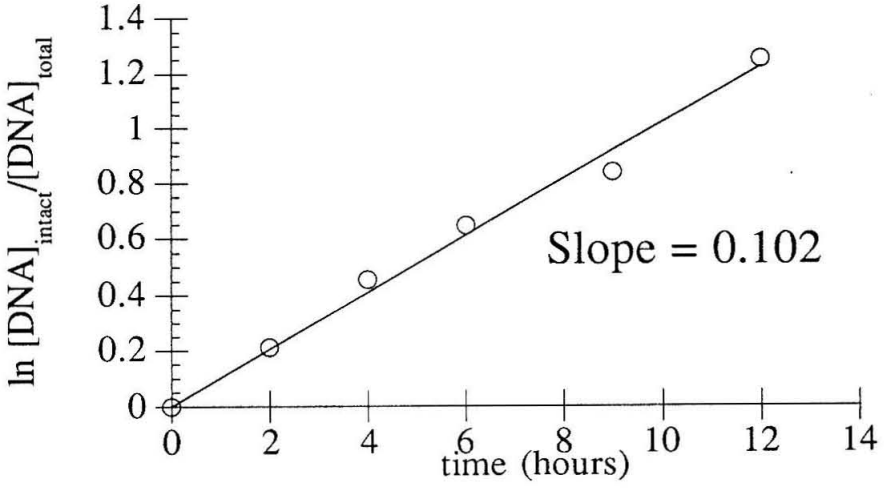


Fig. 4.7. Plot of decrease of intact DNA in the presence of *N*-bromoacetyloligonucleotide **5** as a function of time.

A plot of ln ($[DNA]_{intact}/[DNA]_{total}$) vs. time (pseudo-first order conditions, where micromolar concentrations of oligonucleotide and nanomolar concentrations of duplex insures formation of the triple helical complex) indicates that the reaction between N-bromoacetyloligonucleotide **5** and the double-helical DNA is first order in target DNA concentration with a first order rate constant of 3.1×10^{-5} s^{-1} at 37 °C, corresponding to a half-life for alkylation within the triple helical complex of 6.2 hours (37 °C) (Fig. 4.7). Identical rates are obtained over the range of *N*-bromoacetyloligonucleotide concentrations 50 nM - 2 μM.

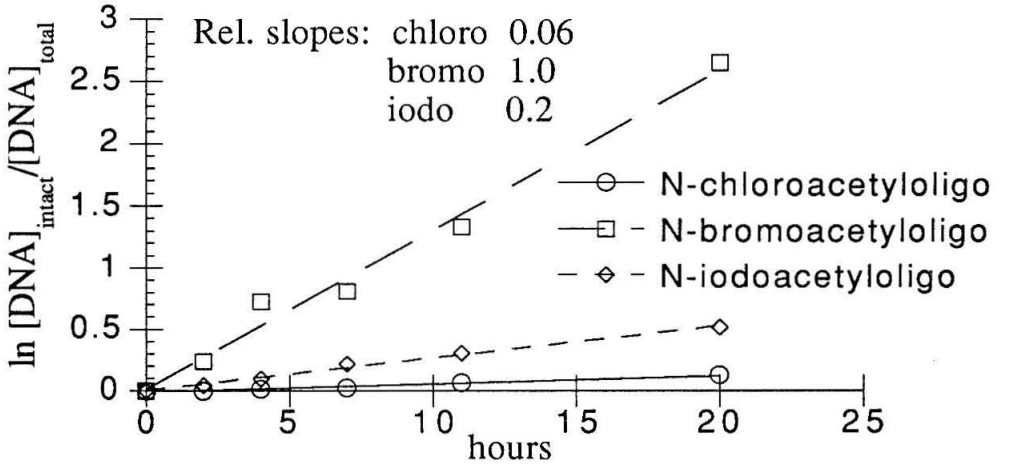


Fig. 4.8. Plots of decrease in amount of intact DNA in the presence of haloacetyloligonucleotides **5, 6,** and **7**.

N-Chloroacetyl- and *N*-iodoacetyloligonucleotides **6** and **7** were synthesized by derivatization of amino-oligonucleotide **4** with *N*-hydroxysuccinimidyl chloro- and iodo- acetates. Determination of the first-order rate constants for the reaction of these oligonucleotides with the double helical target site indicates that these moieties react with relative rates of $k_{iodo}/k_{bromo} = 0.2$ and $k_{chloro}/k_{bromo} = 0.06$ (Fig. 4.8). The slower rates of

reaction for both the chloroacetyl and iodoacetyl derivatives parallel the relative rates observed at N-3 of adenine with the reactions of *N*-bromo-, chloro- and iodoacetyldistamycins bound in the minor groove of double helical DNA.[38]

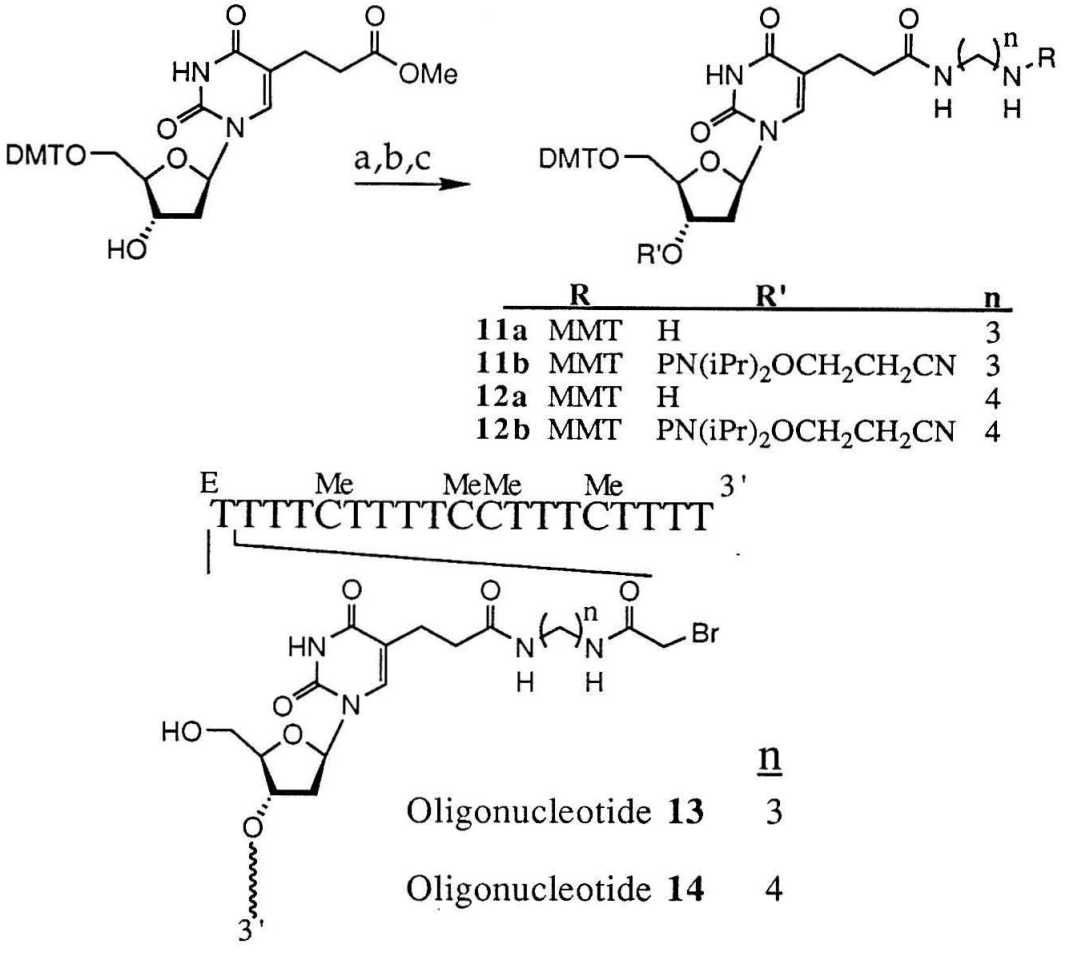


Fig. 4.9. Synthesis of *N*-bromoacetyloligonucleotides **13** and **14** was accomplished via incorporation of phosphoramidites **11** and **12** using automated oligonucleotide synthesis. Phosphoramidites **11b** and **12b** were synthesized by: a) 5-(2-carbomethoxyethyl)-5'-ODMT-2'-deoxyuridine[64] was stirred in 1,3 diaminopropane or 1,4 diaminobutane for 2 days; b) the primary amine was protected with p-anisylchlorodiphenylmethane; and c) the 3' hydroxyl was activated as the phosphoramidite using standard procedures.[67] Synthesis of the oligonucleotides was accomplished using standard automated synthesis. Derivatization was accomplished as previously described with *N*-hydroxysuccinimidyl bromoacetate, and purification with reverse-phase HPLC to yield oligonucleotides **13** and **14**.

Optimization of linker arm. Design of the linker arm in **5** was based on model building studies with the DNA double helical target in an A' configuration. Since the exact structure of the DNA triplex as well as the triplex to duplex junction remains to be elucidated, two additional linker arms, containing one and two additional methylene groups, were studied (Fig. 4.9). Oligonucleotides **13** and **14** were synthesized from the phosphoramidites **11** and **12** and derivatized as before by reaction with N-hydroxysuccinimidyl bromoacetate followed by reverse phase HPLC purification.

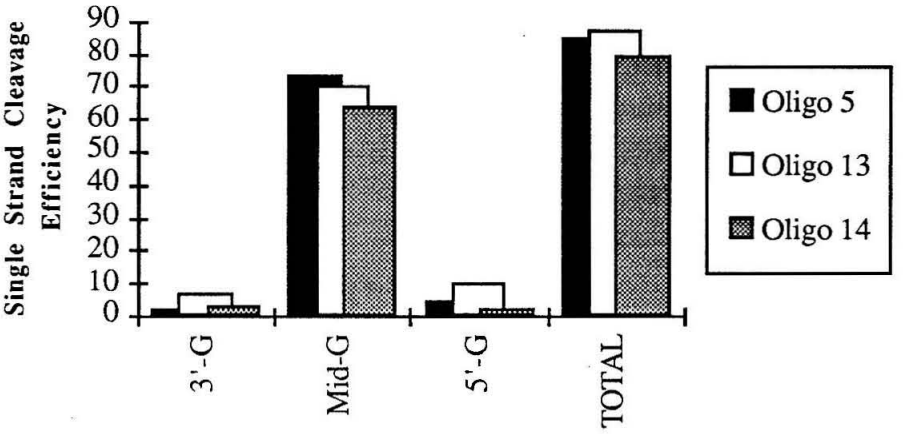


Fig. 4.10. Cleavage efficiencies of the *N*-bromoacetyloligonucleotides 5, 13 and 14 at each of the three guanines 5' to the triplex target site. The 3' guanine is the base abutting the triplex site; the 5' guanine is 3 bases removed from the triplex binding site.

The total cleavage efficiencies and specificities of the *N*-bromoacetyloligonucleotides **5**, **13** and **14** are plotted in Figure 4.10. Incubation of the DNA with *N*-bromoacetyloligonucleotide **13** results in the

largest overall cleavage efficiency; however, oligonucleotide **5** affords the highest cleavage efficiency at the central guanine and is the most specific in terms of reactivity at the central guanine vs. reaction at the flanking bases. These data are drawn from a single gel, and slight differences in oligonucleotide purity might affect the results obtained. The error in these measurements must be considered to be large relative to the deviation in the results; therefore, these experiments simply indicate that single methylene changes in the 9 atom long linker arm do not result in large changes in reactivity or specificity.

Modelling of linker arm. A computer illustration depicting a 6 base pair triple helix with a terminal thymidine-bromoacetyl is shown in Figure 4.11. Since the exact structure of a triple helix or a triplex-duplex juncture is unknown, this model is intended to impart descriptive, not quantitative, information. Indeed, footprinting data suggest that triplex-duplex junctions result in some DNA distortion (see chapter II). The figure allows a pictorial understanding of the design of the linker arm and its relationship in space with the three strands present in the complex.

The triple helix was constructed from coordinates obtained from fiber diffraction data.[76] Three guanine bases were incorporated 5' to the triplex site; thus, this sequence represents that found in the plasmid pUCALK. The linker arm was constructed such that it extends in the direction of the targeted guanine bases, and was subsequently energy minimized with the N-7 of the targeted guanine, the electrophilic carbon, and the bromine atom defining an angle of 180°.

Analysis of the structure indicates that the torsional angle around the Br-C-C-O bond is 88°, close to the ideal value of 90° required to insure maximum orbital overlap during the displacement reaction. Likewise torsional angles

along the linker arm chain are within 10° of ideal values, indicating that the arm does not have to assume a strained conformation to allow alkylation of the DNA. The extended linker arm is ideally positioned to access the targeted guanine base two base pairs removed from the binding site, although the exact structure of the triplex-duplex junction is not known and it could differ substantially from the depicted model.

The model suggests that geometrical considerations make the guanine two bases removed from the triplex target an ideal base for intracomplex reaction. The positioning of this base makes it the dominant site of reaction with all of the *N*-bromoacetyloligonucleotides constructed (oligonucleotides **5, 13,** and **14**. The curvature of the helix makes access of guanine bases farther removed from the triple helix site difficult, since the base is both 3.4 Å farther away in a vertical direction and rotated another 30° around the helical axis. Reaction with the distal guanine base requires tilting of the base in the direction of the bromoacetyl moiety, and extension of the linker arm.

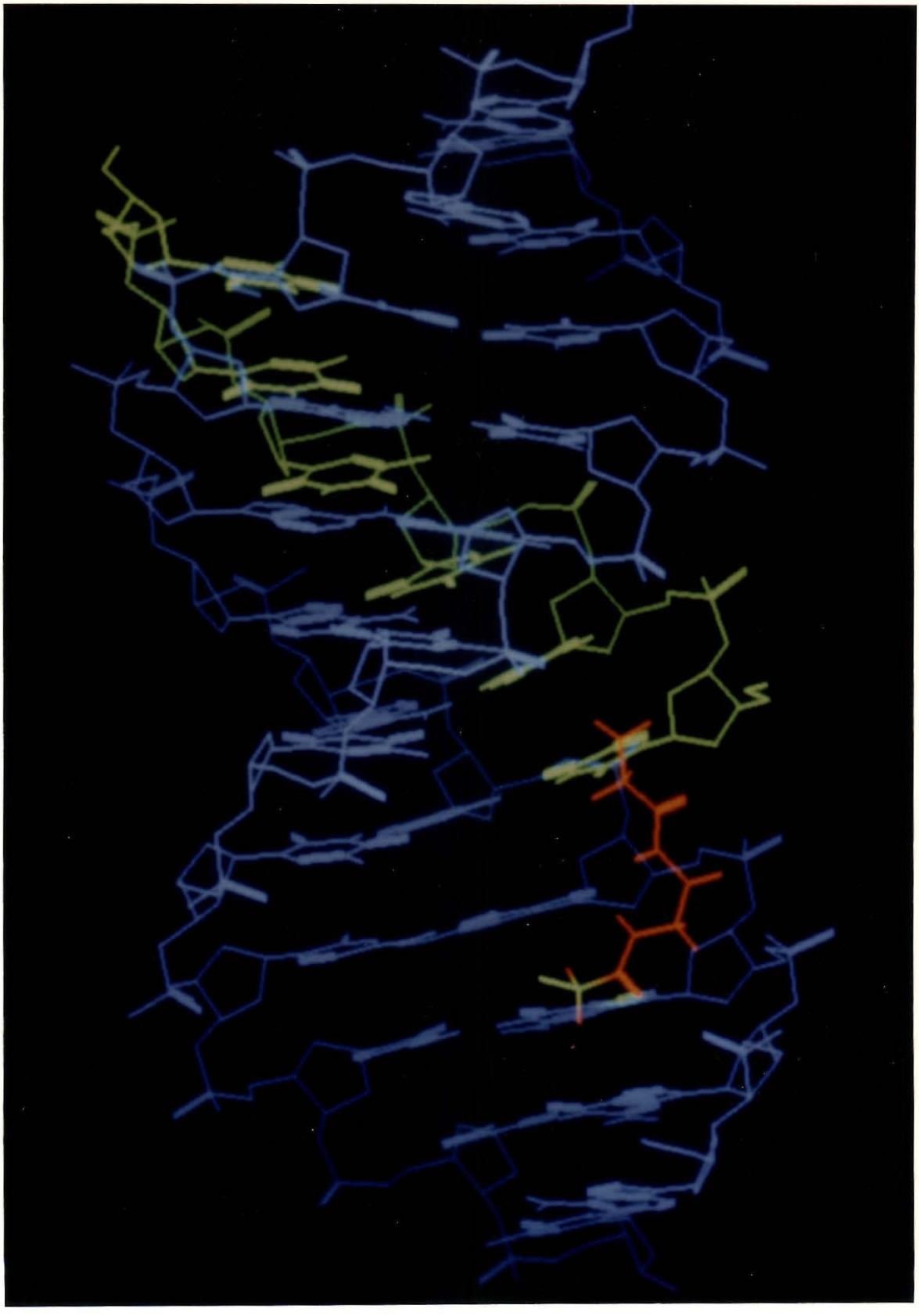
The guanine base immediately adjacent to the triple helix site has actually not rotated sufficiently to be accessible for reaction. Access to this base would require major alterations in the geometry of the linker arm and in the DNA target. The alkylation data suggest that this is highly unfavorable.

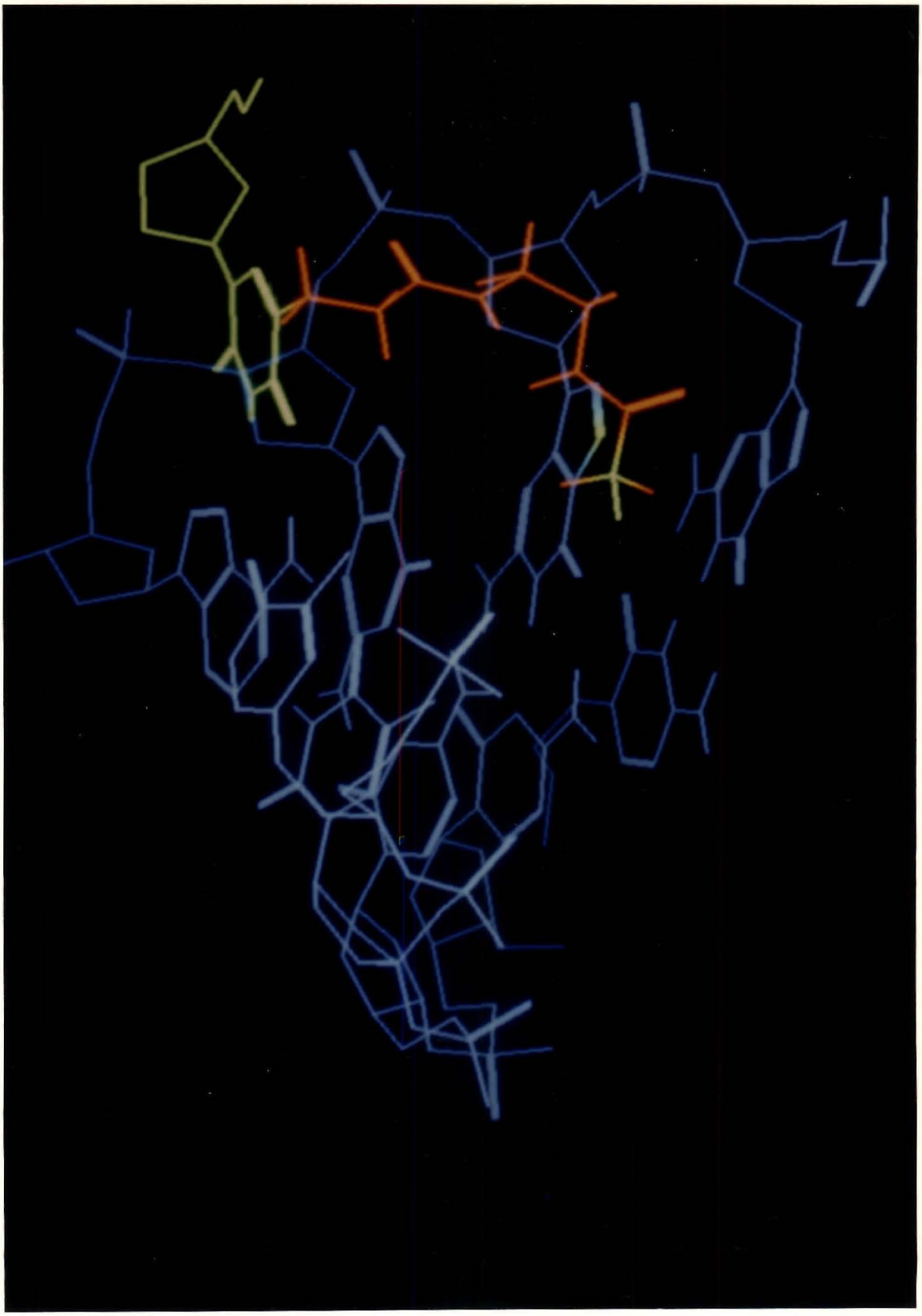
A guanine base two bases removed from the target site is almost directly "below" the modified thymidine, due to the placement of the third strand directly above the purine strand of the duplex.

The linker arm is not able to alkylate the other strand of the target duplex, due to the location of this second strand across the groove of the DNA.

Fig. 4.11a. Model of a 6 base triple helix containing a 5' terminal bromoacetyl-thymidine which is positioned for alkylation of a guanine base 2 base pairs to the 5' side of the triple helical target site. The triple helix was constructed using published coordinates. The linker arm was constructed, a constraint was placed on the geometry of the N7-C-Br angle (180°), and the linker arm was energy minimized. The Watson-Crick double helix is in blue, The third strand is green, and the linker arm is red. The N-7 of guanine, the electrophilic carbon, and the bromine leaving group are highlighted in green.

Fig. 4.11b. Close-up view of the final base triplet and the linker arm. The view has been tilted to afford a view of the guanine base. The highlighted N7 is poised for reaction with the bromoacetyl moiety.





Conclusion

This work demonstrates that a non-diffusible electrophilic moiety judiciously attached to the 5'-end of an oligonucleotide is capable of alkylation of intact double-helical DNA at a single base position in high yield. The half-life for the triple helical complex is 6.2 hours. Predominant modification occurs at a guanine two bases 5' to the triplex target site, with modification of guanines flanking this position with rates 0.03 and 0.06 that of the central guanine. Modification by *N*-chloro- and *N*-iodoacetyloligonucleotides occurs with rates 0.06 and 0.20 that of *N*-bromoacetyloligonucleotides. Changes in the linker arm do not result in large changes in the yield or specificity of the alkylation reaction.

Since oligonucleotide-directed triple-helix formation is sufficiently generalizable and specific to allow the recognition of single sites in genomic DNA,[25-27,77] modification of a single base within megabase-size chromosomes using strictly chemical methods might be possible.

Part II

Double Strand Alkylation and Cleavage of DNA Using *N*-Bromoacetyloligonucleotides

The manipulation of DNA is made possible by the high selectivity and cleavage efficiency afforded by restriction endonucleases, which allow the cleavage of DNA at predetermined sequences to be accomplished in quantitative yield in a manner which allows religation of the products. The use of naturally occurring endonucleases is largely limited to plasmid or cosmid size DNA (10^3 - 10^5 base pairs), due to the limited sequence specificity (4-8 base pairs) of these entities.

The high efficiency and specificity of the alkylation cleavage chemistry developed would offer a number of advantages to DNA manipulations if double strand modification could be achieved. The high efficiency and base specificity of cleavage mimic those of restriction enzymes. The >15 base pair specificity possible using the triple helix motif makes the cleavage at single sites within large chromosomal DNA realizable.

Two motifs for the double strand cleavage of DNA using *N*-bromoacetyloligonucleotides were considered.

Double Strand Alkylation/Cleavage using Cross-over N-Bromoacetyloligonucleotides.

The synthesis and DNA cleavage of a 5'-purine-pyrimidine-3' sequence was accomplished by Horne and Dervan using an oligonucleotide containing

an internal 3'-3' phosphate linkage.[21] This oligonucleotide recognizes purine stretches on each of the two strands of the DNA by crossing-over the major groove at the purine-pyrimidine junction. Oligonucleotides of this type contain two 5' ends.[21] Incorporation of *N*-bromoacetyl-thymidines at each end results in an *N*-bromoacetyloligonucleotide able to bind 18 base pairs of double helical DNA and alkylate each strand 22 base pairs apart.

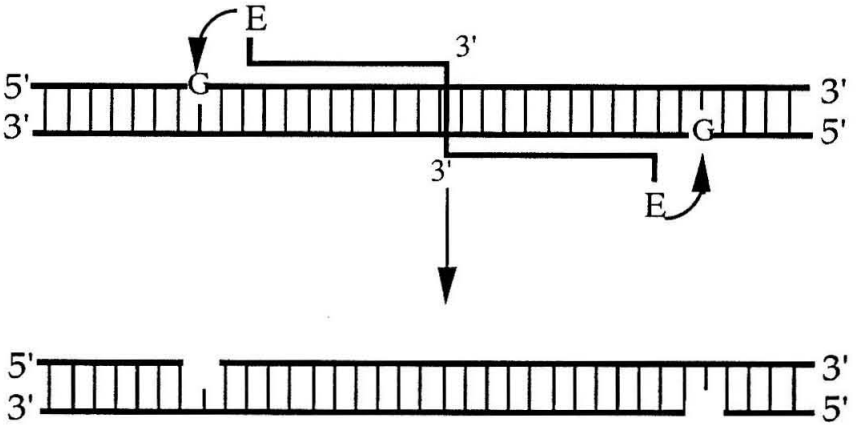


Fig. 4. 12. Motif for the cleavage of double helical DNA via an alkylation/strand scission mechanism using cross-over oligonucleotides of the type described by Horne and Dervan.[21] Cross-over oligonucleotides recognize sequences of the type 5'-purine-pyrimidine-3'. Oligonucleotides containing amino-thymidines at each 5' end are derivatized to afford an oligonucleotide containing 2 5'-*N*-bromoacetyl moieties. The cleavage sites on the individual strands are separated by 22 base pairs.

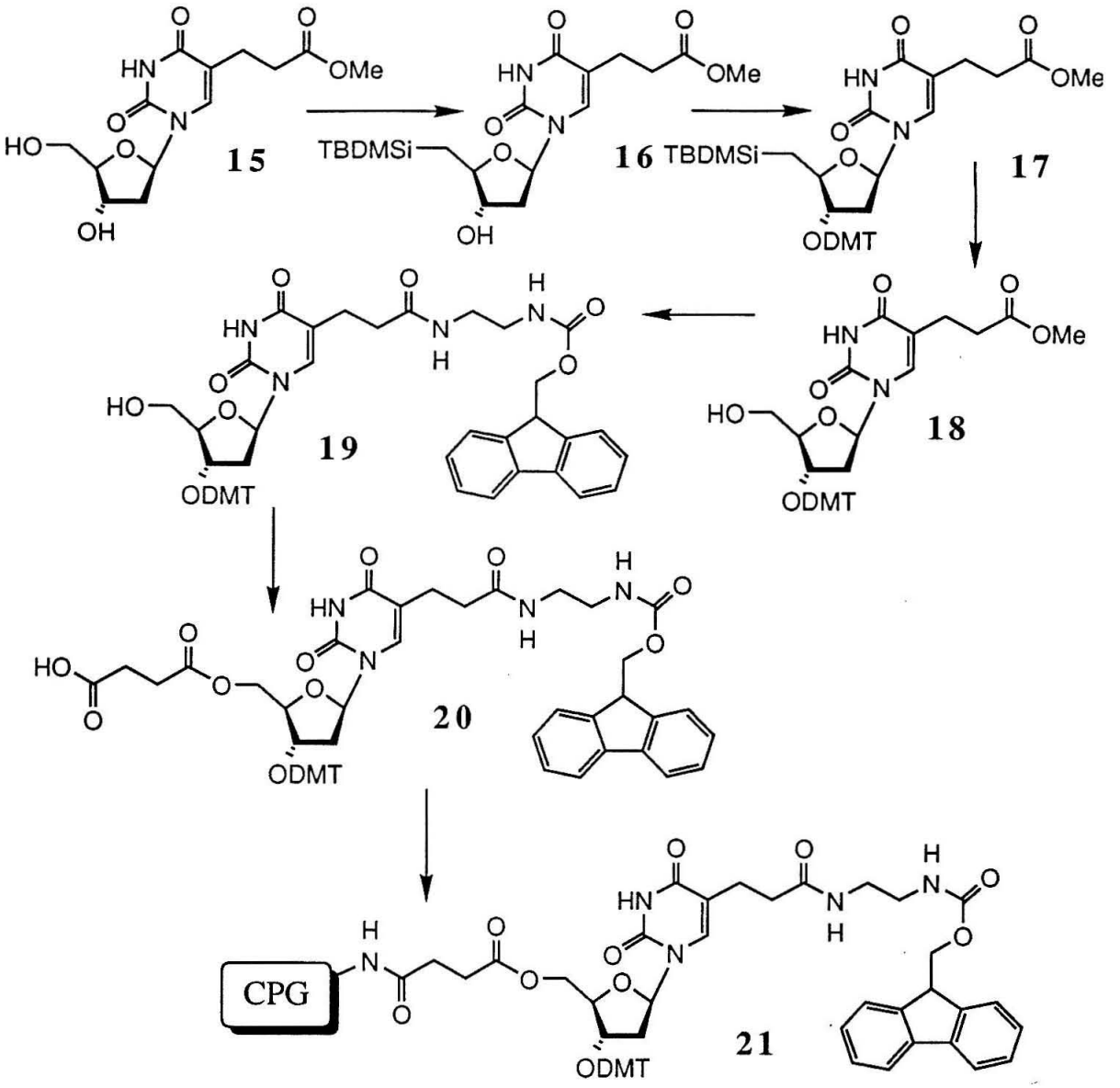


Fig. 4. 13. Synthetic scheme for the synthesis of CPG-Alk-2 solid support for the synthesis of cross-over *N*-bromoacetyloligonucleotides of the type described by Horne and Dervan for the cleavage of double helical DNA via alkylation/strand scission mechanism.

The construction of bis-alkylating cross-over oligonucleotides requires the synthesis of nucleoside **3** attached to a solid support via its 5' hydroxyl (Fig. 4.13). This was accomplished using standard procedures[21,,67] and involves protection of the 3' hydroxyl with a dimethoxytrityl protecting group, attachment of a succinate linker to the 5' hydroxyl, and derivatization of solid CPG-NH_2 support using an activated 5'-p-nitro-phenolate ester (Fig. 4.13). Synthesis of the first half of the oligonucleotide in the reverse 5'-3' direction, coupling of an abasic linker[78,79] as described by Horne and Dervan,[21] and coupling of the remainder of the nucleosides in the conventional 3' to 5' direction affords oligonucleotide **23** (Fig. 4.14).

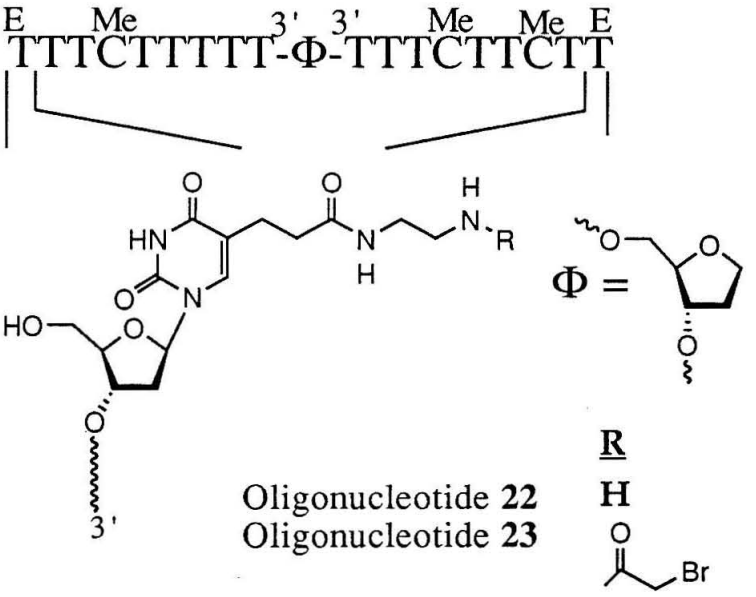


Fig. 4. 14. Oligonucleotides **22** and **23** for the study of double stranded DNA cleavage via alkylation/strand scission mechanism using cross-over oligonucleotides of the type described by Horne and Dervan.[21] Oligonucleotide **22** is converted to the *N*-bromoacetyloligonucleotide **23** as previously described. Purification was effected by HPLC using a Vydac anion exchange column.

The plasmids pUCALK and pUCLEU2D each contain binding sites for cross-over oligonucleotide **23** with guanines positioned on each strand two bases 5' to the end of each purine run (Fig. 4.15). The cross-over oligonucleotide binds the underlined bases, using a 3' - 3' internal phosphate linkage to change polarity and cross-over the major groove at the purine-pyrimidine junction.

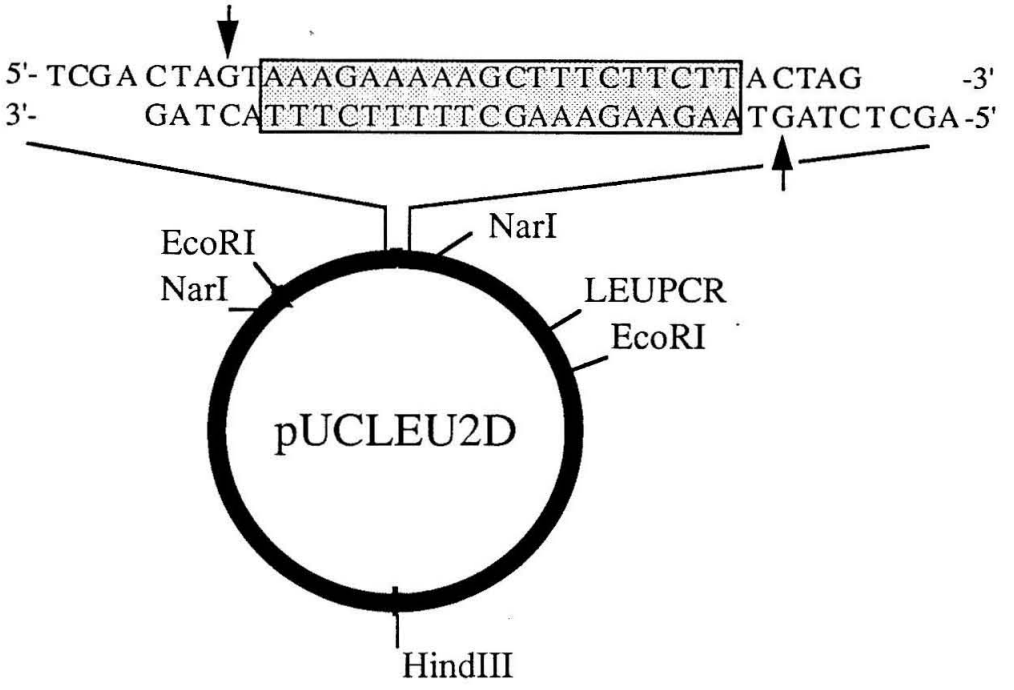


Fig. 4.15. Construction of the plasmid pUCLEU2D from pUCLEU2 via ligation of the oligonucleotides indicated into the Xho I site. Binding site for the *N-bromoacetyloligonucleotide* **23** is boxed, and the targeted guanine bases are indicated. Restriction endonuclease sites used for the isolation of desired restriction fragments are indicated.

Incubation of oligonucleotide **23** with Hind III linearized ^{32}P-labeled plasmid pUCLEU2D indicates that alkylation of each of the guanines can be

accomplished in high yield (>80%). The yield of double strand cleavage was low. Plasmid pUCLEU2D was constructed to allow the insertion of this sequence into yeast chromosome III. Cleavage of yeast chromosome III by oligonucleotide **23** affords yields of 10%. Due to difficulties involved with the handling of chromosomes, it is not possible to manipulate large DNA at temperatures above 55 °C. It is possible that higher cleavage efficiencies are achieved, but that due to the large overhang (22 bases) produced, denaturation and strand separation at this site does not occur. Due to the low cleavage yields, no further studies with this motif were undertaken.

Double Strand Alkylation/Cleavage Using Two *N*-Bromoacetyloligonucleotide Binding Sites.

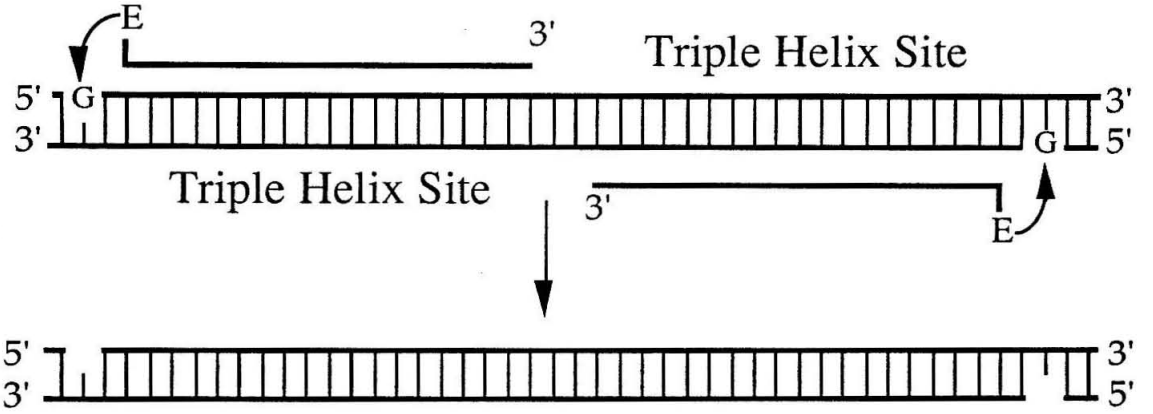


Fig. 4. 16. Strategy for the cleavage of double helical DNA via an alkylation/strand scission mechanism. Two oligonucleotides bind to separate sites on each of the two duplex strands, targeting guanines to the 5' end of the purine strand bound by each oligonucleotide.

Sequence design. The efficient modification (>85%) of one strand of a double helical DNA target sequence can be realized using the *N*-

bromoacetyloligonucleotide **5**.[61] Incorporation of triplex/alkylation sites on each of the strands of the double helix affords a second motif for double helical modification of DNA (Figure 4.16).

Using this strategy, double strand cleavage could be obtained at 5'-G-N-purine-$(N)_n$-pyrimidine-N-C-3' sequences (Fig. 4.16). Sequences of this type are recognized by the cross-over molecules pioneered by Horne and Dervan.[21] The cleavage sites on the two strands, however, would be 2x + 2 base pairs apart (x = triple helix binding site size). Difficulties in denaturation at the cleavage sights might be encountered.

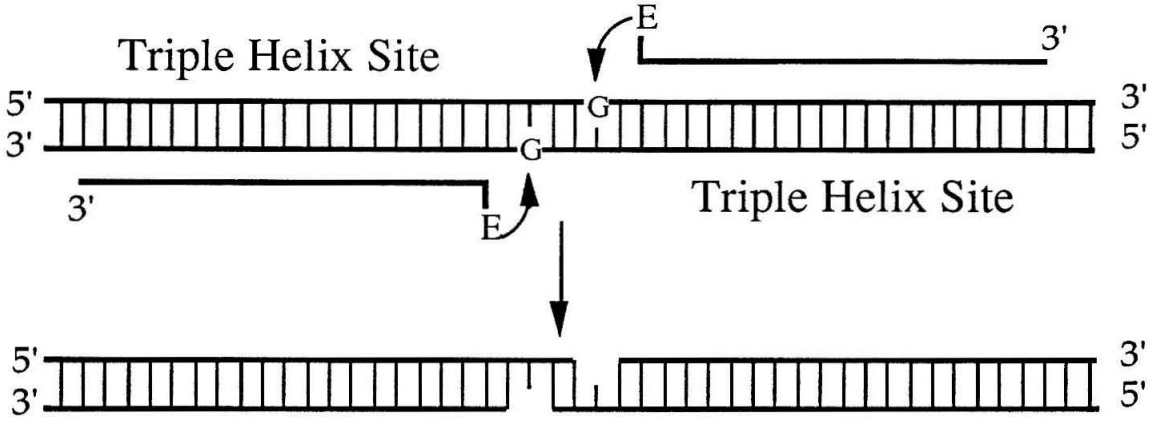


Fig. 4. 17. Strategy for the cleavage of double helical DNA via an alkylation/strand scission mechanism. Two oligonucleotides bind to separate sites on each of the two duplex strands, targeting guanines on the purine strand two base pairs to the 5' end of the triplex site. Recognition of 5'-pyrimidine-N-C-N-N-G-N-purine-3' sites places the sites of cleavage within a few base pairs of each other. Since cleavage is limited to a single nucleotide position, a properly designed overhang might be clonable to a complementary overhang produced by enzymatic cleavage of DNA.

A sequence of the type 5'-pyrimidine-N-G-$(N)_n$-C-N-purine-3' places the cleavage sites on each of the two strands a few bases apart (Fig. 4.17). The designed sequence must contain a guanine two base pairs to the 5' side of each triplex binding site, the position most efficiently alkylated by the bound *N*-bromoacetyloligonucleotide **5**. Incorporation of a single guanine 5' to the binding site insures that specific cleavage at a single nucleotide position on each strand occurs.

Model building indicates that juxtaposition of the two binding sites results in the formation of a four-stranded structure at the junction of the binding sites (Fig. 4.19). To avoid electrostatic repulsions which may make such a structure difficult to form, the binding sites are separated by a half turn of the DNA helix (six base pairs).

Because cleavage is limited to a single nucleotide position, the resulting ends might be compatible with and clonable to those produced by cleavage of DNA with a restriction endonuclease. Due to the nature of the triplex binding sites, the endonuclease must produce 3' overhangs. The base pairs between the two oligonucleotide binding sites were chosen such that ends produced after alkylation/cleavage by *N*-bromoacetyloligonucleotide **5** are complementary to DNA cut with the restriction endonuclease Sfi I.

The plasmid pUCLEU2C was constructed based on the above considerations by ligation of oligonucleotides containing the sequence 5'-<u>TTTTCTTTCCTTTTCTTTT</u>ACTAGT<u>AAAAGAAAAGGAAAGAAAA</u>-3' into the pUCLEU2 plasmid (Fig. 4.18).[26,31] This sequence contains binding sites (underlined) on each strand for the *N*-bromoacetyloligonucleotide **5** (5'-$T_4(Me^5C)T_4(Me^5C)_2T_3(Me^5C)T_4$-3'), previously shown to efficiently modify plasmid pUCALK. Incorporation of a single guanine 5' to the target site insures that cleavage occurs at a single base position. After modification of

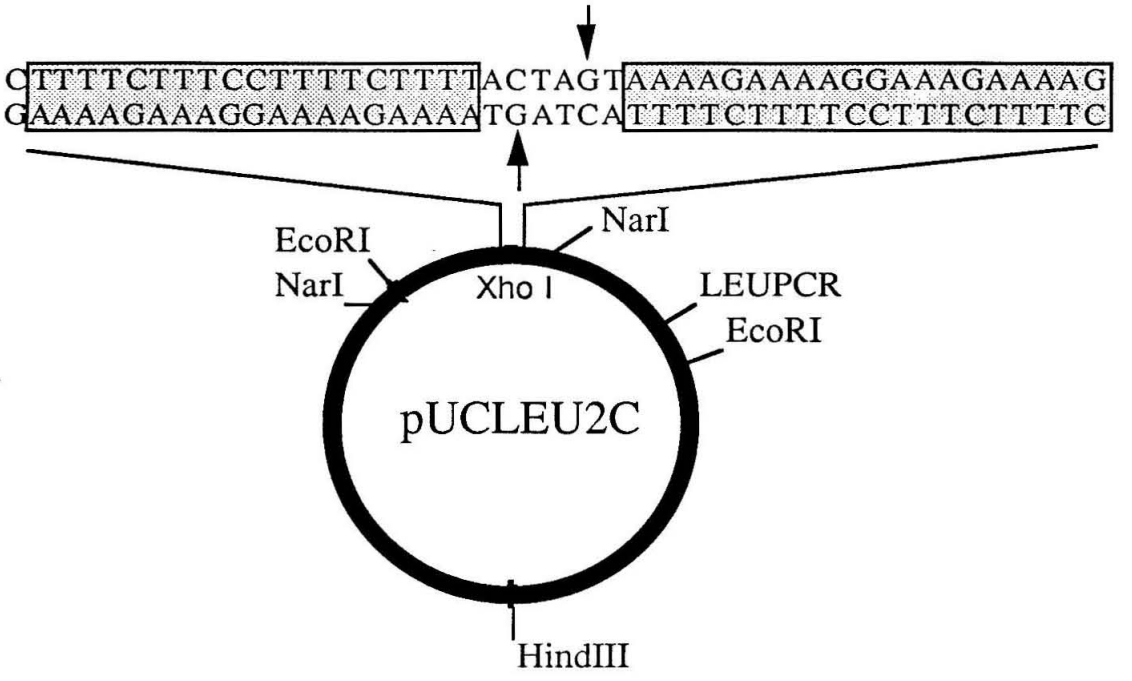


Fig. 4.18. Construction of the plasmid pUCLEU2C from pUCLEU2 via ligation of the oligonucleotides indicated into the Xho I site. Binding sites for the *N*-bromoacetyloligonucleotide **5** are boxed, and the targeted guanine bases are indicated by the arrows. Restriction endonuclease sites used for the isolation of desired restriction fragments are indicated.

the targeted guanine by the *N*-bromoacetyloligonucleotide, base treatment, loss of the guanine base, and strand scission, the resulting overhang (5'-CTA-3') is compatible with the ends generated by cleavage with Sfi I, which produces 5'-CTA-3' and 5'-TAG-3' overhangs. The presence of 6 base pairs between the oligonucleotide binding sites insures that electrostatic interactions do not prevent simultaneous binding of oligonucleotides to both DNA binding sites. For ease of analysis and oligonucleotide synthesis, a sequence containing identical binding sites on each strand of the duplex was designed for this study; however, using two different *N*-bromoacetyloligonucleotides, cleavage at different binding sites could be

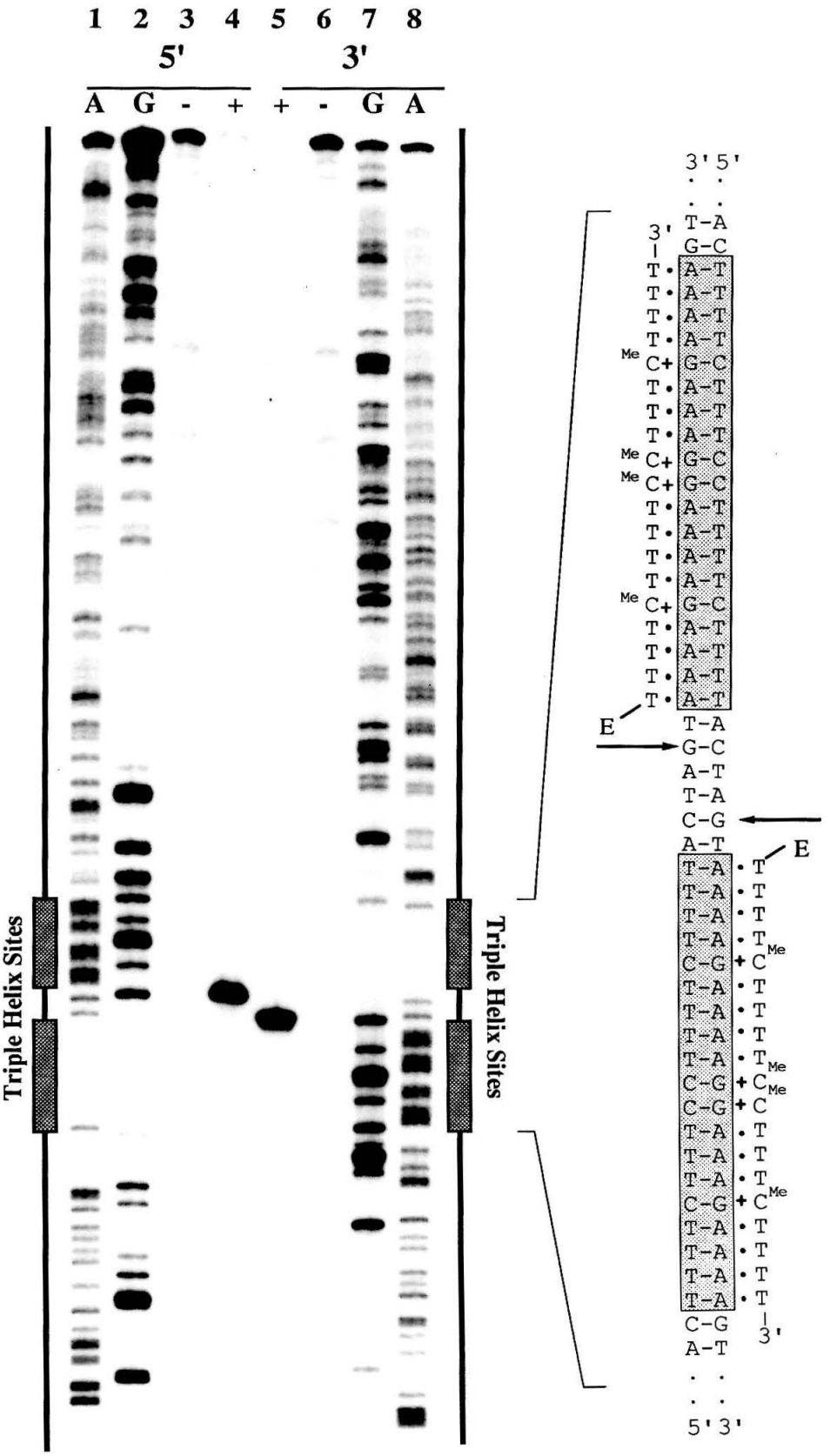
accomplished. The use of the plasmid pUCLEU2 as a vector, which contains a gene needed by the LEU2 deficient yeast strain SEY6210,[80] allows the introduction of the triplex target site into chromosome III of the yeast genome. Analysis of plasmid DNA cleavage products using high resolution polyacrylamide gel electrophoresis and of chromosomal DNA using pulsed field gel electrophoresis allows the study of the cleavage specificity afforded by *N*-bromoacetyloligonucleotides from the single base to the megabase pair level.

Nucleotide resolution of cleavage specificity. To determine the efficiency and specificity of modification of double helical DNA afforded by *N*-bromoacetyloligonucleotide **5** at the targeted guanine bases, a 0.9 kilobase pair Nar I/EcoR I restriction fragment containing the triple helix site was labeled approximately 210 base pairs from the target. DNA ^{32}P-labeled at the Nar I site on either strand (3' or 5' end-labeled) was incubated with *N*-bromoacetyloligonucleotide **5** under conditions previously shown to result in high yields of modification (0.8 mM $Co(NH_3)^{+3}$, 10 mM Hepes pH 7.4). Modification occurs solely at the guanine nucleotide located two bases to the 5' side of each oligonucleotide binding site (Fig. 4.19). The yield of alkylation is >96% on each strand. The higher yield (as compared to that previously reported on plasmid pUCALK)[61] is most probably a result of improvements in the purification (i. e. implementation of anion exchange HPLC as a means of purification) of the derivatized *N*-bromoacetyloligonucleotide **5**.

The rate constants for reaction of *N*-bromoacetyloligonucleotide **5** with each of the two strands were determined using phosphorimaging as previously described. Half lives for reaction of the *N*-bromoacetyloligonucleotide **5** with the 3' and 5' end-labeled strands are 5.5 and 6.5 hours respectively. These values are averages from three separate

Figure 4.19. Left: Autoradiogram of a high-resolution 6% denaturing polyacrylamide gel of cleavage products from reaction of *N*-bromoacetyloligonucleotide **5** with a ^{32}P 3'- and 5'-end-labeled 0.9 kbp restriction fragment (*Nar I/ Ssp I*) from plasmid pUCLEU2C. Reaction conditions were 1 μM concentration of oligonucleotide **5**, 20 mM Hepes, pH 7.4, 0.8 mM $Co(NH_3)^{+3}$, and 10,000 cpm end-labeled DNA in a total volume of 15 μl. Reactions were incubated at 37 °C for 36 hours, precipitated with NaOAc/EtOH, washed with 70% EtOH, and treated with 0.1 M piperidine (90 °C, 30 min.). After lyophilization, cleavage products were analyzed on a 6% 1:20 cross-linked, 48% urea polyacrylamide gel, 0.4 mm thick. Lanes 1-4 are 5'-end-labeled DNA. Lanes 5-8 contain 3' end-labeled DNA. Lanes 1 and 8 are A-specific chemical sequencing reactions for the 5' and 3' end-labeled fragments. Lanes 2 and 7 are G-specific chemical sequencing reactions for the 5' and 3' end-labeled fragments. Lanes 3 and 6 contain DNA incubated for 36 hours in the absence of 5'-*N*-bromoacetyloligonucleotide **5**, followed by piperidine treatment. Lanes 4 and 5 contain DNA incubated for 36 hours with 5'-*N*-bromoacetyloligonucleotide **5**, followed by piperidine treatment.

Right: Sequence of the oligonucleotide-DNA triplex within plasmid pUCLEU2C. The major sites of modification are indicated.


1 2 3 4 5 6 7 8
5'
3'
A G - + + - G A
Triple Helix Sites
Triple Helix Sites
3' 5'
T-A
G-C
E
T-A
G-C
A-T
T-A
C-G
A-T
E
C-G
A-T
5' 3'

determinations. It is not clear why the two rates differ, since the 19 base pair binding sites, as well as the four base pairs flanking *each side of each* binding site, are identical on the two strands. It is possible that small differences in salt concentration or purity of the DNA isolated from the different labelling reactions, or differences in the extended DNA conformations of the two strands, lead to the observed differences in alkylation rate. The rates on each strand are not considerably different from the rate reported for reaction of the *N*-bromoacetyloligonucleotide **5** with the plasmid pUCALK (half life = 6.2 hours).[61] The bases 5' to the purine site in the plasmid pUCALK are 5'-GGG-3'; the corresponding bases in the plasmid pUCLEU2C are 5'-TGA-3'. This suggests that the base pairs flanking the targeted guanine have little effect on the efficiency or rate of the reaction. Model building studies indicate that no combinations of flanking base pairs present significant steric barriers to reaction, although it is difficult to predict the effect of such changes on DNA conformation.

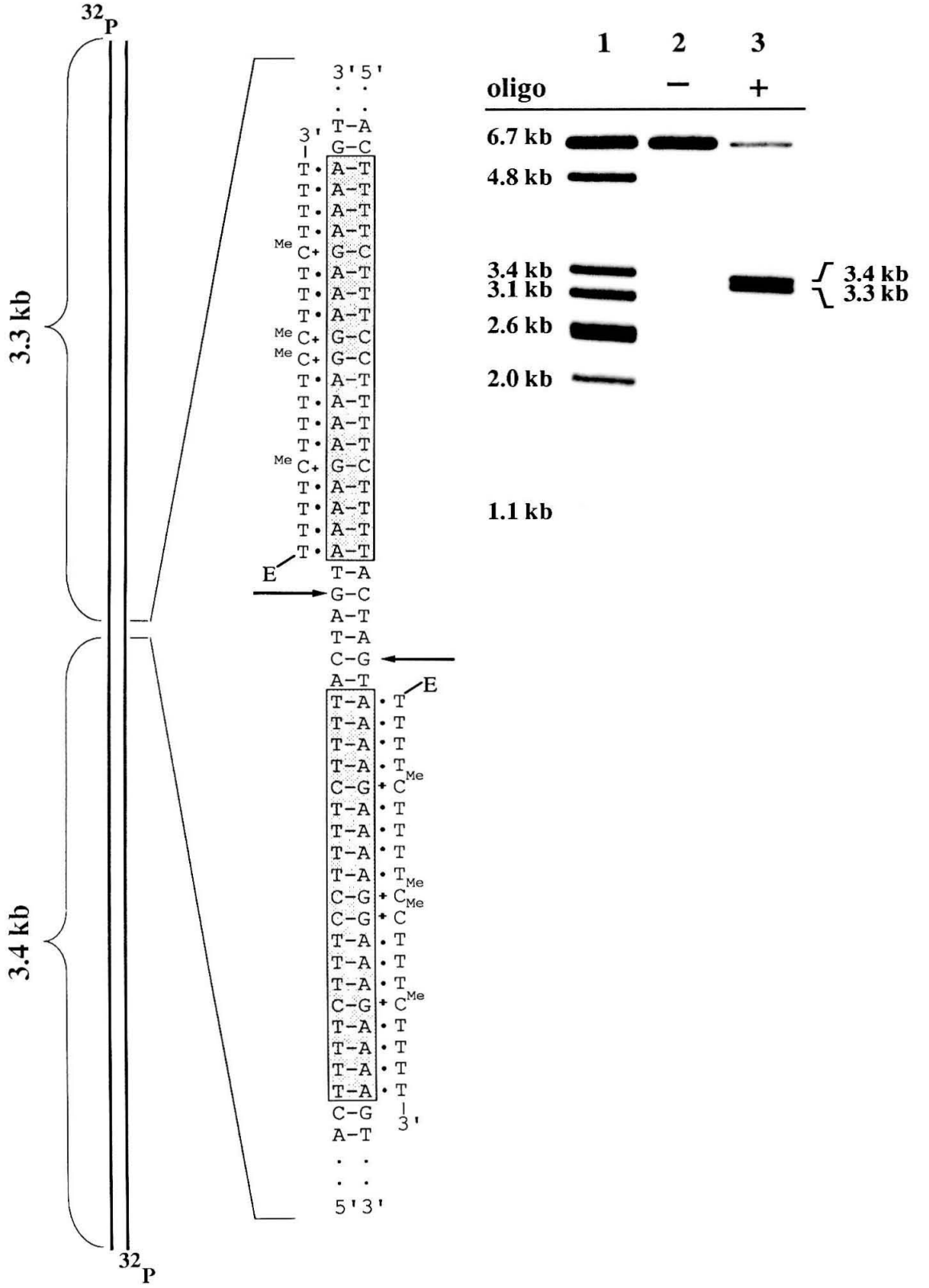
Double Strand DNA Cleavage. To determine if *N*-bromoacetyloligonucleotides can be used to effect efficient double strand cleavage of DNA, the plasmid pUCLEU2C was Hind III linearized and ^{32}P 3' end-labeled. Modification and cleavage at the triple helix target site is expected to result in the production of bands 3.3 and 3.4 kilobase pairs in size.

Standard treatment with piperidine at elevated temperatures (1% aqueous piperidine, 90 °C, 30 min.) results in denaturation of the duplex. Attempts to anneal the two strands together in a series of buffers after piperidine treatment and lyophilization were unsuccessful.

Incubation of ^{32}P end-labeled Hind III linearized pUCLEU2C DNA with *N*-bromoacetyloligonucleotide **5** produces DNA alkylated at the targeted

Figure 4.20. Left: Diagram of linearized plasmid pUCLEU2C indicating location and sequence of triple helical oligonucleotide binding sites. The sites of alkylation and subsequent DNA cleavage are indicated.

Right: Autoradiogram of a 1.0 % agarose gel of cleavage products from the reaction of oligonucleotide **5** with a ^{32}P 3'-doubly end-labeled linearized plasmid pUCLEU2C (6.7 kbp in length). Reaction conditions were 1 μM concentration of oligonucleotide **5**, 20 mM Hepes, pH 7.4, 0.8 mM $Co(NH_3)^{+3}$, and 10,000 cpm end-labeled DNA in a total volume of 15 μl. Reactions were incubated at 37 °C for 36 hours, precipitated with NaOAc/EtOH, washed with 70% EtOH, and treated with 0.1% aqueous piperidine in 100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8.0 (55 °C, > 12 hrs.). Cleavage products were analyzed on a 1% agarose gel. Lane 1: DNA size markers obtained by digestion with restriction endonucleases. Sizes are indicated to left of band. Lane 2: DNA incubated in the absence of oligonucleotide **5**, precipitated, and piperidine treated. Lane 3: DNA incubated with oligonucleotide **5**, precipitated, and piperidine treated.


32P
3.3 kb
3.4 kb
32P
3'5'
3'
E
5'3'
1
2
3
oligo
−
+
6.7 kb
4.8 kb
3.4 kb
3.1 kb
2.6 kb
2.0 kb
1.1 kb
3.4 kb
3.3 kb

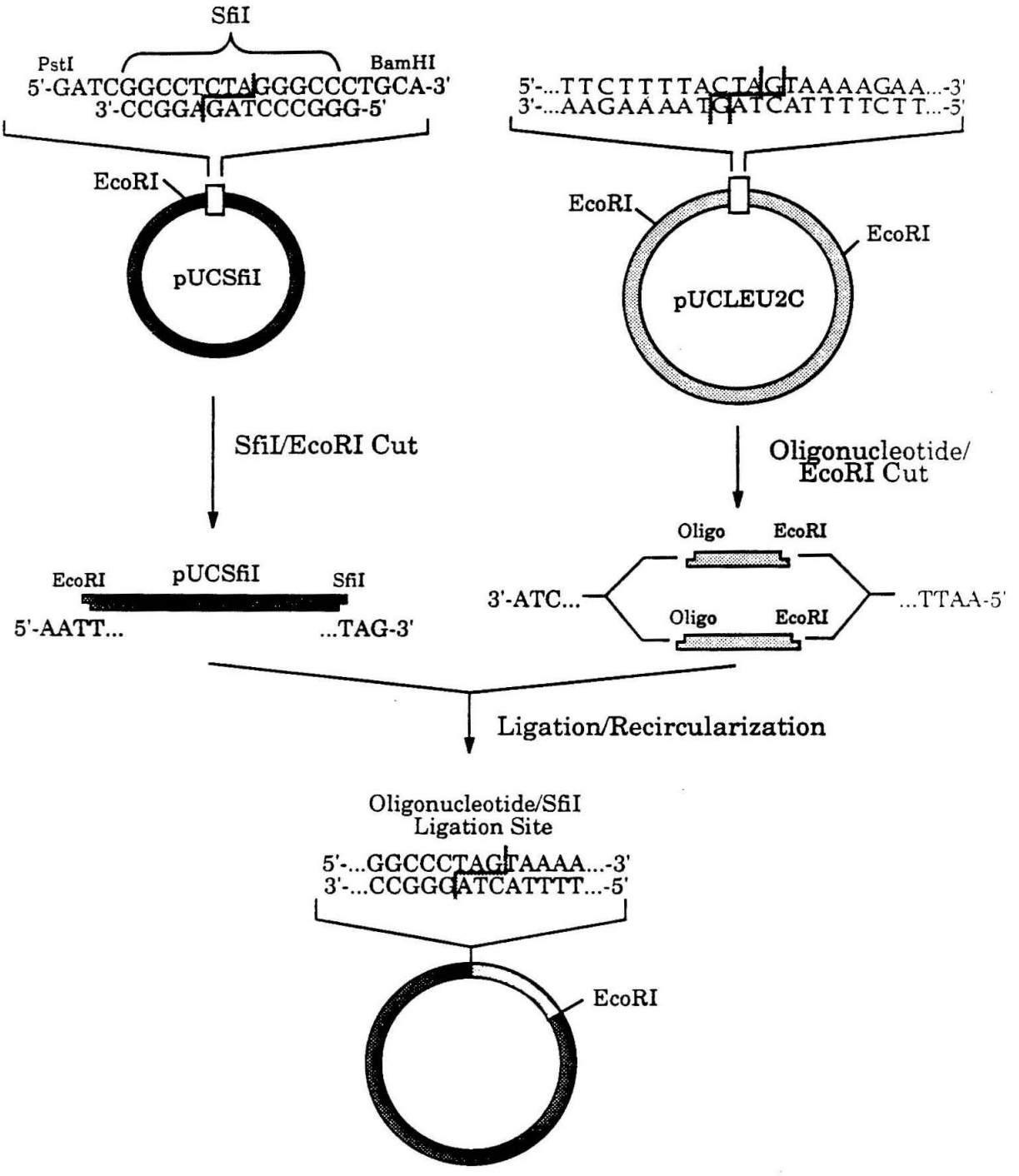
guanine bases. After precipitation from the alkylation reaction, incubation of the DNA in 0.1% piperidine at 55 °C for 12 hrs. in a buffer containing 10 mM EDTA, 10 mM Tris pH 8.0 and 100 mM NaCl results in the production of a single double strand break in high yield with no detectable cleavage at non-modified sights or denaturation of the duplex. The high yields observed indicate that these work-up conditions are sufficient to result in quantitative or near quantitative strand scission at modified sites.

Phosphorimaging of the agarose gel separated products indicates that double strand cleavage efficiencies of greater than 85% were achieved after 36 hours, suggesting single strand cleavage efficiencies of >92% on each strand.

Double strand alkylation and cleavage by *N*-bromoacetyloligonucleotide **5** was studied at both 37 °C and 50 °C. Although the reaction rate is substantially increased at 50 °C, a large increase in background degradation results. Apparently, at these elevated temperatures, the DNA is susceptible to modification by adventitious metals or other trace contaminants. Any depurination or other modification is converted to a cleavage site upon piperidine treatment. The optimal temperature for DNA alkylation/cleavage, representing a compromise between the rates of the specific vs. non-specific reactions, is approximately physiological temperature, 37 °C.

Cloning of N-bromoacetyloligonucleotide cleavage products. Cleavage of pUCLEU2C DNA with the *N*-bromoacetyloligonucleotide produces ends with the 3' overhanging sequence 5'-CTA-3'. The alkylation/piperidine treatment mechanism of DNA cleavage is expected to result in the production of 3' and recessed 5' phosphate ends. Ligation of standard restriction enzyme digested DNA products involves reaction of a 5' phosphate end with a 3' hydroxyl group. Digestion of the plasmid pUCSfiI

Figure 4.21. Cleavage of plasmid pUCSfiI with Sfi I and EcoR I produces as the longer piece of DNA a fragment containing a 5' overhang on one end (5'-AATT) and a 3' overhang on the other (3'-TAG). Incubation of plasmid pUCLEU2C with *N*-bromoacetyloligonucleotide **3**, followed by strand scission effected by piperidine treatment results in double strand cleavage. Cleavage with the restriction endonuclease EcoR I produces two fragments containing 5'-AATT and 3'-ATC overhangs. Both of these products are ligatable with the larger of the two fragments isolated after Sfi I/ EcoR I digestion of pUCSfiI. After circularization and transformation of bacterial cells with the new plasmid, DNA was isolated and sequenced to verify the production of ligatable ends using the *N*-bromoacetyloligonucleotide/piperidine cleavage motif.


SfiI
PstI
BamHI
5'-GATCGGCCTCTAGGGCCCTGCA-3'
3'-CCGGAGATCCCGGG-5'
EcoRI
pUCSfiI
5'-...TTCTTTTACTAGTAAAAGAA...-3'
3'-...AAGAAAATGATCATTTTCTT...-5'
EcoRI
EcoRI
pUCLEU2C
SfiI/EcoRI Cut
Oligonucleotide/
EcoRI Cut
EcoRI
pUCSfiI
SfiI
5'-AATT...
...TAG-3'
Oligo
EcoRI
3'-ATC...
...TTAA-5'
Oligo
EcoRI
Ligation/Recircularization
Oligonucleotide/SfiI
Ligation Site
5'-...GGCCCTAGTAAAA...-3'
3'-...CCGGGATCATTTT...-5'
EcoRI

with SfiI produces the 3' overhanging sequence, 5'-TAG-3', which is complementary to the end of the oligonucleotide digested product (Fig. 4.21).

Because the 3' end of the SfiI product contains a free hydroxyl group, it could be ligated to the expected 5' phosphate end of the *N*-bromoacetyloligonucleotide digested DNA. It is unlikely that the 3' phosphate end of the oligonucleotide digested product is ligatable to the 5' phosphate of the restriction enzyme cut DNA; however, the remaining single strand nick could readily be repaired *in vivo* following transformation.

The two cleavage products obtained by digestion of pUCLEU2C with EcoR I and *N*-bromoacetyloligonucleotide **5** were each ligated into the Sfi I/EcoR I fragment of pUCSfiI using T4 DNA ligase by Scott Strobel. The colonies were screened by α-complementation and five plasmids for each of the two fragments were sequenced. All products were consistent with cleavage by *N*-bromoacetyloligonucleotide **5** at the targeted positions, loss of the guanine base, and simple ligation of the resulting fragments into the SfiI/EcoR I sites of pUCSfiI. Thus, this synthetic cleaving agent results in the chemical modification of DNA in a manner which produces end products that can be readily ligated to restriction enzyme digested DNA.

Alkylation/Cleavage of agarose embedded DNA by N-bromoacetyloligonucleotides. To determine the effect of base specific modification on the specificity of oligonucleotide-directed triple helical cleavage, it is desirable to study the modification and subsequent cleavage at a triple helix site within a large piece of DNA (the 340 kilobase pair yeast chromosome III). To avoid degradation due to shearing, large DNA is routinely manipulated by embedding it within agarose plugs, followed by

diffusion of reagents in and out of the agarose matrix. Oligonucleotides equipped with EDTA·Fe(II) are capable of binding to and cleaving agarose embedded DNA.[26,32] To determine whether alkylation by *N*-bromoacetyloligonucleotides is compatible with such manipulations, Hind III linearized plasmid pUCLEU2C was embedded in 1% low-melting agarose. Incubation of the plug in 0.8 mM $Co(NH_3)_6^{+3}$, 10 mM Hepes pH 7.0, diffusion of the *N*-bromoacetyloligonucleotide **5** into the plug, incubation at 37 °C and piperidine treatment (0.1% piperidine in 100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8.0, 55 °C, 12 hours) results in double strand cleavage of the DNA at the triple helix site with slightly lower yields than those observed in solution. This slight loss in yield could be due to reaction between the bromoacetyl moiety and the agarose saccharides.

To circumvent this problem, DNA was embedded in agarose which had been pre-treated by incubation in 2-bromoacetamide. It was hoped that such incubation would "cap" nucleophilic sites capable of reaction with the *N*-bromoacetyl moiety. Agarose solidified in molds to afford a high surface:volume ratio was incubated in a buffer solution containing 10 mM 2-bromoacetamide at 37 °C for 36 hours, then extensively washed in buffer. The agarose was melted and used to cast plugs containing either plasmid DNA or yeast chromosomal DNA. While agarose treated in this manner retains its gelling characteristics and DNA can easily be embedded within it, its melting temperature and resistance to shearing are reduced. The resulting plugs are susceptible to mechanical destruction during the several pipetting steps required to equilibrate the plug in buffer. Incubation of the plug at 55 °C in 0.1% piperidine (100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8.0) results in partial melting and destruction of the plug. These observations indicate that

reaction between bromoacetamide and agarose results in a loss of gel strength. Such reaction may be responsible for the decreased efficiency of DNA modification by *N*-bromoacetyloligonucleotides observed on DNA embedded in agarose plugs.

Site specific cleavage of yeast chromosomal DNA using N-bromoacetyloligonucleotides. Oligonucleotides equipped with the DNA cleaving moiety EDTA·Fe(II) are capable of binding to and cleaving DNA at a single engineered triple helix site within a yeast chromosome.[26] These molecules, however, are both inefficient (maximum cleavage yield observed was 6%) and, unless conditions are carefully chosen, result in double strand cleavage at partially homologous secondary sites.[26] Coupling of a base specific cleaving moiety with the >15 base pair specificity attainable using the triple helix motif may result in the construction of molecules which, although bound at secondary sites, effect DNA modification and cleavage at a single site. In this study, additional specificity is obtained due to the requirement that potential triple helix sites (i. e. polypurine runs) be present on both strands of the DNA within a reasonable number (<20) of base pairs for double strand modification and cleavage to be observed.

A yeast strain containing the target site found in pUCLEU2C was constructed using homologous recombination between the pUCLEU2C plasmid and *Saccharomyces cerevisiae* yeast strain SEY6210. The strain is *leu2-*, inhibiting growth on leucine deficient media.[80] The plasmid pUCLEU2C contains the desired target site adjacent to the *LEU2* gene. Incorporation of this fragment with the *LEU2* gene into the yeast genome allows selection on leucine deficient media. Transformed yeast strains were selected, and the presence of the target site was verified by PCR amplification of segments of the third chromosome, followed by restriction enzyme

digestion and sequencing.[26,31] Pulsed field gel electrophoresis of the yeast genome to separate the chromosomes, followed by southern blotting with probes specific for the LEU2 gene, verified that the LEU2 gene and the triplex sites were incorporated exclusively into chromosome III of the *Saccharomyces cerevisiae* genome.

DNA from a transformed strain was isolated and embedded in agarose to protect against mechanical shearing during DNA manipulations. After equilibration with $Co(NH_3)_6^{+3}$ (1 mM) and buffer (Hepes, pH 7.4, 10 mM EDTA) (triplex/alkylation buffer), the plug was incubated with *N*-bromoacetyloligonucleotide **5** at 37 °C for 24 hours. The unreacted *N*-bromoacetyloligonucleotide was successively diluted from the duplex in a series of washes using low salt, high pH (10 mM Tris, pH 9.5, 10 mM EDTA) buffer, the plugs were reequilibrated in triplex/alkylation buffer, and a second aliquot of *N*-bromoacetyloligonucleotide **5** was added. After this second incubation, plugs were washed and incubated with piperidine (0.1% aqueous piperidine, 100 mM NaCl, 10 mM EDTA, 10 mM Tris, pH 8.0, 55 °C, 12 hours) to effect strand scission. Pulsed field gel electrophoresis results in complete resolution of the smallest four chromosomes, with partial separation of the larger chromosomes (Fig. 4.22).

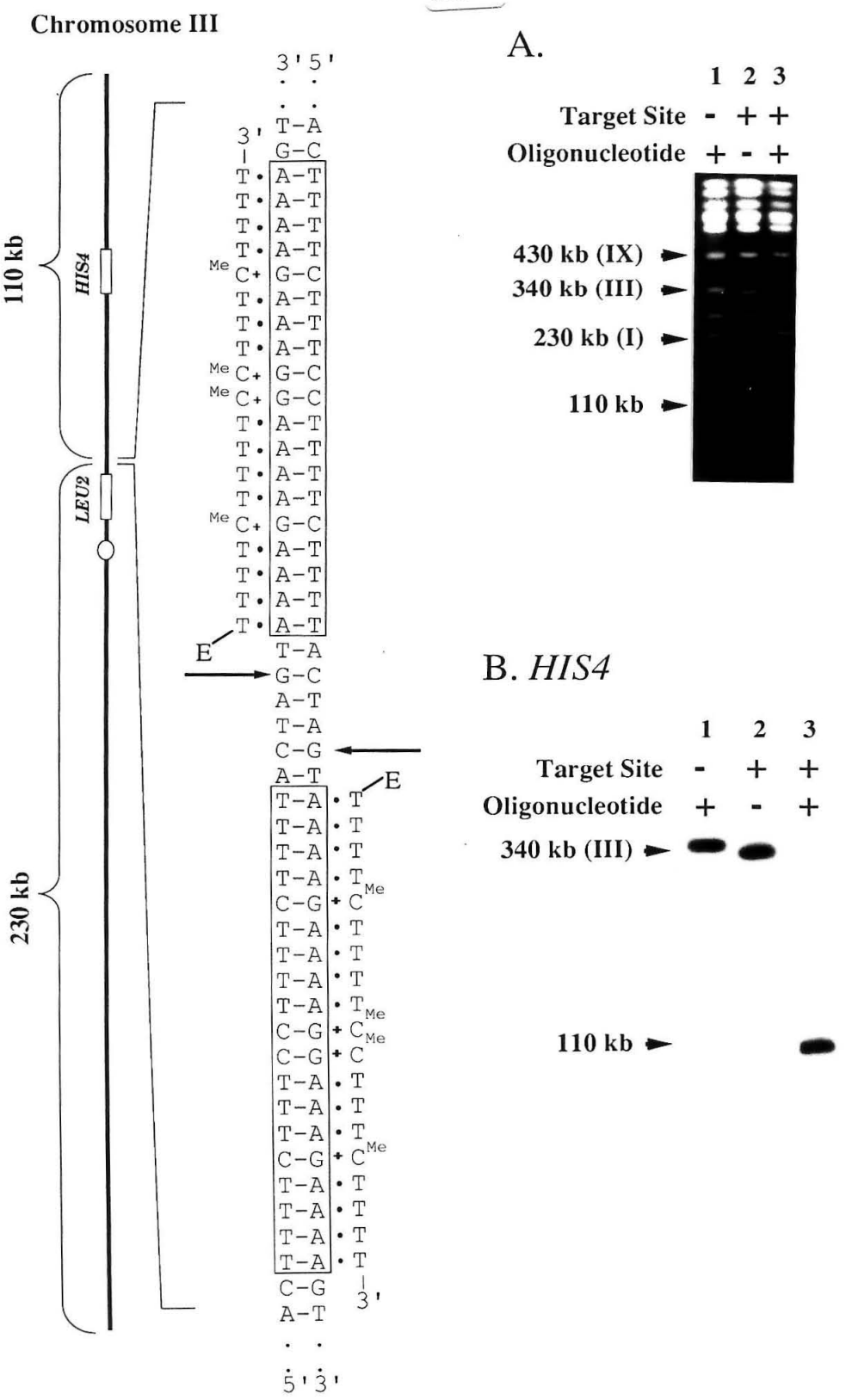
Incubation of yeast chromosomes with *N*-bromoacetyloligonucleotide **5** produces the expected 110 kbp product (Fig. 4.22). The larger of the expected products (230 kbp in size) co-migrates with chromosome I and is not visualized by ethidium staining. The smaller of the products can be faintly observed as a fast running band on the gel at the position indicated. Size determination is based on the migration of 39 kbp size ladders. Ethidium staining indicates that the chromosomes are intact and that no unexpected new bands are produced. These results indicate that: (i) the piperidine

Figure 4.22: Cleavage of yeast chromosome III by *N*-bromoacetyloligonucleotide **5**. Agarose plugs containing DNA from yeast strains SEY6210 (no triplex site) and SEY6210C (triplex site present) were equilibrated in 1.0 mM $Co(NH_3)^{+3}$, 20 mM Hepes pH 7.4, 10 mM EDTA, and 2 μM *N*-bromoacetyloligonucleotide **5**. After 24 hours, the plugs were washed with a Tris pH 9.5 (10 mM), EDTA (10 mM) solution three times (10 min. each wash). Plugs were re-equilibrated with 1.0 mM $Co(NH_3)^{+3}$, 20 mM Hepes pH 7.4, 10 mM EDTA, and 2 μM *N*-bromoacetyloligonucleotide **5**. After a second 24 hour incubation (37 °C), the plugs were washed in 0.1% piperidine, 100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8.0, (3 washes 10 min. each) and incubated in 900 μl of this solution for 12 hours at 55 °C. The plugs were equilibrated with 0.5x TBE buffer, and loaded into a 1.0% agarose gel for pulsed field gel electrophoresis. The gel was run for 24 hours at 200 V. Pulsed times were ramped from 10 - 60 seconds for the first 18 hours, and from 60 - 90 sec. over the last 6 hours.

Left. Diagram of yeast chromosome III indicating location and sequence of triple helical oligonucleotide binding sites and site of binding of *HIS 4* probe used for cleavage analysis by hybridization. The binding site sequence and sites of alkylation and subsequent DNA cleavage are indicated.

A. Ethidium Bromide stained pulsed field agarose gel. Chromosomes IX, III, II and I are clearly resolved and indicated. Resolution of the remaining chromosomes is incomplete. Lane 1: DNA from yeast strain SEY6210 lacking triple helical target site incubated with *N*-bromoacetyloligonucleotide **5** and piperidine treated. Lane 2: DNA from yeast strain SEY6210C incubated and piperidine treated in the absence of *N*-bromoacetyloligonucleotide **5**. Lane 3: DNA from yeast strain SEY6210C incubated with *N*-bromoacetyloligonucleotide **5** and piperidine treated.

B. Autoradiogram of a DNA blot hybridization experiment of gel in A with a 250 bp HIS 4 fragment radio-labeled with ^{32}P by random-probe hybridization.[77] Locations of the intact chromosome III (340 kbp) and the 110 kbp cleavage fragment are indicated.


Chromosome III
110 kb
230 kb
HIS4
LEU2
3' 5'
E
A.
1 2 3
Target Site - + +
Oligonucleotide + - +
430 kb (IX)
340 kb (III)
230 kb (I)
110 kb
B. HIS4
1 2 3
Target Site - + +
Oligonucleotide + - +
340 kb (III)
110 kb

incubation is gentle enough to be used on chromosomal DNA without general destruction of the DNA and (ii) non-specific modification of DNA by the *N*-bromoacetyloligonucleotide **5** occurs infrequently, resulting in no chromosomal degradation. It is interesting to note that non-specific binding by oligonucleotides is greatly disfavored, due to charge repulsion between the oligonucleotide and the target DNA. The ratio of specific to non-specific DNA modification by derivatized oligonucleotides is most likely very high. Most DNA binding molecules and proteins are positively charged, increasing their non-specific as well as specific binding affinities for DNA. Cleavage of large DNA using molecules of net positive charge might result in a large amount of random cleavage and degradation, unless the specificity of the cleaving moiety is very high.[35]

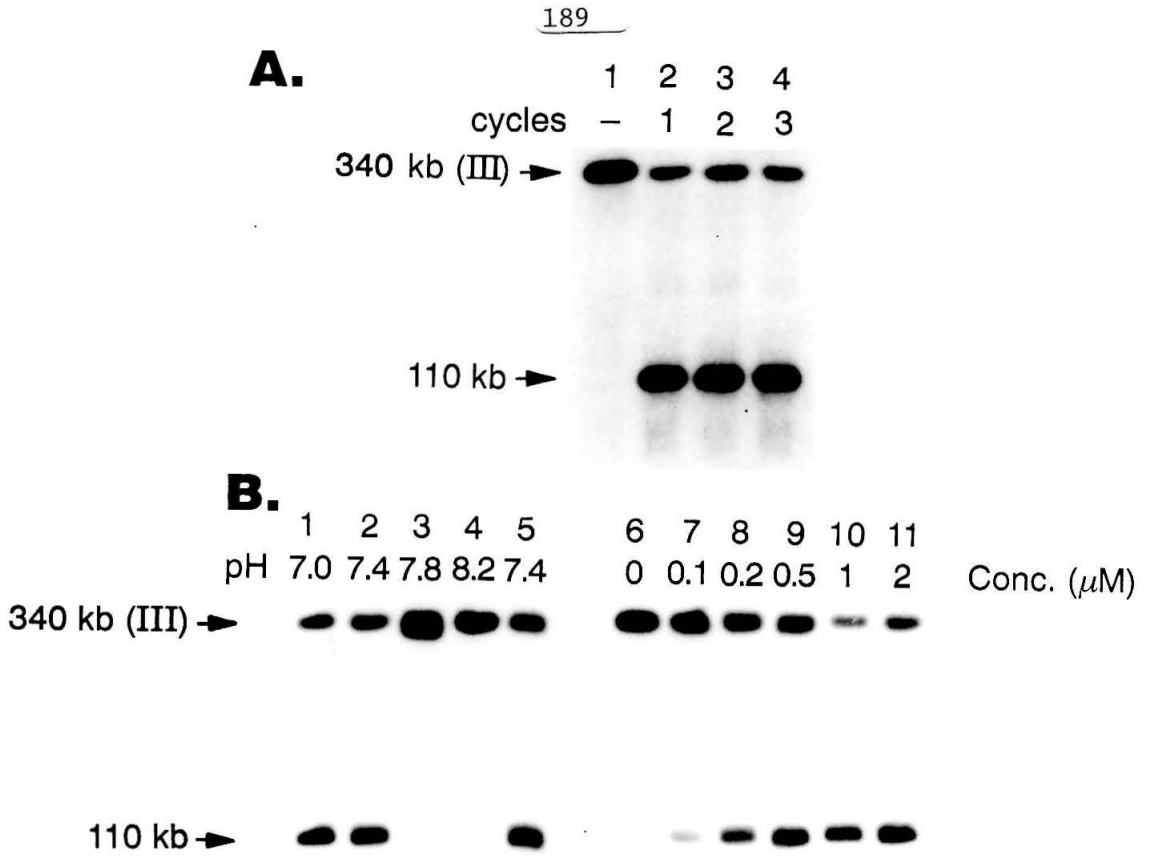
To accurately determine the cleavage efficiency of chromosomal cleavage, as well as to determine if low efficiency cleavage is occurring at secondary sites on chromosome III, the DNA was transferred to a membrane and hybridized with a marker specific for HIS 4, a gene found on the short arm of yeast chromosome III (Fig. 4.22). The cleavage efficiency, as determined by blot hybridization with a HIS 4 probe and analysis of the data by phosphorimaging, is 90%. Furthermore, no cleavage at secondary sites is observed ($\leq \sim 2\%$), indicating that double strand modification occurs at a single site within the 340 kbp of chromosome III.

Effect of oligonucleotide cycling, pH, and concentration. The effect of cycling of the oligonucleotide was studied by incubation of yeast DNA with one, two or three separate aliquots of oligonucleotide **5** (Fig. 4.23, A). In each of the reactions, yeast DNA was incubated for a total of ~ 72 hours with *N*-bromoacetyloligonucleotide **5**. DNA in lane 1 was incubated with a single aliquot of *N*-bromoacetyloligonucleotide **5**, DNA in lane 2 was incubated with

Figure 4.23: Cleavage of yeast chromosome III assayed by pulsed field agarose gel electrophoresis and ^{32}P-HIS 4 probe hybridization. Unless indicated otherwise, all reactions were conducted as follows: agarose plugs containing DNA from yeast strains SEY6210 and SEY6210C were equilibrated in 1.0 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.4, 10 mM EDTA (alkylation/triplex buffer), and *N*-bromoacetyloligonucleotide **5** was added to a final concentration of 2 μM. After 24 hours, the plugs were washed with Tris pH 9.5 (10 mM) and EDTA (10 mM) three times (10 min. each wash). The plugs were re-equilibrated with 1.0 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.4, 10 mM EDTA, and *N*-bromoacetyloligonucleotide **5** was added to a final concentration of 2 μM. After a second 24 hour incubation, the plugs were washed in 0.1% piperidine, 100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8.0, (3 washes, 10 min. each) and incubated for 12 hours at 55 °C in 900 μl of the piperidine solution. The plugs were chilled, washed twice with 0.5x TBE buffer, and loaded into a 1.0% agarose gel for pulsed field gel electrophoresis. The gel was run for 24 hours at 200V. Pulsed times were ramped from 10 - 60 seconds for the first 18 hours, and from 60 - 90 sec. over the last 6 hours. The DNA was transferred to a membrane, and hybridized with a 250 bp HIS4 probe specific for the short arm of chromosome III and ^{32}P-radio-labeled by random priming.

A. Cleavage of yeast chromosomal DNA as a function of number of equilibrations with *N*-bromoacetyloligonucleotide **5**. Autoradiogram of a DNA blotting experiment with a random primer ^{32}P radio-labeled HIS 4 probe specific for the short arm of chromosome III after pulsed field agarose gel electrophoresis. Lane 1: Yeast chromosomal DNA incubated for 24 hours in the absence of *N*-bromoacetyloligonucleotide **5** in alkylation/triplex buffer, washed in high pH, low salt buffer, equilibrated for 20 hours in the absence of *N*-bromoacetyloligonucleotide **5** in alkylation/triplex buffer, washed a second time, equilibrated for 20 hours in the absence of *N*-bromoacetyloligonucleotide **5** in alkylation/triplex buffer, washed and treated with piperidine as stated above. Lane 2: Yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide **5** for 72 hours, followed by piperidine treatment. Lane 3: Yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide **5** for 44 hours, washed in high pH low salt buffer, incubated with *N*-bromoacetyloligonucleotide **5** for 24 hours, and piperidine treated. Lane 4: Yeast chromosomal DNA incubated for 24 hours with *N*-bromoacetyloligonucleotide **5**, washed in high pH low salt buffer, equilibrated for 20 hours with *N*-bromoacetyloligonucleotide **5**, washed again, equilibrated for 20 hours with *N*-bromoacetyloligonucleotide **5**, washed and treated with piperidine as stated above.

B. Cleavage of yeast chromosomal DNA as a function of pH, EDTA, and *N*-bromoacetyloligonucleotide concentration. Autoradiogram of a DNA blotting experiment with a random primer ^{32}P radio-labeled HIS 4 probe after pulsed field gel electrophoresis. Yeast chromosomal DNA incubated for 24 hours with *N*-bromoacetyloligonucleotide **5** (2 μM) in alkylation/triplex buffer, washed in high pH, low salt buffer, equilibrated for 20 hours with *N*-bromoacetyloligonucleotide **5**, and piperidine treated as stated above. Lane 1: Yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide **5** at pH 7.0. Lane 2: Yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide **5** at pH 7.4. Lane 3: Yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide **5** at pH 7.8. Lane 4: Yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide **5** at pH 8.2. Lane 5: Yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide **5** at pH 7.4, in the presence of 1 mM EDTA. Lane 6: Yeast chromosomal DNA incubated in the absence of *N*-bromoacetyloligonucleotide **5**. Lane 7: Yeast chromosomal DNA incubated with 0.1 μM *N*-bromoacetyloligonucleotide **5**. Lane 8: Yeast chromosomal DNA incubated with 0.2 μM *N*-bromoacetyloligonucleotide **5**. Lane 9: Yeast chromosomal DNA incubated with 0.5 μM *N*-bromoacetyloligonucleotide **5**. Lane 10: Yeast chromosomal DNA incubated with 1 μM *N*-bromoacetyloligonucleotide **5** . Lane 11: Yeast chromosomal DNA incubated with 2 μM *N*-bromoacetyloligonucleotide **5** .


A.
1 2 3 4
cycles – 1 2 3
340 kb (III)
110 kb
B.
1 2 3 4 5 6 7 8 9 10 11
pH 7.0 7.4 7.8 8.2 7.4 0 0.1 0.2 0.5 1 2
Conc. (μM)
340 kb (III)
110 kb

two separate aliquots of *N*-bromoacetyloligonucleotide **5**, and DNA in lane 3 was incubated with three separate aliquots of *N*-bromoacetyloligonucleotide **5**.

The cleavage efficiencies for lanes 1-3 are 83, 80 and 83%, indicating that disruption of the triplex followed by incubation with fresh oligonucleotide results in no detectable increase in cleavage efficiency, and that efficient cleavage can be realized using a single lengthy incubation.

The effect of pH, EDTA, and oligonucleotide concentration on the efficiency of chromosomal cleavage by *N*-bromoacetyloligonucleotides was studied (Fig. 4.23, B). Cleavage at pH 7.0 and 7.4 remained high, in agreement with studies on plasmid DNA. No cleavage was observed at pH ≥ 7.8 (B, lanes 3 and 4), in accordance with the known pH dependence of pyrimidine oligonucleotide-directed triple helix formation.[18,19,26,27] The presence of EDTA in yeast cleavage reactions prevents the degradation of the DNA by trace nucleases or metals. The concentration of EDTA has no effect on alkylation/cleavage efficiency (B, lane 5).

Oligonucleotide concentration markedly effects cleavage efficiency. Although efficient cleavage of plasmid DNA in solution is observed at very low (16 nM) *N*-bromoacetyloligonucleotide concentrations, the cleavage of large DNA in agarose requires micromolar *N*-bromoacetyloligonucleotide concentrations (B, lanes 6-11).

Statistically and empirically expected specificity of N-bromoacetyloligonucleotides. Statistically, a 19 base pair purine site flanked by a guanine base two base pairs to the 5' side of the binding site is expected to occur once every 2.1 million base pairs ($2^{19} \times 4$). Statistically, two such sites occur within 20 base pairs of each other once in 1×10^{11} base pairs ($[2 \times 10^6]^2 \times 0.05$). Previous data indicate that a pyrimidine oligonucleotide of similar

composition equipped with EDTA•Fe(II) binds to and cleaves the target DNA at a limited number (4) of secondary sites within the 340 kilobase pairs of chromosome III under conditions similar to those used in the present study.[26] Assuming that this represents a random distribution of partially homologous pyrimidine oligonucleotide binding sites within genomic DNA, this suggests that *N*-bromoacetyloligonucleotides may bind approximately once every 100,000 base pairs of DNA. DNA modification is expected to occur at a subset of these sites (statistically 1 out of every four), due to the requirement for an appropriately positioned guanine base, making the specificity of single strand modification one site per 4×10^5 base pairs. Two such binding sites within 20 base pairs of each other are statistically expected to occur once every 8×10^9 base pairs ($[4 \times 10^5]^2 / 20$).

The empirically observed specificity of pyrimidine oligonucleotides is considerably less than what is statistically expected. This may be due to the known overrepresentation of polypurine sequences in eukaryotic DNA.[33,34] The specificity of double strand cleavage by *N*-bromoacetyloligonucleotides arises in great measure from the requirement for two oligonucleotide binding sites on separate DNA strands. The specificity expected for this motif based on the observations of oligonucleotide-EDTA•Fe(II) specificity made by Strobel[26] is still sufficient to allow for single site cleavage within DNA roughly the size of 2 human genomes. No double strand secondary cleavage sites (> 3% cleavage efficiency) are detected within the 340 kbp of chromosome III, indicating that within the limits of this experiment, the coupling of cleavage and binding specificities results in reagents capable of site specific cleavage within megabase size DNA.

Conclusion

Two strategies for the double strand modification and cleavage of double helical DNA are presented. Cross-over *N*-bromoacetyloligonucleotides containing two 5' ends recognize and modify both strands of the double helix, but double helical cleavage yields are low. Incorporation of triple helix sites on both strands allows DNA modification by two *N*-bromoacetyloligonucleotides at either 5'-purine-N-G-N-C-N-pyrimidine-3' or 5'-pyrimidine-N-C-N-G-N-purine-3' sequences.

Using this motif, *N*-bromoacetyloligonucleotides have be used to effect the base specific, near quantitative modification of double helical DNA in a manner which produces ends which are clonable with compatible ends produced by conventional restriction enzyme digestion. The specificity of this motif is sufficient to produce efficient double strand chemical cleavage at a single nucleotide position within a 340 kbp yeast chromosome.

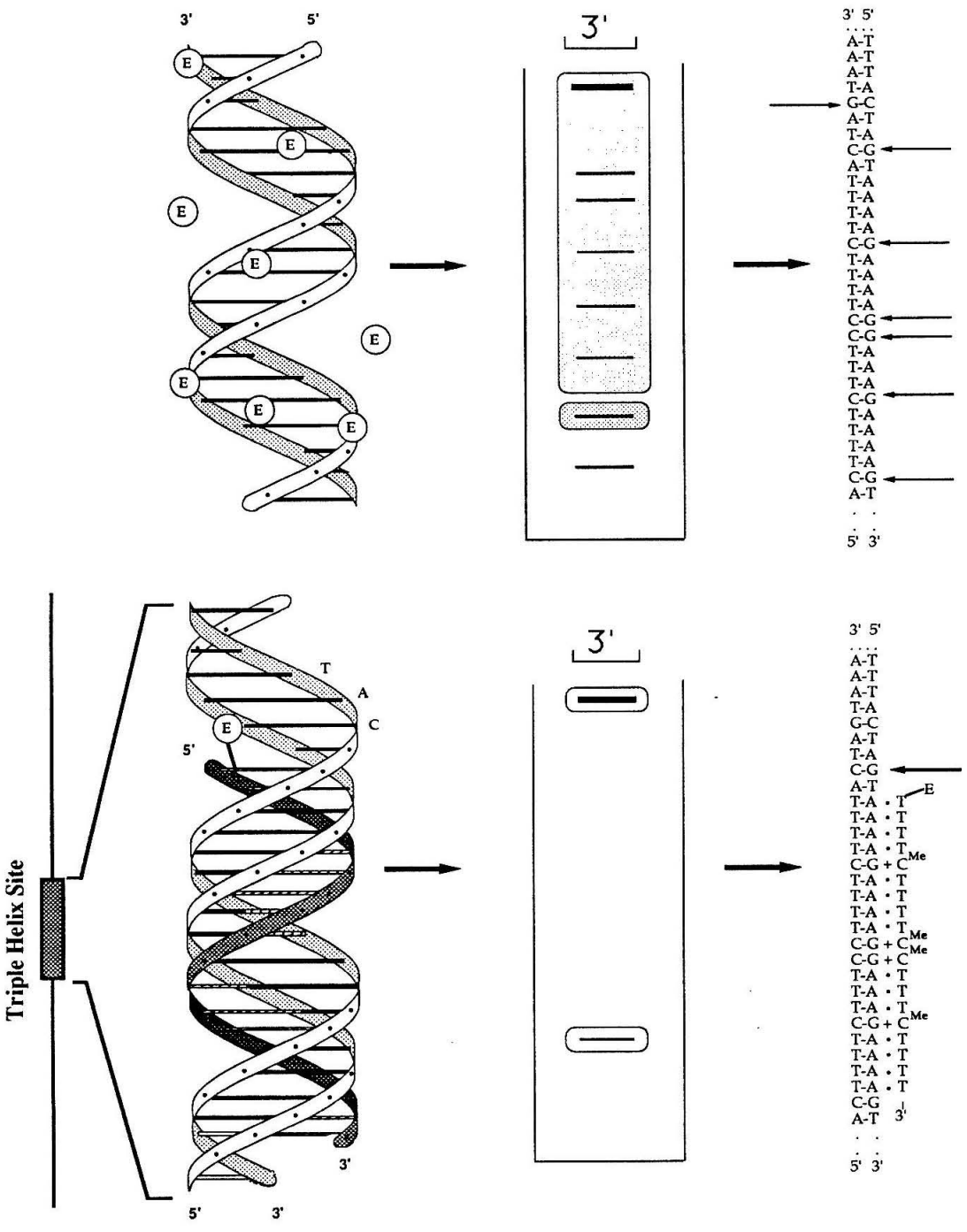
Part III

Determination of the Effective Molarity of a DNA Modification Moiety Tethered to a DNA Binding Molecule

Life is made possible by the presence of natural catalysts which are capable of effecting, under mild conditions (37 °C, neutral pH), reactions which would otherwise simply not occur on any appreciable time scale.[81] A large body of work indicates that significant rate enhancement can be obtained upon incorporation or binding of two reactants within a single molecule or macromolecular complex.[82-91] Since this is one mechanism which can be and is used by natural systems to accelerate reactions under mild conditions, it is not surprising that this phenomenon has been thoroughly studied. This rate enhancement is most often reported as an effective molarity, and values as high as 10^{13} M have appeared in the literature.[88] This value represents the concentration of one of the starting materials required to achieve a rate of bimolecular reaction equal to that observed in the intramolecular case.

Affinity cleaving, a technique developed to elucidate the structural principles for DNA recognition,[1,3-5,7-13,15,16,18,21,32,92,93] is indirectly based on this concept. This technique relies on the use of chemical synthesis to attach an indiscriminate and highly reactive cleaving reagent to a DNA binding molecule. In so doing, an inefficient bimolecular reaction is transformed into a reaction occurring within a complex containing the double helical target DNA, the binding molecule, and the attached reactive moiety (Figure 4.24).[1,2] The covalent tether is designed both to localize the reactive moiety and to

Figure 4.24. (top) Reaction of bromoacetamide (E = electrophile) with DNA is a bimolecular reaction, which results in cleavage at each guanine in the sequence. (bottom) Covalent attachment of the bromoacetamide moiety to an oligonucleotide results in reaction solely within the triple helical complex, and exclusive modification of a single guanine. Shaded areas in the gel indicate bands which were quantitated for the determination of the bimolecular and unimolecular rate constants. The sequence indicated is that proximal to the triple helical complex.


3'
5'
E
3'
3' 5'
A-T
A-T
A-T
T-A
G-C
A-T
T-A
C-G
A-T
T-A
T-A
T-A
T-A
C-G
T-A
T-A
T-A
T-A
C-G
C-G
T-A
T-A
T-A
C-G
T-A
T-A
T-A
T-A
C-G
A-T
5' 3'
Triple Helix Site
T
A
C
5'
3'
5'
3'
3'
3' 5'
A-T
A-T
A-T
T-A
G-C
A-T
T-A
C-G
A-T
E
T-A • T
T-A • T
T-A • T
T-A • T
Me
C-G + C
T-A • T
T-A • T
T-A • T
T-A • T
Me
C-G + C
Me
C-G + C
T-A • T
T-A • T
T-A • T
Me
C-G + C
T-A • T
T-A • T
T-A • T
T-A • T
C-G
A-T
3'
5' 3'

incorporate the flexibility required to avoid modifying the binding properties of the molecule recognizing the DNA.[1,2] It is of interest to understand the rate acceleration obtained upon formation of the DNA-affinity cleaving molecule complex as compared to the rate of the bimolecular solution reaction between the DNA and the reactive moiety. This rate enhancement might help design functionalities which can effect DNA modification/cleavage on desired time scales when tethered to DNA binding molecules.

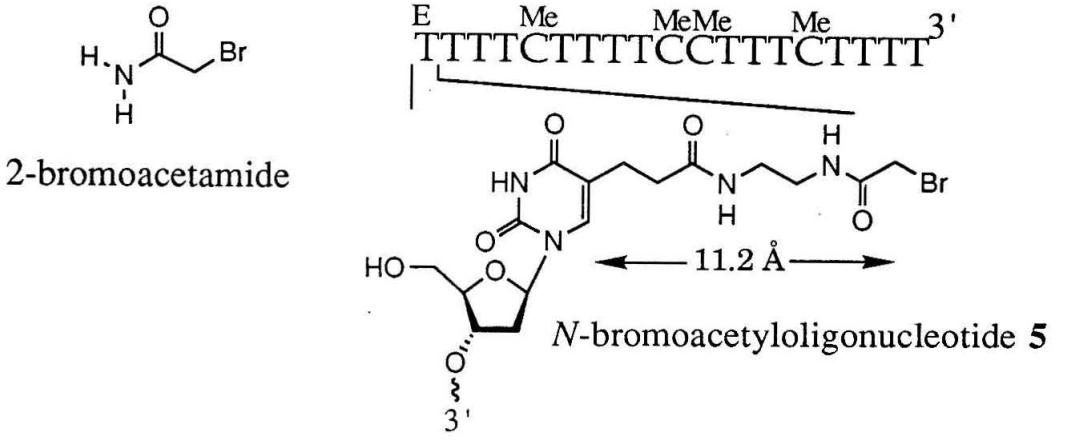


Fig. 4.25. Chemical structure and sequence of the modified terminal thymidine within *N*-bromoacetyloligonucleotide **5**.

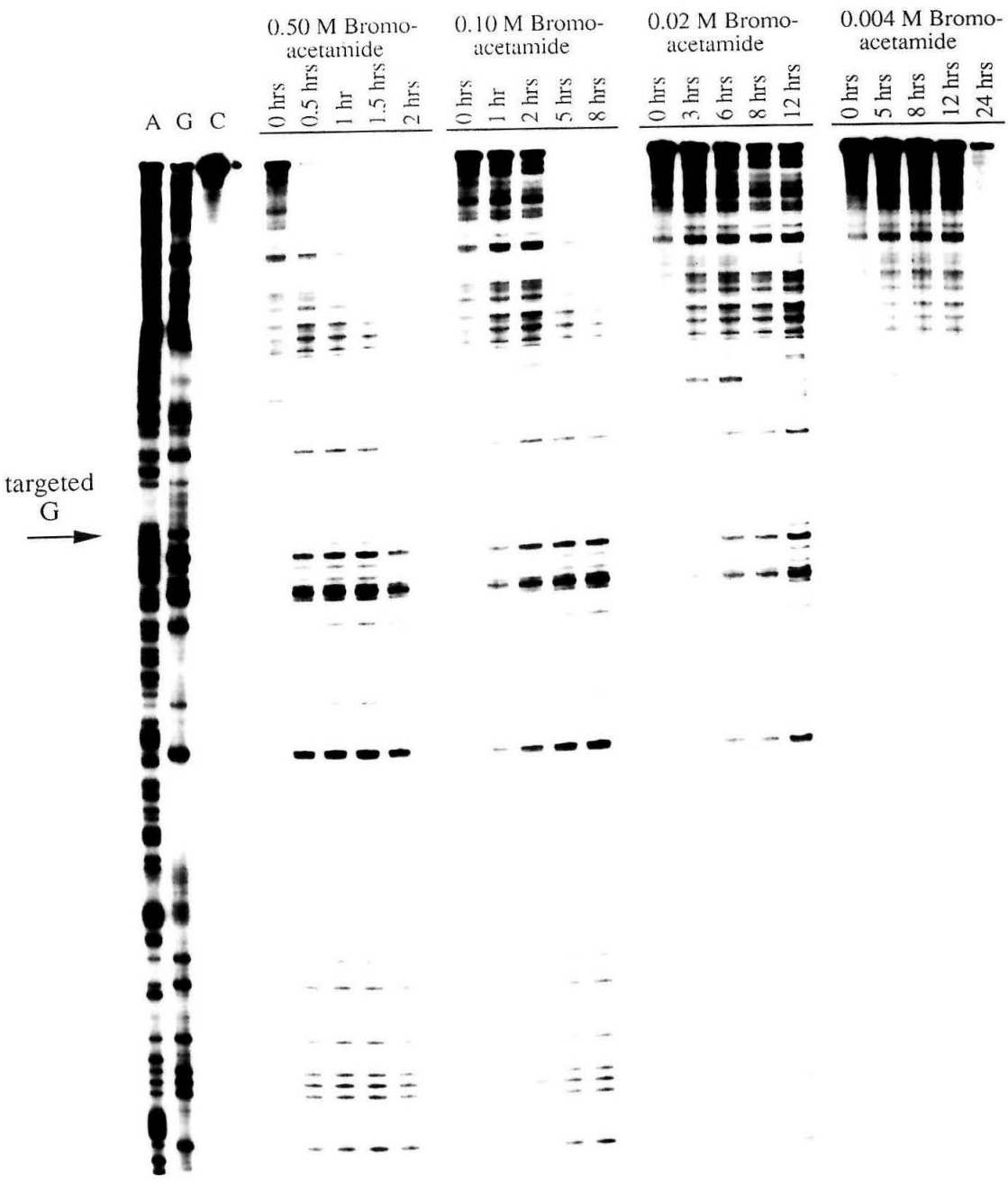
The *N*-bromoacetyloligonucleotide **5** contains a bromoacetamide moiety attached to the 5' terminal thymidine residue of a 19 base oligonucleotide (Figure 4.25),[61] which binds double helical DNA at a 19 base pair purine target site via triple helix formation (Fig. 4.24). *N*-Bromoacetyloligonucleotide **5** is capable of nearly quantitative modification of an appropriately positioned guanine base two base pairs to the 5' side of the purine target site within a 659

base pair fragment of duplex DNA.[61] To determine the rate enhancement obtained upon incorporation of the bromoacetamide functionality within a triple helical complex, the bimolecular and unimolecular rate constants for reaction of 2-bromoacetamide and *N*-bromoacetyloligonucleotide **5** with a single guanine within a DNA restriction fragment have been determined.

A ^{32}P-labeled 0.9 kbp restriction fragment (Nar I/EcoR I fragment from plasmid pUCLEU2C, labeled at the Nar I site) was incubated with various concentrations of 2-bromoacetamide (4-500 mM) in 25 mM NaCl, 20 mM Hepes, pH 7.2. Reactions were quenched at appropriate times by addition of DMS stop solution and ethanol precipitation.[62,63] After piperidine treatment (1.0% aqueous piperidine, 90 °C, 30 min) and lyophilization, the products were separated by gel electrophoresis, and quantitation was accomplished using phosphorimaging (Fig. 4.26).

The reactivity of the targeted guanine was determined by quantitating the amount of DNA which corresponds to reaction at the desired guanine relative to DNA which had not reacted at that guanine (Figure 4.24). Fragments smaller than the desired band are not quantitated, since it is not possible to determine if reaction at the designated guanine has occurred. All fragments larger than those cleaved at the desired band are considered "intact" DNA, as reaction at the targeted base has not occurred. The natural log of "intact" DNA/"total" DNA was plotted versus time to determine a pseudo-first order rate constant (Figure 4.27). The first order rate constant for reaction of this guanine within the *N*-bromoacetyloligonucleotide-DNA complex has been determined to be 3.6×10^{-5} sec^{-1}, corresponding to a half-life of 5.5 hours within the triple helical complex. This value has been measured and does not change over a *N*-bromoacetyloligonucleotide concentration range of 50 nM - 2 μM.

Fig. 4.26. Autoradiogram of a high resolution 6% denaturing polyacrylamide gel of cleavage products obtained after incubation of pUCLEU2C ^{32}P 3'-end-labeled DNA with various concentration of bromoacetamide as a function of time. DNA was dissolved in an aqueous solution of 50 mM NaCl, 40 mM pH 7.4 Hepes. An equal volume of bromoacetamide (2x solution) was added to the appropriate final concentration (in 25 mM NaCl, 20 mM Hepes, pH 7.4), and the reactions incubated at 37 °C. After the indicated time, the DNA was precipitated by addition of ethanol, washed, and treated with piperidine (1% piperidine, 90 °C, 30 min.). After lyophilization, the DNA was dissolved in formamide buffer and loaded onto a 6% polyacrylamide gel.


A G C
0.50 M Bromo-acetamide
0 hrs
0.5 hrs
1 hr
1.5 hrs
2 hrs
0.10 M Bromo-acetamide
0 hrs
1 hr
2 hrs
5 hrs
8 hrs
0.02 M Bromo-acetamide
0 hrs
3 hrs
6 hrs
8 hrs
12 hrs
0.004 M Bromo-acetamide
0 hrs
5 hrs
8 hrs
12 hrs
24 hrs
targeted G

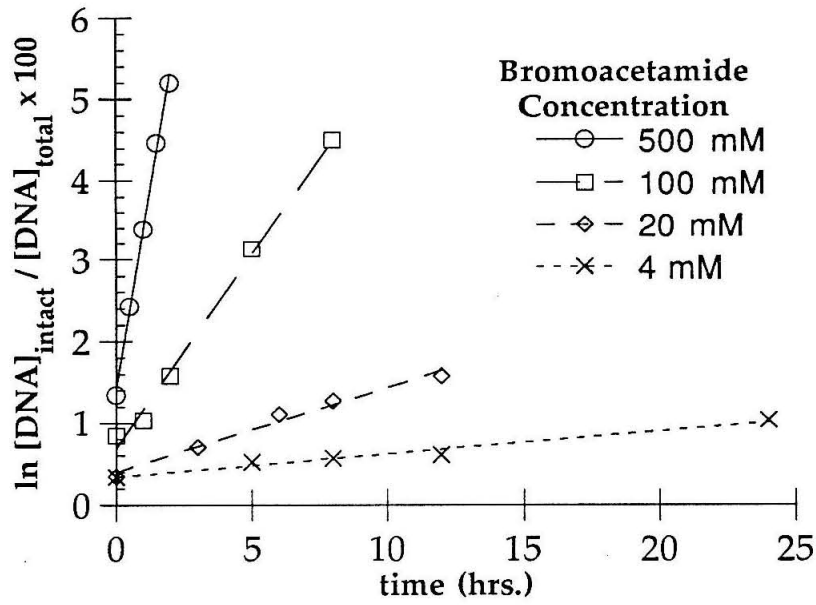


Figure 4.27: Determination of the bimolecular rate constant for reaction of bromoacetamide with double helical DNA. Each line represents a single time course at a different bromoacetamide concentration. Best fit lines give slopes with the following values: 500 mM bromoacetamide: 0.020; 100 mM bromoacetamide: 0.0048; 20 mM bromoacetamide: 0.0010; 4 mM bromoacetamide: 0.00028.

Table 4.1 contains the slopes, bimolecular rate constants, half-lives of the targeted guanine base, effective molarity, and nominal rate acceleration obtained for each time course. Bimolecular rate constants are obtained by division of the slope by the bromoacetamide concentration used for that time course. The determined bimolecular rate constant, the mean of seven separate time course experiments, is $1.5 \pm 0.4 \times 10^{-5}$, M^{-1} sec^{-1}. The plots in Figure 4.27 represent actual data from four of these sets of reactions derived from a single gel.

Bromo-acetamide conc.	slope of plot	bimol. rate constant (M^{-1} sec^{-1})	1/2 life	effective M of BRAA	rate acceleration
500 mM	0.0195	1.1×10^{-5}	36 hours	3.2 M	64×10^6
100 mM	0.0062	1.7×10^{-5}	116 hours	2.0 M	40×10^6
	0.0076	2.1×10^{-5}	92 hours	1.7 M	34×10^6
20 mM	0.0010	1.4×10^{-5}	690 hours	2.5 M	50×10^6
	0.0010	1.4×10^{-5}	690 hours	2.5 M	50×10^6
4 mM	0.00028	1.9×10^{-5}	2500 hours	1.8 M	36×10^6
	0.00015	1.0×10^{-5}	4600 hours	3.3 M	67×10^6

Table 4.1. Determination of bimolecular rate constant and effective molarity of *N*-bromoacetyloligonucleotides for the modification of double helical DNA. Slopes of plots (Fig. 4.27) are used to determine bimolecular rate constant, half-life for the targeted guanine base, effective molarity obtained upon tethering a bromoacetamide moiety to an triple-helix forming oligonucleotide.

The half-life of the targeted guanine represents the length of time required at the particular bromoacetamide concentration for one half of the guanines at that position to be alkylated. For comparison, the half-life of the targeted guanine within the triple helical complex is 5.5 hours. The effective molarity of bromoacetamide for each time course was determined by calculating the concentration of bromoacetamide required to obtain equal reaction rates in the bimolecular and intracomplex reactions. An effective molarity of 2.5 ± 0.5 M is obtained by averaging the values obtained from each time course. The reported rate acceleration is the ratio of the rates of reaction of the targeted guanine base with 50 nM *N*-bromoacetyloligonucleotide **5** vs. 50 nM bromoacetamide. This represents the lowest concentration of *N*-bromoacetyloligonucleotide **5** for which a rate of reaction was measured. Undoubtedly, the rate of the intra-complex reaction does not change drastically until oligonucleotide concentrations below the dissociation constant (< 16 nM) are reached. Since the determined rate acceleration is

dependent on the oligonucleotide concentration used to drive triplex formation, the effective molarity is a more accurate measure of the effect of tethering the reactive moiety to a DNA binding molecule.

There has been extensive discussion on the source of the rate enhancement observed in intramolecular reactions.[82-91] A simple calculation indicates that the rate enhancement observed in this study can be ascribed predominantly to localization of the reactive moiety close to the DNA. Assuming the linker arm to be quite flexible, it is possible to calculate the space accessible to the electrophilic carbon. This space is a hemisphere defined by a radius based on a fully extended all-trans tether (11.2 Å)[94] and by the DNA double helix, which sterically prevents the linker arm from the other half of the sphere. The volume of this hemisphere is approximately 5.9 x 10^{-21} cm^3. Confining a single molecule within this space results in an effective molarity of 0.5 M. The observed effective molarity is five times greater. A higher effective molarity is expected, as the entire volume of the hemisphere is not energetically equally accessible and the tether was designed to access the N-7 of the targeted guanine easily. Thus, the observed rate can be largely accounted for by confinement of the reactive moiety close to the DNA within the space accessible to the tether.

Theoretical concentration of tethered bromoacetyl based on volume calculation:

length of linker arm (radius):	11.2 angstroms
volume of 1/2 sphere:	5.8 x 10^{-24} dm^3
concentration:	3.4 x 10^{23} molecules/liter
effective concentration:	=0.5 M

The observed value also closely approximates that predicted based on entropic arguments proposed by Jencks and others.[82,83] These calculations indicate an intramolecular reaction to be favored over its bimolecular counterpart by approximately 35 e.u. under ideal circumstances. To account for loss of rotational entropy in the product, 3.3 and 6.2 entropy units are subtracted for each sp^3-sp^3 and sp^3-sp^2 bond whose rotation is restricted in the product. Using a crude calculation of this type to predict the effective molarity expected for *N*-bromoacetyloligonucleotide-directed alkylation of DNA, a calculated value of 5 M is obtained.[82,83]

Conclusions

The bimolecular rate constant for reaction of bromoacetamide with a specific guanine base in double helical DNA has been measured and is reported to be $1.5 \pm 0.4 \times 10^{-5}$, M^{-1} sec^{-1}. Delivery of a reactive moiety to a DNA target by covalent attachment with a flexible tether to a DNA binding domain therefore results in an effective molarity at the DNA target site of 2-3 M. The rate of the intra-complex reaction shows no change over the *N*-bromoacetyloligonucleotide concentration range of 50 nM to 2 μM. At a concentration of 50 nM, the ratio of the rates of the bimolecular and oligonucleotide-directed alkylation reactions is 5×10^7.

This work has implications for the design of sequence specific DNA modification reagents, since it defines the inherent reactivity necessary to afford a sequence specific cleaver which is capable of efficient cleavage on a desired time scale.

Part IV

Use of *N*-Bromoacetyloligonucleotides to Measure Binding Affinities and Cooperative Interactions

The high cleavage yield observed using the alkylation/base treatment method of modifying and cleaving DNA makes analysis of alkylation efficiency an ideal method of studying DNA binding by molecules with weak affinities, or in studies where it is desirable to measure small changes in binding affinities. *N*-bromoacetyloligonucleotides have been used to study the binding of short oligonucleotides and to quantitate changes in binding affinity due to external forces such as cooperative interactions.

Effect of oligonucleotide length on alkylation/cleavage efficiency.

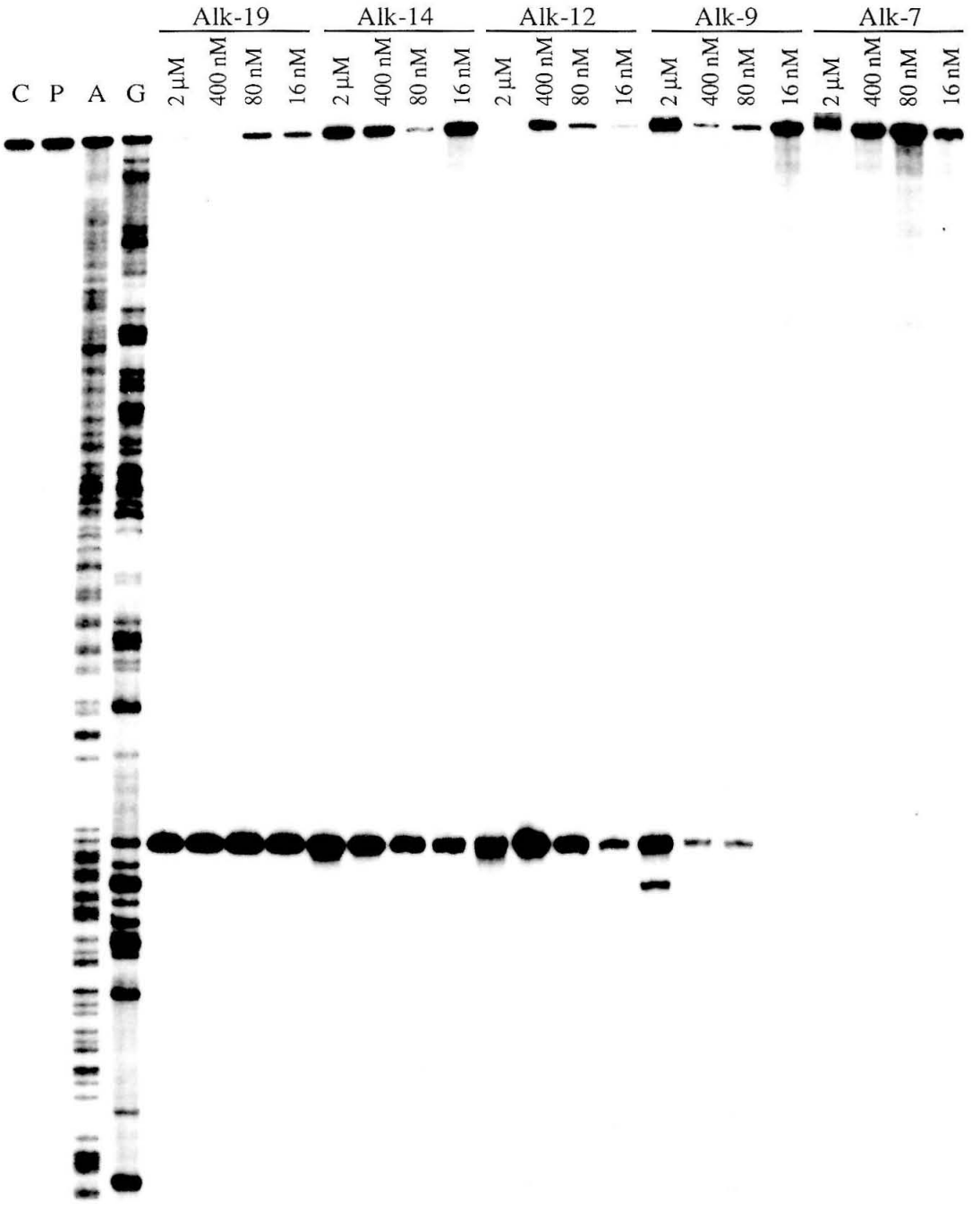
Specific double strand cleavage at a single binding site was observed when the *N*-bromoacetyloligonucleotide Alk-19 was used to cleave yeast chromosome III (340 kilobase pairs of DNA). *N*-bromoacetyloligonucleotide Alk-19 is identical to *N*-bromoacetyloligonucleotide **5**. Oligonucleotides in this section are named Alk-N, where N = length of oligonucleotide for ease of discussion. The specificity observed is a function of both the increased specificity of the cleaving function used and the requirement for oligonucleotide binding sites on opposite strands within a limited distance (< 20 bp) for double strand cleavage to be observed. Double strand cleavage at the target site involves recognition of 40 base pairs of DNA (two 19 base pair target sequences and two guanine bases). Statistically, two purine sites amenable to cleavage by alkylation are expected to occur once every 4×10^{12}

base pairs. Due to the overrepresentation of polypurine sequences in genomic DNA[33,34] and the ability of triple helix forming oligonucleotides to bind to sequences of imperfect homology,[26,27] it is to be expected that the frequency of cleavage will be greater than statistics would indicate. However, if this motif is to be used as a tool for genomic mapping and cleaving at sites of interest, it will be necessary to recognize and cleave sequences with greater frequency. In addition, shorter oligonucleotides are thought to traverse biological membranes with greater efficiencies.[95] These considerations make it desirable to understand the affinities and alkylation efficiencies of short oligonucleotides.

Fig. 4.28. *N*-Bromoacetyloligonucleotides synthesized for the study of oligonucleotide length on DNA alkylation/cleavage efficiency. *N*-Bromoacetyloligonucleotides bind to sites (boxed) on each of the two strands of the DNA target helix, alkylating a single guanine (indicated by the arrow) on each strand. Upon piperidine treatment, cleavage at the site of alkylation is expected.

Binding affinity as a function of oligonucleotide length. Oligonucleotides Alk-19 through Alk-7 were synthesized from T, Me^5C, and T^E (**3**) phosphoramidites to study the effect of oligonucleotide length on alkylation efficiency, and to determine if DNA modification can be

Fig. 4.29. Alkylation of plasmid DNA as a function of *N*-bromoacetyloligonucleotide length and concentration. Autoradiogram of a high-resolution 6% denaturing polyacrylamide gel of cleavage products from the reaction of oligonucleotides CT-07 through CT-19 with a ^{32}P 3'-end-labeled 0.9 kbp restriction fragment (*Nar I/ EcoR I*) from plasmid pUCLEU2C. Reaction conditions were *N*-bromoacetyloligonucleotide at stated concentration, 20 mM Hepes, pH 7.0, 1.0 mM $Co(NH_3)^{+3}$, 10 mM EDTA, and 10,000 cpm end-labeled DNA in a total volume of 15 µl. Reactions were incubated at 37 °C for 36 hours, precipitated with NaOAc/EtOH, washed with 70% EtOH, and treated with 0.1 M piperidine (90 °C, 30 min.). After lyophilization, cleavage products were analyzed on a 6% 1:20 cross-linked, 48% urea polyacrylamide gel, 0.4 mm thick. Lane 1: untreated DNA. Lane 2: piperidine treated DNA. Lane 3: A sequencing reaction. Lane 4: G sequencing reaction. Lanes 5-9: DNA incubated with *N*-bromoacetyloligonucleotide Alk-19 at 2 µM (lane 5), 400 nM (lane 6), 80 nM (lane 7), and 16 nM (lane 8). Lanes 10-13: DNA incubated with *N*-bromoacetyloligonucleotide Alk-14 at 2 µM (lane 10), 400 nM (lane 11), 80 nM (lane 12), and 16 nM (lane 13). Lanes 14-17: DNA incubated with *N*-bromoacetyloligonucleotide Alk-12 at 2 µM (lane 14), 400 nM (lane 15), 80 nM (lane 16), and 16 nM (lane 17). Lanes 18-21: DNA incubated with *N*-bromoacetyloligonucleotide Alk-09 at 2 µM (lane 18), 400 nM (lane 19), 80 nM (lane 20), and 16 nM (lane 21). Lanes 22-25: DNA incubated with *N*-bromoacetyloligonucleotide Alk-07 at 2 µM (lane 22), 400 nM (lane 23), 80 nM (lane 24), and 16 nM (lane 25).


Alk-19
Alk-14
Alk-12
Alk-9
Alk-7
C P A G
2 µM
400 nM
80 nM
16 nM

achieved in high yield with shorter sequence requirements (Fig. 4.28). Derivatization of the oligonucleotides was accomplished using *N*-hydroxysuccinimidyl bromoacetate. Each was purified by anion exchange HPLC, with small changes in the solvent gradient to account for the change in charge associated with differences in oligonucleotide length. Separation and ease of purification increased inversely with oligonucleotide length: shorter oligonucleotides eluted in sharper bands with larger differences in the retention times of the amino- and *N*-bromoacetyl-oligonucleotides.

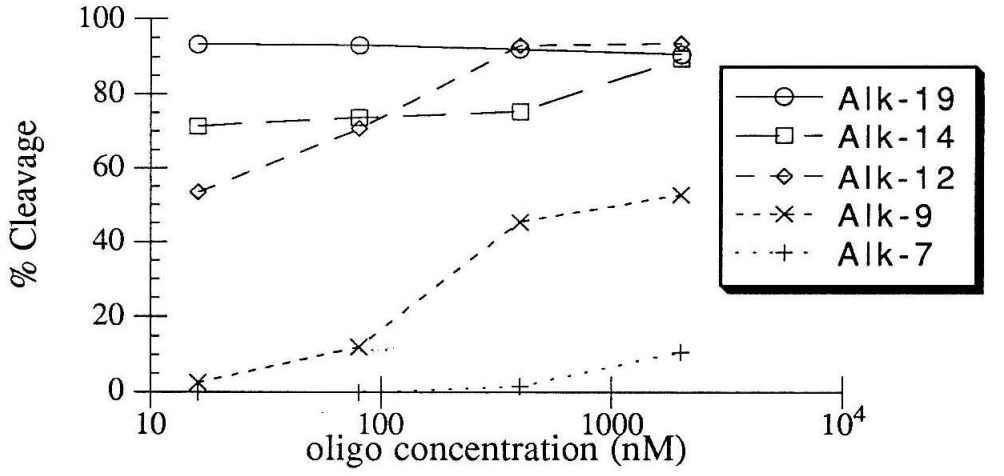


Fig. 4.30. Single strand alkylation/cleavage efficiency by *N*-bromoacetyloligonucleotides of various lengths as a function of oligonucleotide concentration. Data are obtained by phosphorimaging of gel in Figure 4.29.

Figures 4.29 and 4.30 indicate the efficiencies of alkylation observed by oligonucleotides of various lengths as a function of concentration.

Approximate dissociation constants for the *N*-bromoacetyloligonucleotides are estimated from the data in Figure 4.30 and are listed below:

N-Bromoacetyloligonucleotide	Dissociation constant
Alk-19	< 10 nM
Alk-14	< 10 nM
Alk-12	~ 15 nM
Alk-9	~ 400 nM
Alk-7	>> 2 μM

There is a sharp decrease in site saturation under the conditions used (37 °C, neutral pH) with oligonucleotides < 12 nucleotides long. This is corroborated by results obtained using DMS footprinting (Chapter II), which indicate that at an oligonucleotide concentration of 2 μM, detectable binding is observed for the 9 base oligonucleotide CT-09 at 0° and 12.5°, but not at room temperature. These studies agree that efficient binding at physiological temperature and pH requires oligonucleotides ≥ 12 nucleotides in length.

Specificity of N-bromoacetyloligonucleotides. Oligonucleotides Alk-12, Alk-14 and Alk-19 modify a single guanine within the 0.9 kilobase pair restriction fragment. Oligonucleotide Alk-9 modifies at least 2 guanines: the targeted guanine 5' to the 19 base purine site, and at least one of the two central guanines within the purine target sequence. This indicates that oligonucleotide Alk-9 binds at least two sites within the purine sequence. Binding to the second site requires the formation of at least one mismatch and occurs with lower affinity, since, at low oligonucleotide concentration, cleavage occurs only at the guanine 5' to the target site. No modification of any of the adenines within the target sequence is observed, indicating the base specificity of the cleavage reaction.

Statistical significance of cleavage by short N-bromoacetyloligonucleotides. Double strand cleavage by *N*-bromoacetyloligonucleotides requires the presence of two oligonucleotide binding sites on opposite strands with guanines properly positioned for attack. The two sites must be located within a reasonable distance (≤ 20 base pairs) to allow the strands to denature and separate. Statistically, sites amenable to cleavage by *N*-bromoacetyloligonucleotides nine bases long are expected to occur once every 2×10^5 base pairs ($[0.5^9 \times 0.25]^2 \times 20$). Sites for cleavage by two 12 or 14 base oligonucleotides are expected once every 1.3×10^7 ($[0.5^{12} \times 0.25]^2 \times 20$) and once every 2.2×10^8 ($[0.5^{14} \times 0.25]^2 \times 20$) base pairs. These calculations assume that only purine sites can be bound (i. e. mismatches are not tolerated), and that a statistical distribution of purine sites exists. Since the triple helix motif can be used to recognize sites which contain all four base pairs of DNA,[20,96] and since purine sequences are overrepresented in genomic DNA,[33,34] it is expected that potential cleavage sites will occur more frequently than statistically expected.

To be used as tools for genetic mapping, rare cutting cleaving strategies must be able to cut DNA at genetic markers of interest. Markers are usually within cosmid size DNA (50,000-250,000 base pairs). These calculations indicate that 9-10 base *N*-bromoacetyloligonucleotides could be used to produce cleavage with the frequency required to find and cut sites within this size range of DNA.

Cleavage of DNA using degenerate N-bromoacetyloligonucleotides. Techniques for the determination of triple helical target sites in unsequenced DNA based on the use of degenerate oligonucleotides have recently been developed.[31] Oligonucleotides are synthesized from a mixture of T and C nucleoside phosphoramidites. 5-Bromo-2'-deoxyuridine and 5-methyl-2'-

deoxycytidine are often substituted for thymidine and 2'-deoxycytidine to increase the binding affinities of the resulting oligonucleotides. The synthesis of degenerate oligonucleotides results in the production of 2^n different sequences, where n is the number of degenerate positions. While oligonucleotide binding affinity increases with oligonucleotide length, the concentration of the single perfectly matched oligonucleotide for a given target site decreases as the degeneracy increases.

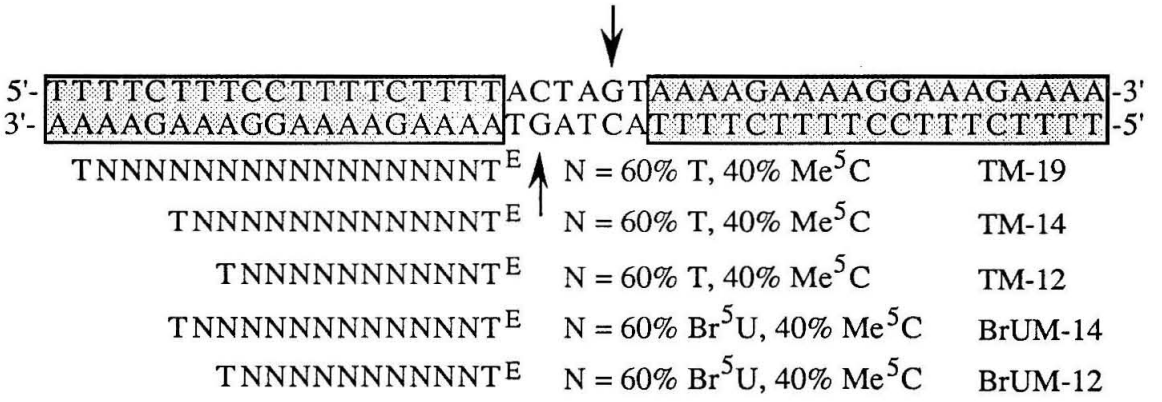


Fig. 4.31. Oligonucleotides degenerate in T and Me^5C or Br^5U and Me^5C for the determination of possible triple helix/alkylation sites within unsequenced DNA. The compositions of nucleotides used to synthesize these oligonucleotides are listed. Oligonucleotides bind to target sites on both strands (boxed), alkylating a single G on each strand (arrow), effecting double strand cleavage.

Since perfectly matched oligonucleotides < 12 bases in length display low binding affinities even at high concentrations, degenerate *N*-bromoacetyloligonucleotides ≥ 12 bases (12, 14 and 19 bases) in length containing 10, 12, and 17 degenerate positions were synthesized (Fig. 4.31). Each of these oligonucleotides contains a 3' terminal thymidine and a 5'

terminal amino-thymidine residue. The oligonucleotides were synthesized using a 60:40 mixture of T:Me^{5}C or Br^{5}U:Me^{5}C to skew the distribution of oligonucleotides toward those with higher T content. Oligonucleotides with high C content are not expected to form stable pyrimidine•purine-pyrimidine triple helical complexes (see Chapters I and II). The synthesis of degenerate *N*-bromoacetyloligonucleotide 12-, 14-, and 19mers, degenerate at 10, 12 and 17 positions results in the production of 1024, 4096, and 131072 different sequences. At a total oligonucleotide concentration of 2 μM, a concomitant decrease in the concentration of the perfectly matched oligonucleotide from 2 to 0.5 to 0.015 nM results. At high oligonucleotide concentrations and low pH, the binding of partially mismatched oligonucleotides is expected.

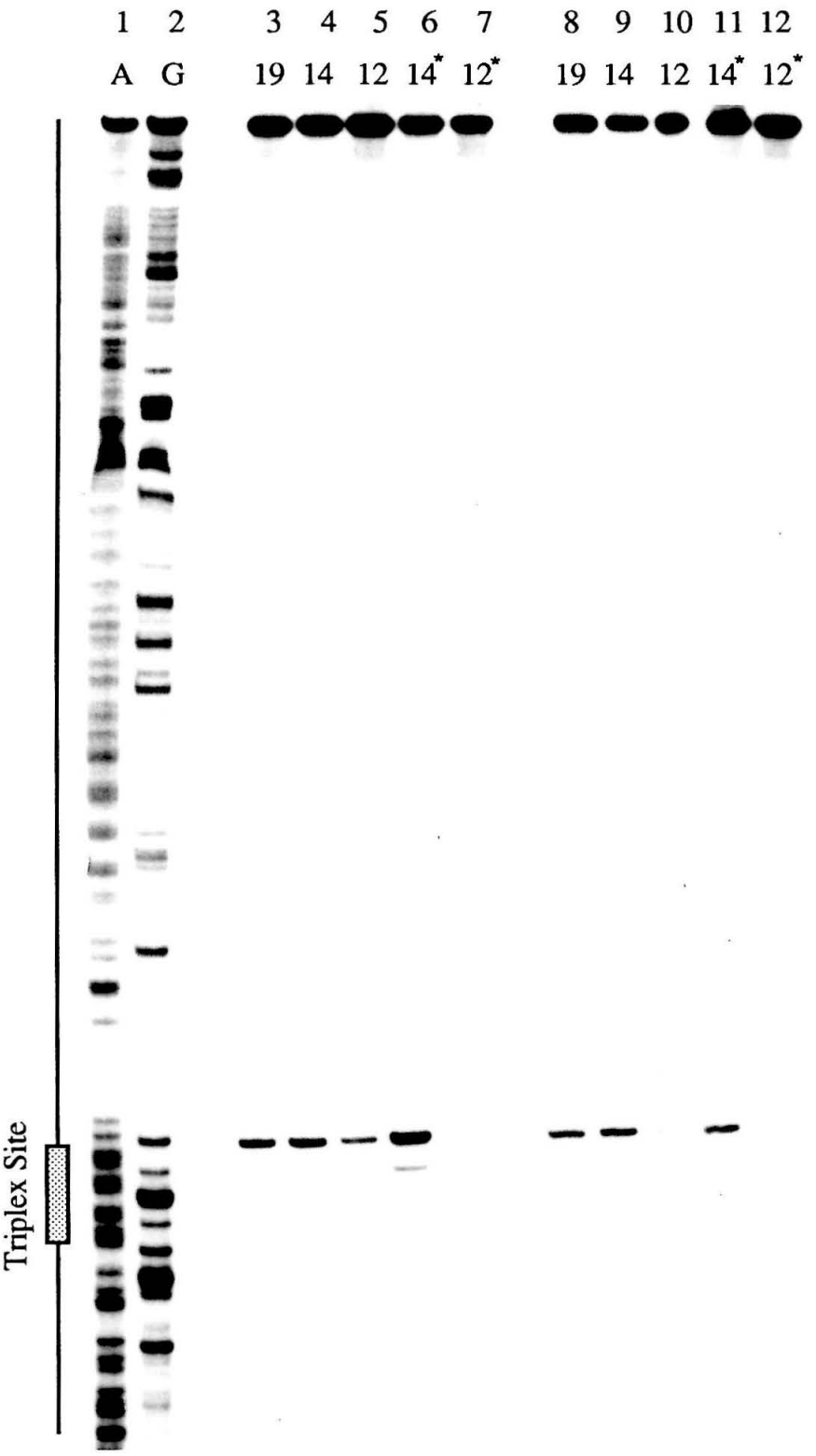
After synthesis and deprotection, amino-oligonucleotides were purified by polyacrylamide gel electrophoresis as a single band of normal size. Derivatization of the oligonucleotides with *N*-hydroxysuccinimidyl bromoacetate was accomplished as previously described. Derivatized degenerate *N*-bromoacetyloligonucleotides could not be purified by HPLC, because these oligonucleotides eluted from the column in a peak several minutes wide. There was no detectable change in the HPLC profile upon derivatization. Purification was accomplished by removal of excess *N*-hydroxysuccinimidyl bromoacetate using Sephadex chromatography (1 ml G-25 spin column[77]) for the T/Me^{5}C degenerate oligonucleotides. It was later observed that more efficient removal of salts could be accomplished using Sep-pak reverse phase cartridges. Degenerate Br^{5}U/Me^{5}C oligonucleotides were purified in this manner. Purification by Sep-pak cartridges was accomplished by washing the cartridge with 20 ml of water to effect complete removal of salts.

Derivatization of amino-oligonucleotides typically results in a high yield of bromoacetylation (> 90%). Derivatization of the degenerate oligonucleotides used in this study is expected to result in a similar yield of conversion to the *N*-bromoacetyloligonucleotides. It is possible that the binding affinities of the amino-oligonucleotides exceed that of derivatized *N*-bromoacetyloligonucleotides, due to the presence of the protonated amine. Since the amino-oligonucleotides are not separated from the derivatized product, competition for binding sites and a decrease in alkylation efficiency might result.

Binding of degenerate *N*-bromoacetyloligonucleotides to potential oligonucleotide binding sites containing guanines appropriately positioned for double strand alkylation and cleavage is expected to result in the production, albeit in low yield, of cleavage bands on high resolution sequencing (polyacrylamide) gels and coarse resolution double strand (agarose) gels. Isolation and sequencing of the bands produced could be used to identify triple helix/alkylation sites within cosmids or genetic markers of unknown sequence.

To determine whether degenerate oligonucleotides can bind, alkylate and effect cleavage at a triple helix site, the plasmid pUCALK2C, containing the known triple helix/alkylation site previously studied, was incubated with degenerate *N*-bromoacetyloligonucleotides of various lengths. Analysis for single stranded modification and cleavage using polyacrylamide sequencing gels indicates that degenerate oligonucleotides cleave each strand of the duplex with an efficiency of 15 - 20% on each strand (Figs. 4.32 and 4.33). This efficiency is observed at an ideal pyrimidine oligonucleotide binding site. Binding and cleavage efficiency at less ideal sites (i. e. higher G content, presence of mismatches) is expected to be lower. These efficiencies are slightly

Figure 4.32. Alkylation by degenerate oligonucleotides. Autoradiogram of a high-resolution 6% denaturing polyacrylamide gel of cleavage products from the reaction of degenerate oligonucleotide CT-19, CT-14, CT-12, Br^5UC-14, and Br^5UC-12 with a ^{32}P 3'-end-labeled 0.9 kbp restriction fragment (*Nar I/ EcoR I*) from plasmid pUCLEU2C. Reaction conditions were 2 or 10 μM total concentration of *N*-bromoacetyloligonucleotide, 20 mM Hepes, pH 6.8, 1.0 mM $Co(NH_3)^{+3}$, and 10,000 cpm end-labeled DNA in a total volume of 15 μl. Reactions were incubated at 37 °C for 36 hours, precipitated with NaOAc/EtOH, washed with 70% EtOH, and treated with 0.1 M piperidine (90 °C, 30 min.). After lyophilization, cleavage products were analyzed on a 6% 1:20 cross-linked, 48% urea polyacrylamide gel, 0.4 mm thick. Lane 1: A sequencing reaction. Lane 2: G sequencing reaction. Lanes 3-7: pUCLEU2C DNA incubated with 10 μM degenerate *N*-bromoacetyloligonucleotide. Lanes 8-12: pUCLEU2C DNA incubated with 2 μM degenerate *N*-bromoacetyloligonucleotide. Lanes 3 and 8: pUCLEU2C DNA incubated with degenerate oligonucleotide TC-19. Lanes 4 and 9: pUCLEU2C DNA incubated with degenerate oligonucleotide TC-14. Lanes 5 and 10: pUCLEU2C DNA incubated with degenerate oligonucleotide TC-12. Lanes 6 and 11: pUCLEU2C DNA incubated with degenerate oligonucleotide Br^5U/Me^5C-14. Lanes 7 and 12: pUCLEU2C DNA incubated with degenerate oligonucleotide Br^5U/Me^5C-12.


1 2 3 4 5 6 7 8 9 10 11 12
A G 19 14 12 14* 12* 19 14 12 14* 12*
Triplex Site

lower than those observed by Strobel using degenerate oligonucleotides to block the action of a methylase at a methylase/triple helix site.[31] Degenerate oligonucleotides afford approximately 25% protection from methylation at a similar ideal triple helix site. The T/Me^{5}C were desalted using Sephadex columns, the Br^{5}U/Me^{5}C oligonucleotides were purified using Sep-pak cartridges. The decreased cleavage efficiencies observed with the T/Me^{5}C degenerate oligonucleotides at higher concentrations is due to the incomplete removal of salts from the oligonucleotide derivatization reaction, as judged by the UV spectrum of these oligonucleotides.

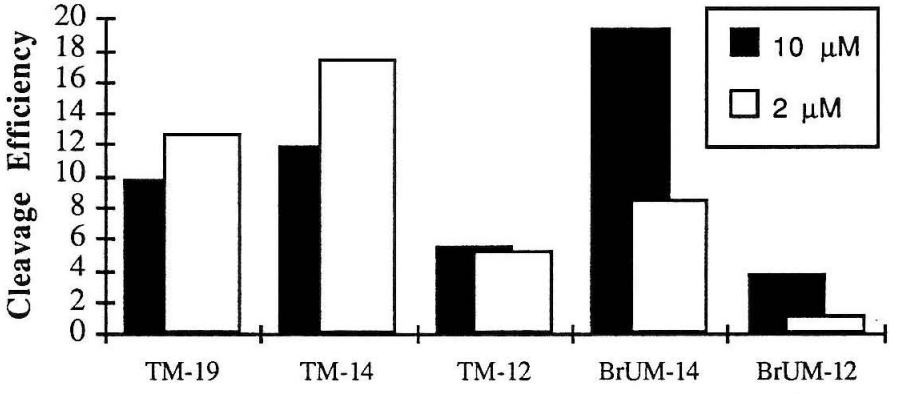


Fig. 4.33: Cleavage efficiencies observed at known triple helix/alkylation site upon incubation of degenerate *N*-bromoacetyloligonucleotides of various lengths and base composition with DNA from pUCLEU2C at total oligonucleotide concentrations of 10 μM and 2 μM.

Effect of degenerate oligonucleotide length on cleavage efficiency. Degenerate *N*-bromoacetyloligonucleotides 14 bases in length alkylate DNA with greater efficiencies than 12 or 19 base long oligonucleotides. This suggests that the increase in binding affinity observed upon formation of

more than 14 base triplets does not compensate for the lower oligonucleotide concentration due to the increased degeneracy. Decreasing the number of base triplets also lowers binding affinity. The four fold increase in individual oligonucleotide concentration does not compensate for the loss of two base triplets. As was observed with perfectly matched oligonucleotides, 12-14 bases represents the minimal oligonucleotide length required for binding at physiological pH and temperature. These data suggest that the rapid screening of unsequenced DNA by degenerate oligonucleotides can most efficiently be accomplished using oligonucleotides approximately 14 bases in length.

The expected yield of double strand cleavage is the square of the single strand modification efficiency. Degenerate *N*-bromoacetyloligonucleotides are therefore expected to effect double strand cleavage with efficiencies of less than 5% $(0.2)^2$. This has been verified using agarose gel analysis of pUCLEU2C cleavage and pulse field gel electrophoretic analysis of yeast chromosomal cleavage by *N*-bromoacetyloligonucleotides.

Specificity of cleavage. Degenerate oligonucleotides can bind in a number of ways within the 19 base purine target sequence in plasmid pUCLEU2C. Alkylation occurs exclusively at the guanine bases, indicating the base specificity of the alkylation reaction. *N*-Bromoacetyloligonucleotides 19 bases in length bind the entire site and cleave only the single targeted guanine. *N*-Bromoacetyloligonucleotides 14 and 12 bases in length bind within the target site in a manner which allows modification of guanines both internal and 5' to the target sequence. This indicates the ability of degenerate oligonucleotides to bind to varying sequences, and suggests that cleavage at

long purine runs frequently found in eukaryotic DNA[33,34,97,98] will occur non-specifically at guanines located throughout such sequences.

Use of degenerate N-bromoacetyloligonucleotides to determine triplex/alkylation sites in unsequenced DNA. The use of degenerate *N*-bromoacetyloligonucleotides to find and sequence triplex/alkylation sites in unsequenced DNA requires: (i) isolation of the fragments produced by site specific double helical cleavage; (ii) digestion of the products to produce blunt-ended fragments; and (iii) cloning of the fragments into a vector to facilitate sequencing. Due to the low yield of double helical cleavage observed with *N*-bromoacetyloligonucleotides, these manipulations are expected to be difficult to accomplish with the small amount of DNA produced.

Furthermore, sites recognizable by the bis-alkylation motif are expected to arise much less frequently than triplex/restriction enzyme (Achilles heel) sites. For instance, efficient binding requires the use of oligonucleotides ≥ 12 bases in length. A 12 base pair purine site overlapping a 6 base pair restriction enzyme site is statistically expected to occur once every ~ 500,000 base pairs (a 6-base pair restriction enzyme cuts one in 2048 sequences; one out of every 256 of these sites ($2^9/2$) will be flanked by nine consecutive purines on either of the two sides). On the other hand, double strand cleavage by 12 base *N*-bromoacetyloligonucleotides requires <u>two</u> 12 base pair binding sites within 20 base pairs of each other, each with an appropriately positioned guanine. Such a sequence is statistically expected to occur once every 13.4×10^6 base pairs (a 12 base purine site occurs once every 2^{12} bases; one out of four of these contains an appropriately positioned guanine base, two such sites within 20 base pairs of each other occur once every $[4 \times 2^{12}]^2 / 20$, or 13.4×10^6 base pairs). These calculations are based on the requirements for perfect purine runs, and are

presented to compare the relative frequency of these types of sites. These considerations suggest that double strand cleavage sites for *N*-bromoacetyloligonucleotides probably do not occur with a frequency which will allow this motif to play a useful role as a rare-cutting strategy to be used in genomic mapping.

Cooperative Binding of *N*-Bromoacetyloligonucleotides to Abutting Sites

Biological systems often use cooperative interactions to increase the sensitivity of DNA binding molecules both to base mismatches and to small concentration changes.[99-103] Cooperative binding by oligonucleotides to DNA may offer several advantages: (i) shorter oligonucleotides may more readily penetrate membrane barriers around the cell and the nucleus;[95] and (ii) the sensitivity of binding affinity to the presence of mismatches may in some cases be increased.

Strobel and Dervan have reported that the binding affinity of a oligonucleotide-EDTA•Fe(II) 9 bases in length increases by a factor of 3.4 in the presence of an oligonucleotide which binds an abutting site.[25] Binding of oligonucleotides to adjacent sites on single stranded RNA has been shown to increase the effectiveness of short oligonucleotides as inhibitors of transcription.[104,105] The enhanced binding is likely due to favorable stacking interactions and/or induced conformation changes propagated to adjacent binding sites.[25] Oligonucleotides equipped with dimerization domains capable of forming Watson-Crick hydrogen bonds bind cooperatively to adjacent triple helix sites with a 15 fold increase in binding affinity.[23]

The cooperative binding of two oligonucleotides to abutting sites was studied using alkylation as a measure of site saturation. The high efficiency of alkylation allows accurate determination of any increase in binding affinity due to cooperative binding of the two oligonucleotides.

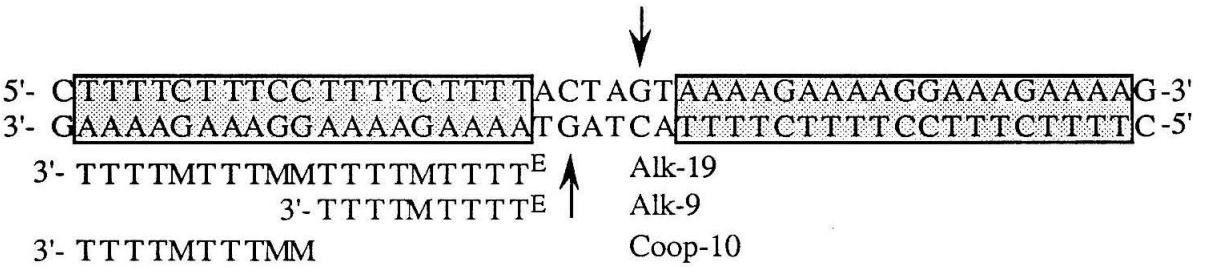
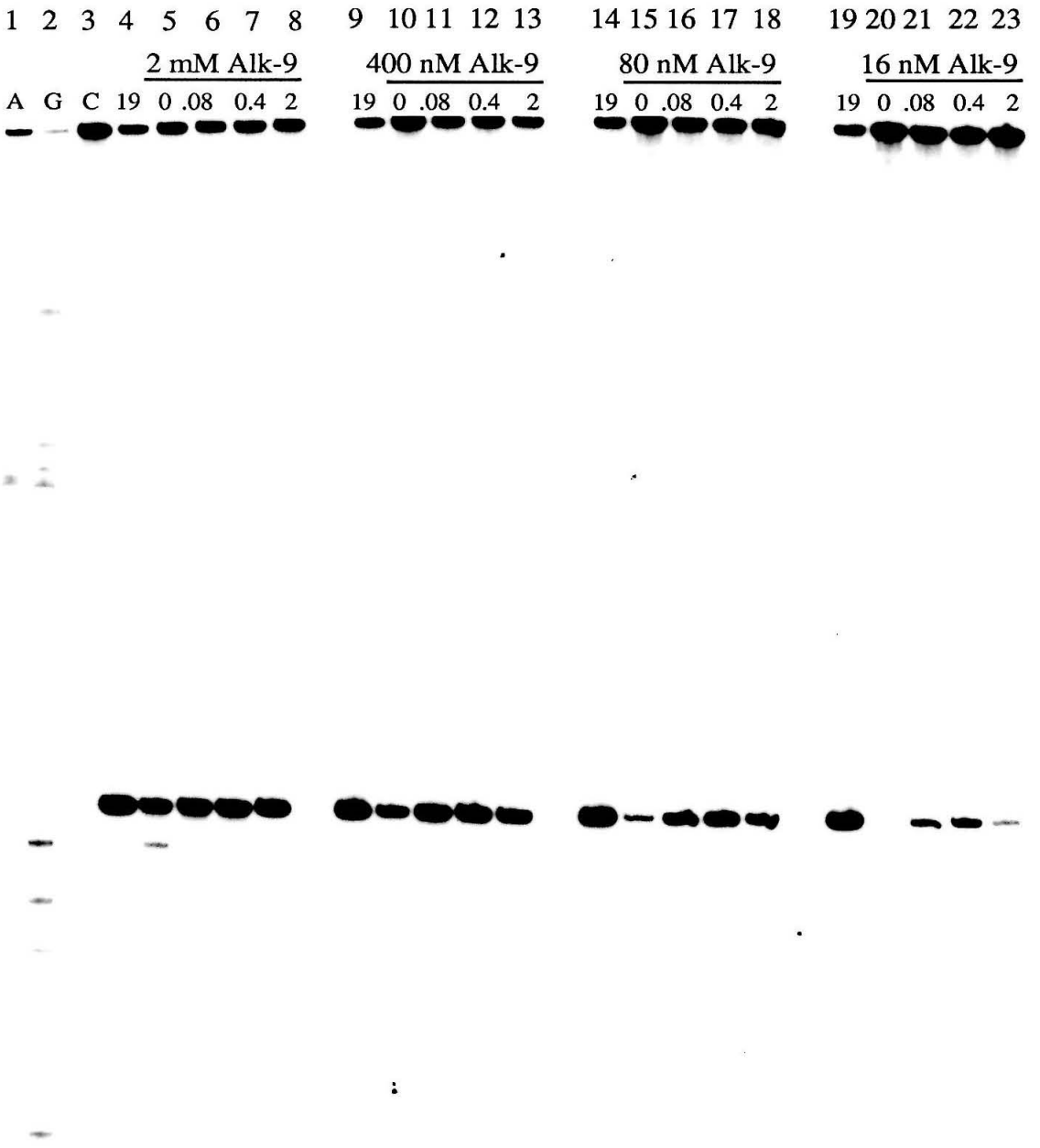


Fig. 4.34. Oligonucleotides used to study cooperative end-to-end interactions in oligonucleotide-directed triple helix formation. Oligonucleotide binding sites are indicated by the boxes and guanine bases targeted for alkylation are indicated by the arrows. Oligonucleotides bind to sites on each of the two strands; thus, both single and double strand cleavage can be studied.

Oligonucleotide Alk-9 binds a 9 base pair site, with a guanine positioned for efficient alkylation two base pairs to the 5' side of the binding site. Oligonucleotide Coop-10 binds a 10 base pair site adjacent to the binding site for Alk-9 (Fig. 4.34). The efficiency of alkylation of the targeted guanine base by oligonucleotide Alk-9 was determined at four different Alk-9 concentrations in the presence of several concentrations of Coop-10 (Fig. 4.35). The binding affinity of the 19 base long oligonucleotide Alk-19 was also determined, to allow comparison of the effects of non-covalent (cooperative) vs. covalent linking of the oligonucleotides.

Alk-9 alkylates guanines both 5' and internal to the 18 base purine site sequence, indicating that the oligonucleotide binds in at least two different locations. Addition of Coop-10 inhibits binding at the secondary site, due to

Fig. 4.35. Cooperative Binding of Oligonucleotides Alk-9 and Coop-10 as assayed by extent of alkylation. Autoradiogram of a high-resolution 6% denaturing polyacrylamide gel of cleavage products from the reaction of *N*-bromoacetyloligonucleotide Alk-19 or *N*-bromoacetyloligonucleotide Alk-9 with a ^{32}P 3'-end-labeled 0.9 kbp restriction fragment (*Nar I/ EcoR I*) from plasmid pUCLEU2C. Reaction conditions were *N*-bromoacetyloligonucleotide Alk-19, *N*-bromoacetyloligonucleotide Alk-9, and/or oligonucleotide Coop-10 at stated concentrations, 20 mM Hepes, pH 7.0, 1.0 mM $Co(NH_3)^{+3}$, and 10,000 cpm end-labeled DNA in a total volume of 15 µl. Reactions were incubated at 37 °C for 36 hours, precipitated with NaOAc/EtOH, washed with 70% EtOH, and treated with 0.1 M piperidine (90 °C, 30 min.). After lyophilization, cleavage products were analyzed on a 6% 1:20 cross-linked, 48% urea polyacrylamide gel, 0.4 mm thick. Lane 1: A sequencing reaction. Lane 2: G sequencing reaction. Lane 3: pUCLEU2C restriction fragment incubated without oligonucleotide and piperidine treated. Lane 4: 2 µM *N*-bromoacetyloligonucleotide Alk-19. Lane 9: 400 nM *N*-bromoacetyloligonucleotide Alk-19. Lane 14: 80 nM *N*-bromoacetyloligonucleotide Alk-19. Lane 19: 16 nM *N*-bromoacetyloligonucleotide Alk-19. Lanes 5-8: 2 µM *N*-bromoacetyloligonucleotide Alk-9. Lanes 10-13: 400 nM *N*-bromoacetyloligonucleotide Alk-9. Lanes 15-18: 80 nM *N*-bromoacetyloligonucleotide Alk-9. Lanes 20-23: 16 nM *N*-bromoacetyloligonucleotide Alk-9. Lanes 6, 11, 16, 21: 80 nM oligonucleotide Coop-10. Lanes 6, 12, 17, 22: 400 nM oligonucleotide Coop-10. Lanes 8, 13, 18, 23: 2 µM oligonucleotide Coop-10.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
2 mM Alk-9
400 nM Alk-9
80 nM Alk-9
16 nM Alk-9
A G C 19 0 .08 0.4 2 19 0 .08 0.4 2 19 0 .08 0.4 2 19 0 .08 0.4 2

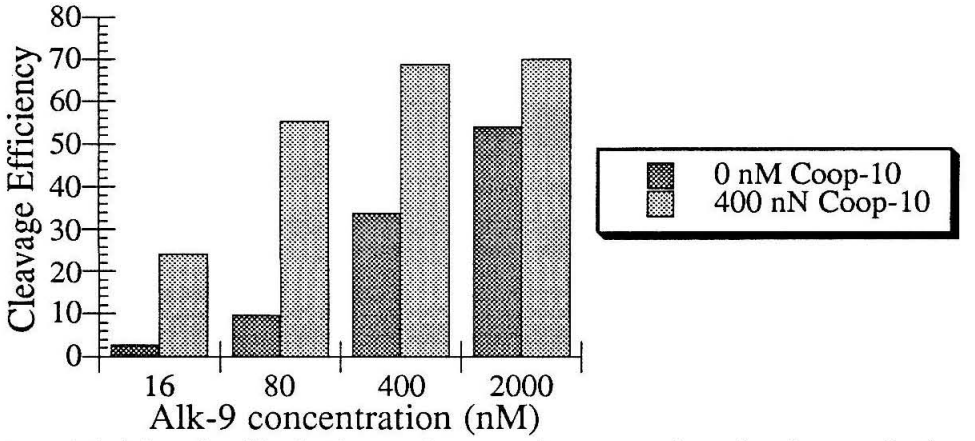


Fig. 4.36. Yield of alkylation observed upon incubation of the *N*-bromoacetyloligonucleotide Alk-9 with pUCLEU2C DNA at four different concentrations of Alk-9 in the presence and absence of 400 nM of the abutting oligonucleotide Coop-10.

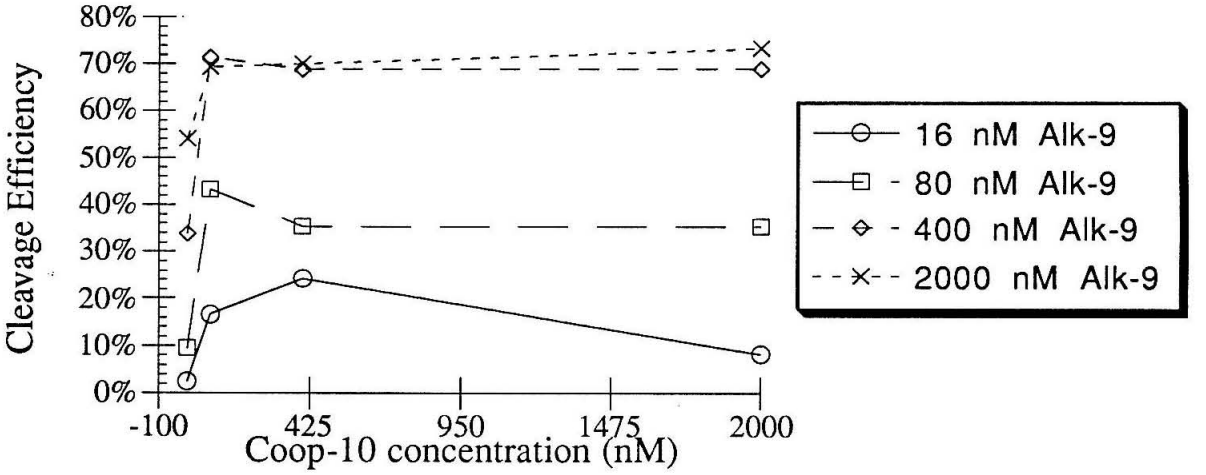


Fig. 4.37. Yield of alkylation observed upon incubation of *N*-bromoacetyloligonucleotide Alk-9 with pUCLEU2C DNA as a function of the concentration of oligonucleotide Coop-10, which binds an abutting site.

the higher affinity of Coop-10 for this site. The yield of alkylation by *N*-bromoacetyloligonucleotide Alk-9 increases by a factor of 10 upon addition of 400 nM Coop-10. At 16 nM Alk-9, the presence of 400 nM Coop-10 increases the binding affinity of Alk-9 by a factor of 10.

The effect of Coop-10 concentration on the binding affinity of Alk-9 is depicted in Figure 4.37. Addition of small amounts of Coop-10 results in an increase in binding affinity. Higher concentrations of Coop-10 do not increase cleavage efficiency, and, at high Coop-10:Alk-9 ratios, eventually lead to decreases in binding of Alk-9, due to competition between the two oligonucleotides for the Alk-9 binding site.

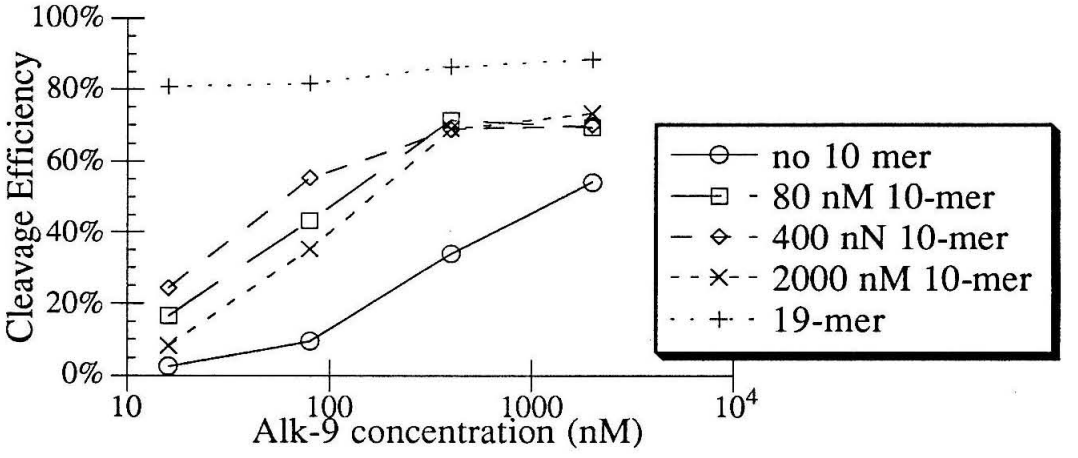


Fig. 4.38. Yield of alkylation observed upon incubation of *N*-bromoacetyloligonucleotide Alk-9 with pUCLEU2C DNA as a function of the concentration of oligonucleotide Coop-10, which binds an abutting site. Approximate dissociation constants for the oligonucleotide Alk-9 at each concentration of Coop-10 can be inferred from this plot.

Cleavage efficiencies as a function of Alk-9 concentration are plotted in Figure 4.38. Approximate binding affinities for Alk-9 at each Coop-10 concentration can be determined. The dissociation constant for the 19 base oligonucleotide Alk-19 is listed for comparison.

Coop-10 (nM)	~ Dissociation Constant of Alk-9
0	400 nM
80	80 nM
400	40 nM
2000	100 nM
Oligo Alk-19	< 10 nM

The dissociation constant for Alk-9 in the absence of Coop-10 is approximately 400 nM. In the presence of 400 nM Coop-10, the binding affinity increases by about a factor of 10 (dissociation constant decreases to 40 nM). Higher Coop-10 concentrations do not result in increased binding affinities. In no case do these cooperatively binding systems approach the binding affinity of the 19 base oligonucleotide Alk-19, which has a dissociation constant of less than 10 nM.

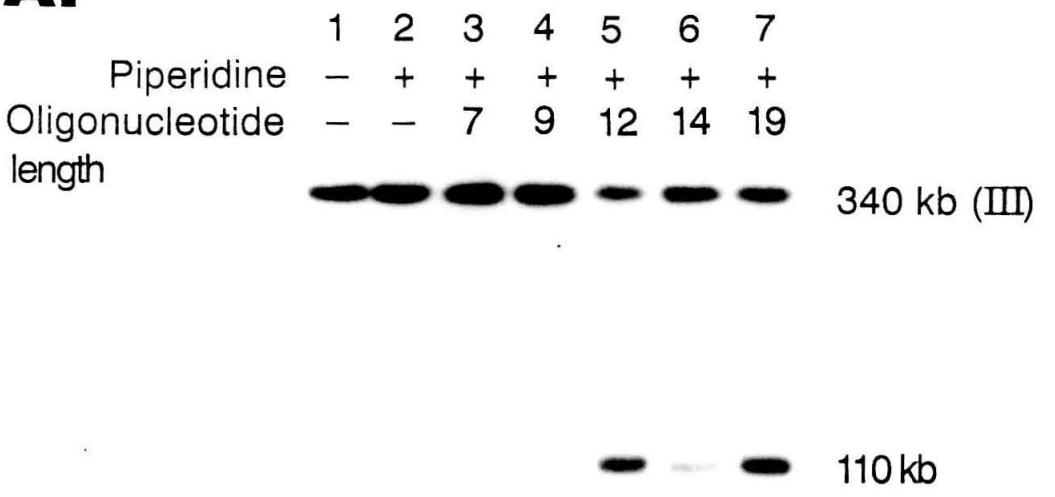
Effect of oligonucleotide length and cooperative interactions on cleavage of large chromosomal DNA. The effect of oligonucleotide length and the ability of oligonucleotides to cooperatively bind to large DNA was studied on yeast chromosomal DNA. Cleavage of yeast chromosome III by oligonucleotides Alk-7 through Alk-19 was accomplished (Figure 4.39). Oligonucleotides > 12 bases in length display high cleavage yields under these experimental conditions (micromolar in *N*-bromoacetyloligonucleotide). Rapid decreases in cleavage efficiency are observed with oligonucleotides less

Figure 4.39. Cleavage of yeast chromosomal DNA by *N*-bromoacetyloligonucleotides Alk-7 through Alk-19. Autoradiogram of a DNA blotting experiment after pulse field gel electrophoresis.

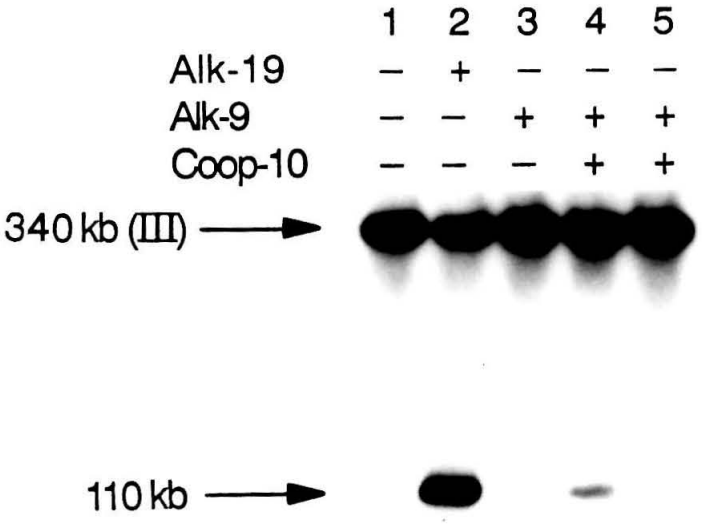
A. Alkylation efficiency as a function of oligonucleotide length. Agarose plugs containing DNA from the yeast strain SEY6210C were equilibrated in 1.0 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.0 and 2 μM *N*-bromoacetyloligonucleotide **5**. After 60 hours, the plugs were washed in 0.1% piperidine, 100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8.0 (3 washes 10 min. each) and incubated for 12 hours at 55 °C. The plugs were then equilibrated with 0.5x TBE buffer, and loaded into a 1.0% agarose gel for pulse field gel electrophoresis. Separation was accomplished by electrophoresis at 200 V for 24 hours. Pulse times were ramped from 10 - 60 seconds for the first 18 hours, followed by 60 -90 sec. over the last 6 hours. Lane 1: yeast chromosomal DNA incubated in the absence of oligonucleotide and not treated with piperidine. Lane 2: yeast chromosomal DNA incubated in the absence of oligonucleotide and treated with piperidine. Lane 3: yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide Alk-7. Lane 4: yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide Alk-9. Lane 5: yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide Alk-12. Lane 6: yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide Alk-14. Lane 7: yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide Alk-19.

B. Alkylation efficiency by oligonucleotides binding abutting target sequences within yeast chromosomal DNA. This experiment was conducted at a pH of 6.8, and all lanes were treated with piperidine. Lane 1: yeast chromosomal DNA incubated in the absence of oligonucleotide. Lane 2: yeast chromosomal DNA incubated with *N*-bromoacetyloligonucleotide Alk-19. Lane 3: yeast chromosomal DNA incubated with 2 μM *N*-bromoacetyloligonucleotide Alk-9. Lane 4: yeast chromosomal DNA incubated with 2 μM *N*-bromoacetyloligonucleotide Alk-9 and oligonucleotide Coop-10 (2 μM). Lane 5: yeast chromosomal DNA incubated with 2 μM *N*-bromoacetyloligonucleotide Alk-9 and oligonucleotide Coop-10 (10 μM).

A.


1 2 3 4 5 6 7
Piperidine – + + + + + +
Oligonucleotide length – – 7 9 12 14 19
340 kb (III)
110 kb

B.


1 2 3 4 5
Alk-19 – + – – –
Alk-9 – – + + +
Coop-10 – – – + +
340 kb (III)
110 kb

than 12 bases in length, in agreement with data obtained on plasmid DNA in solution.

Cooperative binding to chromosomal DNA was investigated using two different concentrations of Coop-10 at 2 μM *N*-bromoacetyloligonucleotide Alk-9. In the presence of 2 μM Coop-10, a two-fold increase in cleavage efficiency is observed. Increasing the Coop-10 concentration results in a decrease in cleavage efficiency, presumably due to competition of the two oligonucleotides for the same binding site. This indicates that two oligonucleotides can bind in a cooperative manner to a site found among almost 14×10^6 base pairs of DNA.

Conclusions

Efficient DNA alkylation can be accomplished using *N*-bromoacetyloligonucleotides ≥ 12 bases in length. Approximate dissociation constants for oligonucleotides 7, 9, 12, 14 and 19 bases in length have been determined. Degenerate *N*-bromoacetyloligonucleotides modify DNA with 15-20% efficiencies on each strand, thus double strand modification efficiencies are low. This makes the identification of triplex/alkylation sites using degenerate *N*-bromoacetyloligonucleotides difficult.

Oligonucleotides binding to abutting sites on double helical DNA can interact in a cooperative manner which results in a ten-fold increase in binding affinity. A two-fold increase in binding affinity was observed for these oligonucleotides binding to yeast chromosomal DNA. This indicates that the cooperative binding of short oligonucleotides can be used to bind to sites on large DNA.

Part V

Preliminary Studies on Use of *N*-Bromoacetyloligonucleotides to Effect Downstream Inhibition of Transcription

Inhibition of Transcription by *N*-bromoacetyloligonucleotides downstream from the site of transcriptional initiation.

Triple helix formation offers a method of targeting single sites within genomic size DNA, and the therapeutic potential of this motif has spawned great interest. Maher, Wold and Dervan have studied the effect of triple helix formation near the site of transcriptional initiation; however, preliminary data indicate that oligonucleotide-directed triple helices within the actual gene itself are not sufficiently stable to inhibit elongation of transcripts.[106] Cross-linking of the oligonucleotide to the duplex, or site specific modification of the duplex by *N*-bromoacetyloligonucleotides, may enable the use of *N*-bromoacetyloligonucleotides as inhibitors of DNA transcription downstream from promoter sequences.

Plasmid pTPTR was constructed to determine the effect of oligonucleotide-directed triple helices on the transcriptional efficiencies of the bacterial RNA polymerases T3 and T7. Plasmid pTPTR was constructed by Jim Maher via ligation of oligonucleotides containing binding and alkylation sites (properly positioned guanine bases) into the plasmid pLJM1. The plasmid contains T3 and T7 promoter sites on opposite strands; therefore, it is possible to study the effect of triple helix formation and subsequent duplex modification both on the transcribed and nontranscribed strands.

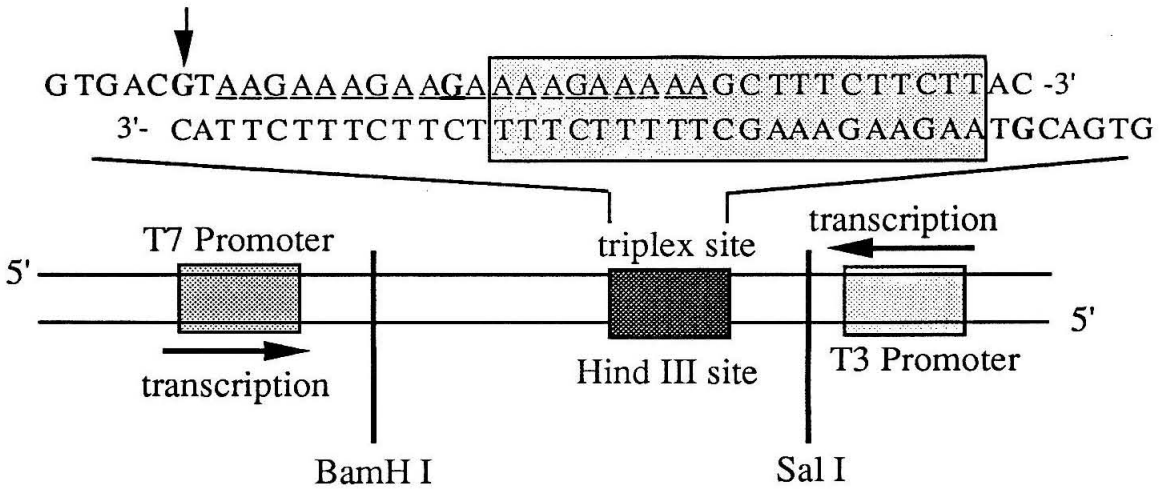


Fig. 4.40 Coarse resolution map of transcribed region of plasmid pTPTR. Plasmid pTPTR was constructed to allow the study of the effect of *N*-bromoacetyloligonucleotide-directed modification of double helical DNA on transcriptional efficiency by the bacterial polymerases T7 and T3. The polymerases transcribe in the direction and along the strand indicated. Use of both polymerases allows determination of the effect of strand modification on either the transcribed or non-transcribed strand. To study transcription by T7, DNA cut with Sal I is used. Studies with T3 polymerase are conducted using DNA linearized with BamH I. In this way, full length transcripts from each promoter are approximately 400 bases in length. Inhibition at the triplex site should result in the production of prematurely terminated transcripts ~ 300 (T7 transcript) and ~ 100 (T3 transcript) in length. Cleavage of the DNA using Hind III results in a transcript whose length is approximately equal to that of transcripts terminated at the triplex site.

Transcription by the polymerases T3 and T7 was studied on fragments obtained by cleavage of the plasmid pTPTR by BamH I and Sal I respectively. Full length transcripts from each promoter are approximately 400 bases in length. The triplex site is located approximately 100 base pairs from the T3 promoter and 300 base pairs from the T7 promoter. Transcription of a DNA template previously cut using Hind III results in the production of a fragment of the size expected if termination of transcription occurs at the triplex site. If transcriptional termination occurs at the alkylated guanines, the T3 and T7 transcripts will be slightly longer than Hind III marker transcript, since the

modified guanines on each strand are slightly downstream of the Hind II site for each promoter.

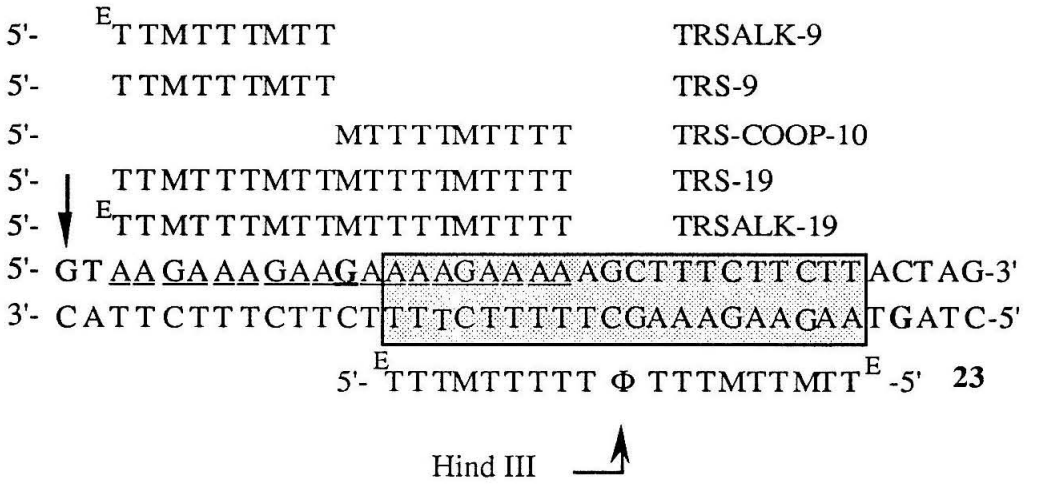
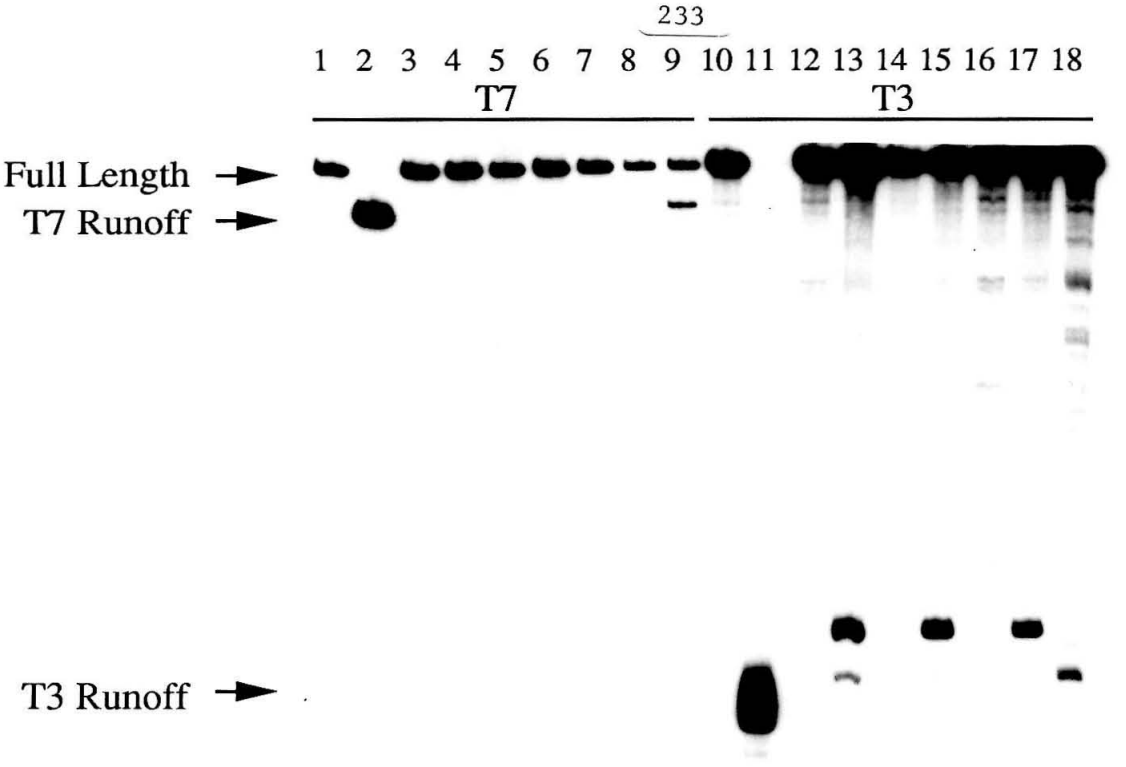


Fig. 4.41. Nucleotide resolution map of the triplex site within the transcribed region of pTPTR. The target sequence constructed within plasmid pTPTR and the oligonucleotides synthesized for studies of transcriptional inhibition by unmodified and *N*-bromoacetyl-oligonucleotides downstream from promoter sites are indicated. Oligonucleotides are aligned above the bases they are designed to bind, with the exception of cross-over oligonucleotide **23**. The *N*-bromoacetyloligonucleotide **23** binds and alkylates both strands of the duplex by crossing over in the major groove as described by Horne and Dervan.[21] Guanine bases targeted for alkylation by oligonucleotide **23** are in bold. The guanine base targeted for alkylation by oligonucleotides TRSALK-9 and TRSALK-19 is indicated by the arrow. These oligonucleotides bind and alkylate only the strand transcribed by the T3 polymerase. The central location of the Hind III site (5'-AAGCTT-3') is indicated.

The plasmid pTPTR was digested with enzymes to produce DNA fragments truncated approximately 400 base pairs downstream from each promoter site. The DNA was incubated with both *N*-bromoacetyl and unmodified oligonucleotides. Transcription of the DNA results in the

Fig. 4.42. Transcriptional inhibition downstream of promoter sites by oligonucleotides and N-bromoacetyloligonucleotides. Double stranded template from plasmid pTPTR was cut with Sal I for studies with polymerase T7; with BamH I for studies with polymerase T3. Linearized DNA was incubated with 2 μM oligonucleotide at 37 °C in 1 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.2. After 24 hours, the reactions were precipitated and washed, than dissolved in 1x transcription buffer and polymerase and ATP, CTP, GTP, and α-^{32}P-UTP added. After 1 hour at 37 °C, formamide was added, and the reactions loaded directly onto a 6% polyacrylamide denaturing gel. Lanes 1-9: DNA template linearized with Sal I. Lanes 10-18: DNA template linearized with BamH I. Lane 1 and 10: transcription of linearized DNA. Lanes 2 and 11: transcription of linearized DNA cut at the triplex site with Hind III. Lanes 3 and 12: Incubation with oligonucleotide TRS9. Lanes 4 and 13: Incubation with *N*-bromoacetyloligonucleotide TRSALK-9. Lanes 5 and 14: Incubation with oligonucleotide TRS19. Lanes 6 and 15: Incubation with *N*-bromoacetyloligonucleotide TRSALK-19. Lanes 7 and 16: Incubation with oligonucleotide TRS9 and Coop-10. Lanes 8 and 17: Incubation with *N*-bromoacetyloligonucleotide TRSALK-9 and Coop-10. Lanes 9 and 18: Incubation with *N*-bromoacetyl-cross-over oligonucleotide **23**.


1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
T7
T3
Full Length
T7 Runoff
T3 Runoff

production of 400 base full length transcripts. Inhibition of transcription at the triplex site results in the production of shorter truncated transcripts.

DNA was incubated with a series of unmodified and *N*-bromoacetyloligonucleotides under conditions where efficient alkylation occurs (1.0 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.2). After incubation at 37 °C for 24 hours, the DNA was precipitated and dissolved in transcription buffer (provided by the manufacturer) containing spermidine (2 mM), buffer (Tris, pH 7.8), and Mg^{+2} (6 mM). Transcription by both polymerases with each type of oligonucleotide was studied. Fig. 4.42 indicates that inhibition of transcription occurs only in cases where the DNA was incubated with *N*-bromoacetyloligonucleotides. Unmodified oligonucleotides have no effect on the rate of transcription. Inhibition of transcription occurs only on the strand modified by the bromoacetyl group. For instance, T7 transcription is inhibited only by the cross-over oligonucleotide **23**, since it is the only oligonucleotide which targets for modification a guanine within the T7 transcribed strand. Cross-over oligonucleotides inhibit transcription on both strands, since both strands are modified by this type of *N*-bromoacetyloligonucleotide.

Termination of transcription at modified sites occurs close to the site of modification. Transcription of Hind III cleaved DNA results in a fragment slightly shorter than the transcript from DNA modified by *N*-bromoacetyloligonucleotide **23**. This oligonucleotide modifies the DNA 10 bases downstream from the Hind III cleavage site (Fig. 4.41). The *N*-bromoacetyloligonucleotide TRSALK-19 modifies the DNA an additional 11 bases downstream. Although no sequencing analysis was conducted, comparison of the lengths of transcripts obtained from DNA templates alkylated at these two positions indicates that transcriptional termination

occurs approximately 10 and 20 bases from the Hind III site. This places the site of termination close to the site of DNA modification. The exact site of termination might be determined using a template extension assay.

In the experiment described above, the triple helix may not be stable under the transcription conditions. To determine if triplexes formed during the transcription reaction inhibit transcription, oligonucleotides were incubated with target DNA in buffer compatible with triplex formation and transcription (2 mM spermine, 6 mM Mg^{+2}, 20 mM Hepes pH 7.1). Transcription was initiated by direct addition of triphosphates and polymerase. Under these conditions, premature termination was observed in cases where the transcribed strand contained a guanine targeted for modification by a *N*-bromoacetyloligonucleotide (data not shown). No inhibition of transcription was observed upon incubation of the template DNA with unmodified oligonucleotides. These data indicate that transcriptional inhibition does not occur unless oligonucleotide-directed triple helix formation results in modification of the duplex on the strand being transcribed.

The effect of buffer composition and the type of lesion at the triple helical complex was studied. In this experiment, the DNA was incubated with *N*-bromoacetyloligonucleotide Alk-19 under several different conditions. After incubation for 24 hours at 37 °C, the DNA was treated in a manner which attempted to produce predominantly a single type of lesion (see subsequent discussion). The efficiency of transcription by T3, which transcribes the strand modified by TRSALK-19, was determined. The results are summarized in Fig. 4.43.

Effect of buffer composition. Duplex DNA was incubated with the *N*-bromoacetyloligonucleotide TRSALK-19 for 24 hours in a solution containing

one of the following salt compositions: no salts; Hepes (25 mM, pH 7.1) alone; Hepes with magnesium (6 mM); Hepes with spermine (2 mM); Hepes, spermine and magnesium; or Hepes and cobalt hexamine (1 mM). After 24 hours, the reactions were precipitated and the DNA resuspended in transcription buffer. Transcription was than effected by addition of triphosphates and T3 polymerase.

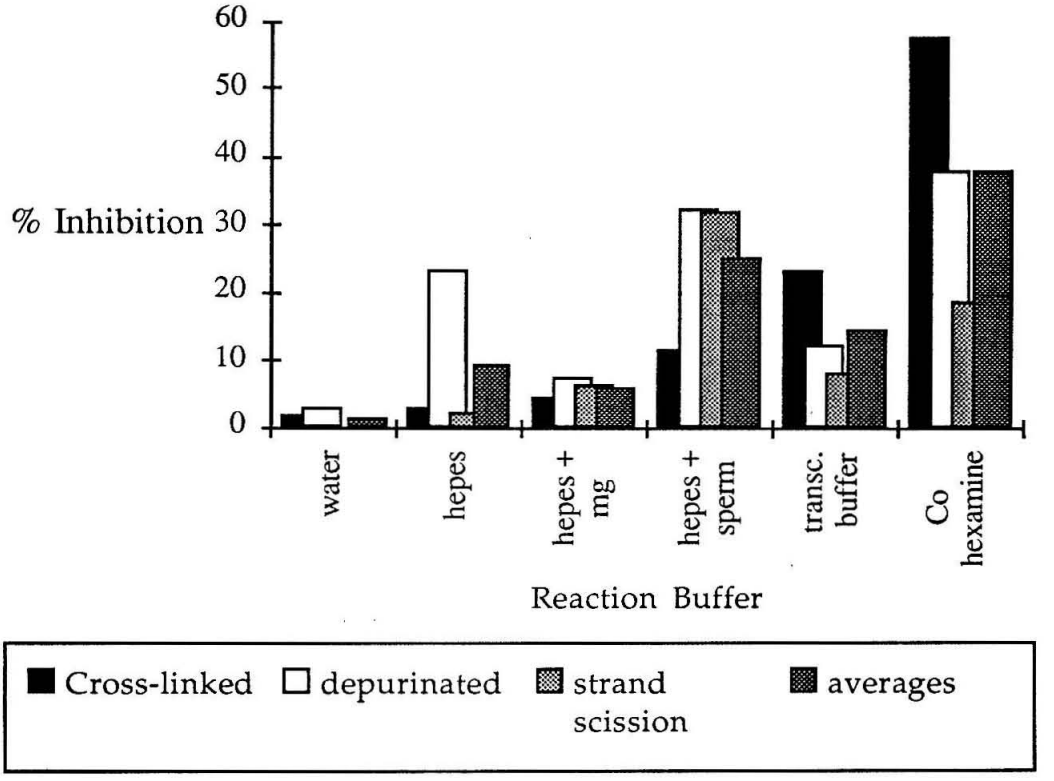


Fig. 4.43. Inhibition of the polymerase T3 at the site of triple helix formation as a function of the type of legion and the reaction buffer.

Fig. 4.43 indicates that limited transcriptional inhibition is observed in buffers containing no salt or little salt. This is presumably due to low modification yields due to the instability of the triplex. Addition of magnesium does not improve the efficiency of inhibition, despite the known

stabilizing effect magnesium ions have on triplexes (see Chapter II).[105,107] Addition of spermine increases the amount of premature transcript termination. Incubation of the DNA with the *N*-bromoacetyloligonucleotide in a buffer containing Hepes, spermidine, and magnesium (transcription buffer) results in lower transcriptional inhibition. This suggests that magnesium ions inhibit modification of DNA by *N*-bromoacetyloligonucleotides. Incubation of the DNA and *N*-bromoacetyloligonucleotides in $Co(NH_3)_6^{+3}$ and Hepes, followed by transcription in transcription buffer, results in the highest yields of prematurely terminated transcripts. It was previously reported that lower yields of DNA modification by *N*-bromoacetyloligonucleotides are observed in buffers containing spermine. The increased yields of inhibition observed upon incubation of *N*-bromoacetyloligonucleotides with DNA in optimal alkylation buffers is most probably due to the increased yield of DNA modification in the absence of spermine.

Effect of type of lesion. Binding of *N*-bromoacetyloligonucleotides to DNA can result in the formation of any of several types of complexes: (1) a noncovalent triple helical complex; (2) a complex containing a covalently attached oligonucleotide to the N-7 position of guanine; (3) a complex containing a depurinated guanine in which the oligonucleotide remains bound to the DNA with the depurinated guanine stacked and hydrogen bonded in the DNA duplex. To determine the ability of each of the complexes listed above to inhibit transcription, the DNA was treated in one of the following ways: (1) incubation of the *N*-bromoacetyloligonucleotide with the DNA in one of the solutions previously described; it is suggested that this results in the production of predominantly covalent oligonucleotide-double helical complexes; (2) incubation of the *N*-bromoacetyloligonucleotide with

the DNA, followed by heating to effect depurination. This produces predominantly the depurinated complex containing a guanine intercalated and hydrogen bonded in the duplex DNA; (3) incubation of the DNA with *N*-bromoacetyloligonucleotide, followed by piperidine treatment to effect strand scission.

It is expected that complete termination of transcription should occur at sites of strand scission. Strand scission effected by piperidine treated does not increase the extent of inhibition of transcription. This indicates that the covalent and depurinated complexes are as efficient at inhibiting transcription as a break in the phosphodiester backbone, and suggests that *in vitro* the bacterial polymerase T3 cannot read through these modifications, or that depurination and/or strand scission occurs spontaneously under the conditions of the reaction (37 °C, 24 hours).

It should be noted that a variety of repair mechanisms exist in cells which rapidly identify and repair methylated bases and depurinated sites. It is unknown how quickly or with what efficiencies these enzymes could repair the lesions produced by *N*-bromoacetyloligonucleotides. These enzymes may be inhibited from the repair of lesions associated with *N*-bromoacetyloligonucleotides by the presence of a triplex at the site of modification.

Conclusions

Inhibition of transcription downstream from promoter sites via triple helix formation occurs only if modification of the transcribed strand occurs. The site of termination of the transcript is near the site of modification. Such modification is inhibited by the presence of Mg^{+2} and spermine. The efficiency of transcriptional termination is not increased by effecting strand

scission at the site of modification, indicating that DNA lesions caused by reaction of *N*-bromoacetyloligonucleotides with the DNA target are sufficient to result in 100% inhibition of transcription *in vitro*.

Part VI

Initial Studies for the Determination of the Termini and End Products Obtained upon Alkylation of Double Helical DNA by *N*-Bromoacetyloligonucleotides

Presented here are preliminary data concerning analysis of the nature of the mechanism of cleavage effected by *N*-bromoacetyloligonucleotides. This data represents part of a continuing project within the group to fully characterize the site of reaction of *N*-bromoacetyloligonucleotides with double helical DNA. As such, the data is not complete, and this part of the thesis merely represents a summary of work accomplished in this area.

Analysis of DNA termini produced by DNA alkylation/cleavage mechanism

The proposed mechanism of cleavage predicts the production of 3' and 5' phosphate ends identical with those observed with Maxam-Gilbert sequencing reactions (Fig. 4.2).[62,63] *N*-Bromoacetyloligonucleotide **5** was incubated with 3' and 5' ^{32}P labeled DNA fragments containing the triplex target site approximately 40 bp removed from the ^{32}P label. The cleavage products were analyzed on a 15% polyacrylamide high resolution gel, and migrate with those obtained in a standard Maxam-Gilbert G reaction, known to be 3' and 5' phosphate (Fig. 4.44).[62,63] Treatment of 3'-end-labeled DNA with calf alkaline phosphatase removes terminal 5' phosphates.[77] Treatment of *N*-bromoacetyloligonucleotide cleaved DNA with CAP results in a fragment with slightly slower mobility in a polyacrylamide gel (Fig. 4.44A, lane 3). Subsequent treatment of this fragment with kinase and ATP results in addition of a phosphate to the 5' hydroxyl.[77] The product thus obtained

again co-migrates with the fragment obtained directly from the alkylation reaction before removal of the 5' phosphate (Fig. 4.44A, lane 4).

5'-End-labeled DNA was cleaved using *N*-bromoacetyloligonucleotides. The product migrated with that produced by methylation/strand scission using dimethyl sulfate/piperidine as prescribed by Maxam and Gilbert,[62,63] a procedure which produces 3' phosphate ends (Fig. 4.44B, lane 1 and 2). This DNA was subsequently treated with kinase under conditions known to result in removal of 3' phosphates[108] A decrease in the migration rate of the band is observed (Fig. 4.44B, lane 3). These data indicate that alkylation of DNA by *N*-bromoacetyloligonucleotides followed by piperidine treatment results in the production of 5' and 3' phosphate end products.

Fig. 4.44. Enzymatic analysis of the ends produced upon cleavage of DNA using the *N*-bromoacetyloligonucleotide alkylation/piperidine cleavage mechanism.

A. Autoradiogram of a high resolution 15% denaturing polyacrylamide gel of cleavage products from the reaction of oligonucleotide **5** with 3' ^{32}P end-labeled fragment (Pst I/Eco0109 I) from plasmid pUCALK. Reaction conditions were 1 μM oligonucleotide **5**, 20 mM Hepes, pH 7.4, 0.8 mM $Co(NH_3)^{+3}$, and 10,000 cpm end-labeled DNA in a total volume of 15 μl. Reactions were incubated at 37 °C for 36 hours, precipitated with NaOAc/EtOH, washed with 70% EtOH, and treated with 1.0% aqueous piperidine (90 °C, 30 min). After lyophilization, the reactions were treated as follows. Lane 2: no additional treatment; Lane 3: treated with calf alkaline phosphatase under conditions recommended by the manufacturer to dephosphorylate 5' phosphates;[77] Lane 4: treated as DNA in lane 3, followed by precipitation to purify the DNA, and treatment with kinase and ATP under conditions which phosphorylate 5' hydroxyl groups.[77] Lane 1 is a standard DMS Maxam-Gilbert G reaction.[62,63]

B. Autoradiogram of a high resolution 15% denaturing polyacrylamide gel of cleavage products from the reaction of oligonucleotide **5** with 5' ^{32}P end-labeled fragment (EcoR I/AlWN I) from plasmid pUCALK. Reactions were conducted as in (A), and were subsequently treated to produce a single type of lesion. Lane 2: no additional treatment; Lane 3: treated with kinase as described in Hertzberg, resulting in dephosphorylation of 3' phosphate groups.[108] Lane 1 is a standard DMS Maxam-Gilbert G reaction.[62,63]

A.

	1	2	3	4
DMS/piperidine	+	-	-	-
N-bromoacetyl alkylation/cleavage	-	+	+	+
CAP	-	-	+	+
Kinase	-	-	-	+

B.

	1	2	3
DMS/piperidine	+	-	-
N-bromoacetyl alkylation/cleavage	-	+	+
Kinase	-	-	+

Preliminary attempts to identify the product produced during the alkylation/cleavage reaction

The mechanism indicated in Fig. 4.2 indicates that upon depurination, an guanine base remains covalently attached to the oligonucleotide. Digestion of the oligonucleotides with phosphodiesterase and dephosphorylation with calf alkaline phosphatase should allow isolation of a thymidine-linker arm-N7-guanine adduct.

```
PA-3:     5'-GCTAGTAAAAGAAAAGGAAAGAAAAGTCG -3'
PA-4:     3'-CGATCATTTTCTTTTCCTTTCTTTTCAGC -5'
```

Fig. 4.45. Sequence of oligonucleotides used to identify product from alkylation/depurination of double helical DNA by *N*-bromoacetyloligonucleotides.

N-Bromoacetyloligonucleotide **5** was incubated with oligonucleotides PA-3 and PA-4 in 1 mM $Co(NH_3)_6^{+3}$, 10 mM pH 7.0 Hepes, and 30% ethanol. The three oligonucleotides were heated to 55 °C and allowed to cool to room temperature, and than to 4 °C for several hours, after which the reaction was transferred to 37 °C. The DNA was precipitated with ethanol, and incubated with 0.3 M NaOAc (pH 5.2) for 20 min. at 90 °C. The DNA was precipitated a second time, washed, and incubated in phosphodiesterase buffer with 3 units of calf alkaline phosphatase and phosphodiesterase. After 45 min at 37 °C, the reaction was analyzed using reverse phase HPLC. Analysis was conducted using a reverse phase C-18 column (Brownlee Aquapore OD-300 C-18 column (200 mm x 4.6 mm) on a Hewlett-Packard HP 1090 with the following solvents and gradient:

Solvent A: 10 mM ammonium acetate, pH 5.2
Solvent B: 10 mM ammonium acetate, pH 5.2, 80% methanol.

Gradient:

time	% B
0-10	10%
10-20	10-20%
20-30	20-30%

Flow Rate: 0.5 ml/minute

Retention times under these conditions are as follows:

amino-thymidine	7.4 min.
N-bromoacetylthymidine	28.7 min.
product from alkylation rxn.	25.4 min.

A new product with a retention time of 25.4 minutes was observed.

Independent synthesis and purification of a product with an equal retention time was accomplished (Fig. 4.46). The known nucleoside **24**[64] was incubated with ethylene diamine. Following removal of the solvent and drying of the compound, the aminonucleoside was derivatized using *N*-hydroxysuccinimidyl bromoacetate in water. Reaction with 5'-O-DMT-2'-deoxyguanosine in DMF, followed by purification of the product using prep reverse phase HPLC afforded compound **27**. NMR is consistent with an adduct containing both guanine base and a thymidine with the linker arm. The UV of the product shows maxima at both 252 and 273 nm, with absorption tailing to > 310 nm.

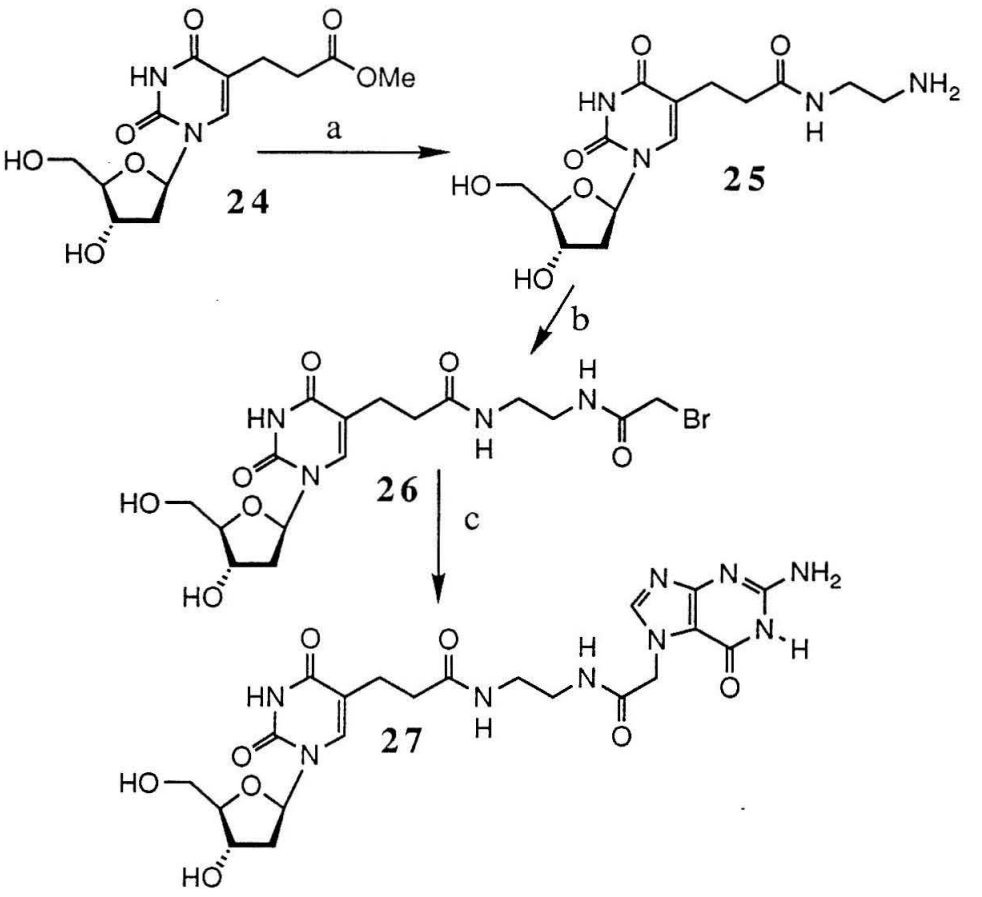


Fig. 4.46. Scheme for the synthesis of expected product from alkylation-depurination reaction. a) ethylene diamine, rt, 2 days; b) *N*-hydroxysuccinimidyl bromoacetate, water; c) 5'-O-DMT-2'-deoxyguanosine, DMF, 2 days, 40 °C.

Part VII

Materials and Methods

^{1}H nuclear magnetic resonance (NMR) spectra were recorded on a Jeol JNM-GX400 FT 400 MHz spectrometer and are reported in parts per million (ppm) from tetramethylsilane. ^{31}P NMR spectra were recorded on a Varian Associates EM-390 and are reported in parts per million from phosphoric acid. Ultraviolet-visible (UV-Vis) spectra were recorded on a Perkin-Elmer UV spectrophotometer. Mass spectra were obtained from the regional mass spectroscopy facility at the University of California, Riverside. Flash chromatography was carried out under positive air pressure using EM Science Kieselgel 60 (230-400 Mesh). Elemental analyses were performed by the analytical laboratory, California Institute of Technology.

High pressure liquid chromatography (HPLC) was accomplished using a Hewlett-Packard 1090 liquid chromatograph with a diode array detector and a Hewlett-Packard 79994A analytical work station. Analytical reverse phase and oligonucleotide preparative HPLC was performed on a Brownlee Aquapore OD-300 C-18 column (200 mm x 4.6 mm). Analytical and oligonucleotide preparative anion exchange HPLC was performed on a Vydac oligonucleotide anion exchange 4.6 x 200 mm column.

Collidine, fluorenylmethyl chloroformate, p-anisylchlorodiphenylmethane, and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite were purchased from Aldrich and used without further purification. Bromoacetic acid, iodoacetic acid, chloroacetic acid, bromoacetamide, and p-nitrophenol were purchased from Aldrich and used as received. *N*-hydroxysuccinimide, succinic anhydride, and

triethylamine were used as received from Fluka. CPG-amine solid support was purchased from Sigma, and was washed with DMF, methanol, and methylene chloride prior to use. Diisopropylethylamine (Aldrich) was distilled from CaH_2 prior to use. Ethylene diamine, 1,3-diaminopropane, and 1,4-diaminobutane (all from Aldrich) were distilled just prior to use. Dry solvents (DMF, methylene chloride, dioxane) were obtained from Fluka over molecular sieves when possible. Methylene chloride used for the synthesis of phosphoramidites was dried by passing over a column of basic alumina immediately before use.

Restriction enzymes were obtained from New England BioLabs or Boehringer Mannheim, and used with the manufacturer's supplied buffers. Polynucleotide kinase and the Klenow fragment were obtained from Boehringer Mannheim. Terminal transferase was obtained from BRL with 5x reaction buffer. T3 and T7 RNA polymerases and RNase inhibitor were obtained from Stratagene. Radioactive nucleotides (^{32}P-α-dATP, ^{32}P-γ-ATP, ^{32}P-α-UTP, ^{32}P-α-ddATP, ^{32}P-α-dCTP, ^{32}P-α-TTP,and ^{32}P-α-dGTP), were purchased from Amersham. Prep-a-Gene matrix was purchased from Bio-Rad. Large scale plasmid purification was accomplished using Maxi-plasmid purification kits from Quiagen as suggested by the manufacturer. pUCLEU2 was constructed by Scott Strobel for the subsequent construction of yeast strains containing desired sequences.

Cobalt (III) hexamine trichloride was obtained from Kodak. Phosphoramidites for oligonucleotide synthesis were obtained from Cruachem. Plasmid pUC19 used for the construction of pUCALK was obtained from Bio-rad. XL-1 competent cells were obtained from Stratagene. Dideoxy sequencing was accomplished using kits obtained from USB (Sequenase 2.0). Sep-Pak cartridges were from Waters. Sephadex was obtained

from Sigma and was equilibrated with water. Water was purified using a Milli-Q system, and filtered through a 0.45 μM filter.

Synthesis of Nucleosides

Nucleoside 2.[65,66] The known aminonucleoside **1**[64] (820 mg, 1.27 mmol lyophilized from dioxane) was dissolved dry dioxane (5 mL). Collidine (500 μL, 3.9 mmol, 3 equivalents) was added, and the solution chilled to 0 °C. Fluorenylmethyl chloroformate (670 mg, 2.6 mmol, 2 equivalents, dissolved in 5 mL dry dioxane) was added over a ten minute period, and the reaction warmed to room temperature. After 20 minutes, TLC (5% MeOH in $MeCl_2$) indicated formation of a single product. The solution was frozen and lyophilized. The product was purified by flash chromatography in 2% MeOH in $MeCl_2$ to afford 1.02 gm of pure product (92% yield). ^1H NMR ($CDCl_3$): δ 7.72 (m, 2H), 7.55 (m, 3H). 7.13-7.39 (m, 14H), 6.80 (q, 4H), 6.35 (t, J = 7 Hz, 1H, $H_{1'}$), 6.18 (bs, 1H), 5.52 (bs, 1H), 4.48 (m, 1H, $H_{3'}$), 4.37 (d, 2H-Fmoc), 4.15 (t, 1H-Fmoc), 4.04 (m, 1H, $H_{4'}$), 3.74 (s, 6H, OCH_3), 3.38 (m, 2H, $H_{5'}$), 3.06-3.20 (m, 4H, NCH_2), 2.37 (m, 2H, $H_{2'}$) 2.18 (m, 4H). IR (KBr) cm^{-1}: 3367, 3065, 2933, 2836, 2370, 1701, 1685, 1676, 1654, 1648, 1508, 1448, 1298, 1251, 1177, 1152, 1091, 1032, 909, 829, 759, 714, 728. MS: (positive ion FAB) m/z 867 (0.7), 319 (2.2), 304 (23.6), 303 (100), 179 (12.5), 178 (11), 155 (11), 152 (4), 138 (13), 137 (24). HRMS calculated for $C_{50}H_{51}O_{10}N_4$ ($M+H^+$): 867.3607. Found 867.3580.

Phosphoramidite 3. Nucleoside **2** (150 mg, 176 μmol) was dissolved in dry methylene chloride (3 mL). Diisopropylethylamine (350 μL, 2 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (100 μL, 448 μmol) were added, and the reaction stirred for 3 hours. The reaction was quenched with the addition of ethanol (0.5 ml) and diluted with ethyl acetate. The

organic layer was washed with an aqueous solution of saturated $NaHCO_3$, water, and twice with an aqueous solution of saturated NaCl. After drying over Na_2SO_4 and concentration under reduced pressure, flash chromatography (500:15:5 methylene chloride:isopropanol:triethylamine) afforded 110 mg of product (60% yield). ^1H NMR ($CDCl_3$): δ 8.40 (bs, 1H), 7.72 (t, J= 7.3 Hz, 2H), 7.55 (m, 3H). 7.13-7.39 (m, 13H), 6.79 (m, 4H), 6.33 (m, 1H, $H_{1'}$), 5.49 (m, 1H), 4.59 (m, 1H, $H_{3'}$), 4.39 (d, 2H-Fmoc, J= 6.6 Hz), 4.15 (m, 1H), 4.09 (m, 1H), 3.77 (s, 6H, OCH_3), 3.54 (m, 2H, $H_{5'}$), 3.52-3.65 (m, 4H), 3.12-3.24 (m, 4H), 2.57 (t, J= 6.1 Hz, 1H), 2.37 (t, J= 6.1 Hz, 1H), 2.26 (m, 2H), 1.98-2.06 (m, 4H), 1.00-1.14 (m, 12H). IR (KBr) cm^{-1}: 3385, 3065, 2966, 2930, 2373, 1715, 1706, 1701, 1685, 1648, 1508, 1458, 1449, 1251, 1180, 1154, 1081, 1032, 978.6, 830.1, 759.6, 742.5, 700.2. ^{31}P NMR ($CDCl_3$): δ 149.081, 148.812.

Nucleoside 11a. 5-(2-carbomethoxyethyl)-5'-O-DMT-2'-deoxyuridine (767 mg, 1.24 mmol) was stirred in 1,3 diaminopropane (5 mL) overnight. The solvent was removed by rotary evaporation under vacuum, and the product dissolved in dioxane and lyophilized to dryness. The white solid was dissolved in dry pyridine (12 mL). Monomethoxytrityl chloride (1.07 gms, 3.5 mmol) was added over a period of 12 hours. The reaction was diluted with water, and extracted twice with diethyl ether. After concentration, flash chromatography afforded 745 mg of product (64.5 % yield). ^1H NMR ($CDCl_3$): δ 7.46 (s, 1H), 7.14-7.42 (m, 21H), 6.77-6.84 (m, 6H), 6.32 (t, 1H), 5.67 (t, 1H), 4.47 (m, 1H), 3.98 (m, 1H), 3.75 (s, 9H), 3.36-3.41 (m, 2H), 3.18 (m, 2H), 2.19-2.24 (m, 4H), 2.10 (m, 4H), 1.52 (m, 2H). MS: (positive ion FAB) m/z 930 (1), 461 (5), 401 (5), 399 (4), 355 (5), 325 (6), 303 (67), 289 (6), 281 (13), 274 (23), 273 (100), 267 (6), 221 (14), 207 (17), 191 (7), 147 (45), 137 (10), 136 (19), 105 (7). HRMS calculated for $C_{56}H_{58}O_9N_4$ (M^+): 930.4206. Found 930.4204.

Nucleoside 11b. Nucleoside 11a (224 mg, 240 μmol) was dissolved in dry methylene chloride (5 mL). Diisopropylethylamine (400 μL, 2.3 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (100 μL, 450 μmol) were added, and the reaction stirred for 10 minutes. TLC (64:32:5 ethyl acetate:hexane:methanol) indicated that the reaction was complete. The reaction was quenched with ethanol (5 mL) and diluted with ethyl acetate (40 mL). The organic layer was washed three times with an aqueous saturated solution of $NaHCO_3$, and once with an aqueous saturated sodium chloride solution. After drying (Na_2SO_4) and concentration under reduced pressure, flash chromatography (97:2:1 methylene chloride:isopropanol:triethylamine) afforded 185 mgs (70% yield) of nucleoside **11b**. The compound was used directly in the machine synthesis of oligonucleotide **13**.

Nucleoside 12a. 5-(2-carbomethoxyethyl)-2'-deoxyuridine (417 mg, 620 μmol) was stirred in freshly distilled 1,4-diaminobutane (5 mL) for two days, with the round bottom flask held just above a heat block, set on "low" to prevent crystallization of the solvent. The diaminobutane was removed by distillation under vacuum. The product was dissolved in dioxane and lyophilized. After lyophilization, the nucleoside was dissolved in dry pyridine (6 mL), and monomethoxytrityl chloride (550 mg, 1.78 mmol) was added over a period of three hours. The reaction was diluted with water and extracted twice with diethyl ether. After concentration, flash chromatography in 4% MeOH/$MeCl_2$ yielded 273 mgs of product (52% overall yield). ^1H NMR ($CDCl_3$): δ 7.13-7.47 (m, 21H), 6.78-6.91 (m, 6H), 6.32 (t, 1H), 5.27 (m, 1H), 4.48 (m, 1H), 3.98 (m, 1H), 3.76 (d, 9H), 3.43-3.46 (m, 1H), 3.32-3.36 (m, 1H), 3.02 (m, 2H), 2.06-2.37 (m, 8H), 1.37 (m, 4H). MS: (positive ion FAB) m/z 944 (<1), 304 (16), 303 (50), 281 (5), 289 (5), 274 (24), 273 (100),207 (7), 154 (7), 147 (16), 136 (14). HRMS calculated for $C_{57}H_{59}O_9N_4$ ($M-H^+$): 944.4363. Found 943.4274.

Nucleoside 12b. Nucleoside 12a (53 mgs, 56 μmol) was dissolved in dry methylene chloride (3 mL). Diisopropylethylamine (150 μL, 0.9 mmol) and 2-cyanoethyl-*N*,*N*-diisopropylchlorophosphoramidite (24 μL, 100 μmol) were added, and the reaction stirred for 10 minutes. TLC (64:32:5 ethyl acetate:hexane:methanol) indicated that the reaction was complete. The reaction was quenched with ethanol (1 mL) and diluted with ethyl acetate (20 mL). The organic layer was washed three times with an aqueous saturated solution of $NaHCO_3$, and once with an aqueous saturated sodium chloride solution. After drying (Na_2SO_4) and concentration under reduced pressure, flash chromatography (230:2:2 methylene chloride:methanol:triethylamine) afforded 48 mgs (75% yield) of nucleoside **12b**. The compound was used directly in the machine synthesis of oligonucleotide **14**.

***N*-hydroxysuccinimidyl bromoacetate**. *N*-hydroxysuccinimidyl bromoacetate was synthesized as described previously.[68-70] *N*-hydroxysuccinimide (2.75 gm, 27.2 mmol) and bromoacetic acid (3.25 gms, 23.4 mmol) were dissolved in dioxane (30 mL). Dicyclohexylcarbodiimide (6.09 gms, 29.5 mmol) was added, and the reaction stirred for 80 minutes. After removal of the urea by filtration, petroleum ether was added and the product crystallized. ^1H NMR ($CDCl_3$): δ 4.64 (s, 1H), 2.82 (s, 2H). *N*-hydroxysuccinimidyl iodo- and chloro- acetates were synthesized analogously using iodo- and chloroacetic acids.

Synthesis of CPG-Alk-2. Synthesis of 3'-O-DMT-nucleosides linked to CPG solid support via the 5' hydroxyl have been described.[110] This general procedure was used to link the amino-thymidine derivative to CPG for the synthesis of cross-over molecules of the type described by Horne and Dervan.[21]

Nucleoside 16.[111,112] 5-(2-carbomethoxyethyl)-2'-deoxyuridine[64] (**15**) (430 mgs, 1.37 mmol) and imidazole (252 mg, 3.7 mmol) were dissolved in DMF (10 mL) and chilled to 0 °C. tert-Butyl-dimethylsilylchloride (236 mg, 1.57 mmol) was added in one portion, and the reaction stirred for 11 hours at 4 °C. The reaction was partitioned between diethyl ether and water. The aqueous layer was extracted with diethyl ether, the combined organic layers were concentrated under reduced pressure, and pure product obtained after flash chromatography in 5% MeOH/methylene chloride to afford 472 mgs (80% yield) of product. ^{1}H NMR ($CDCl_3$): δ 8.26 (s, 1H), 7.52 (s, 1H, H_6), 6.32 (t, J = 6 Hz, 1H, $H_{1'}$), 4.47 (m, 1H, $H_{3'}$), 4.00 (m, 1H), 3.86 (d, J = 3.4 Hz, 2H, $H_{5'}$), 3.66 (s, 3H), 2.60 (m, 4H), 2.36 (m, 1H), 2.14 (m, 1H), 0.91 (s, 9H), 0.11 (m, 6H).

Nucleoside 17.[113] 5-(2-carbomethoxyethyl)-5'-O-tert-butyldimethylsilyl-2'-deoxyuridine (**16**, 472 mgs, 1.10 mmol), dimethoxytrityl chloride (453 mgs, 1.34 mmol) and pyridine (350 μL, 4.33 mmol) were refluxed in methylene chloride for three hours. The reaction was diluted with methylene chloride and washed with water. After separation, the organic layer was extracted twice with methylene chloride, and the organic layers combined and concentrated under reduced pressure. Flash chromatography (1% MeOH/methylene chloride) yielded 646 mg of product (80% yield). ^{1}H NMR ($CDCl_3$): δ 8.35 (s, 1H), 7.59 (s, 3H), 7.31-7.39 (m, 7H), 6.92 (d, 4H), 6.45 (m, 1H, $H_{1'}$), 4.34 (d, 1H, $H_{3'}$), 4.11 (s, 1H, $H_{4'}$), 3.88 (s, 6H), 3.71 (s, 3H), 3.71-3.74 (m, 1H, $H_{5'}$), 3.39 (bd, 1H, $H_{5'}$), 2.62 (m, 4H), 1.66-1.81 (m, 2H, $H_{2'}$), 0.89 (s, 9H), 0.09 (s, 3H), 0.04 (s, 3H).

Nucleoside 18.[111,112] Nucleoside **17** (860 mg, 1.17 mmol) was dissolved in THF (10 mL). Tetrabutylammonium fluoride (1.0 M, 4 mL.) was added and the reaction stirred at room temperature for three hours. The reaction was diluted with methylene chloride, washed with an aqueous solution of

saturated $NaHCO_3$, an aqueous solution of saturated NaCl, dried (Na_2SO_4), and concentrated under reduced pressure. Flash chromatography in 1% MeOH/$MeCl_2$ afforded complete separation from a yellow impurity which was otherwise difficult to remove. Yield: 483 mgs, 67%. ^1H NMR ($CDCl_3$): δ 8.27 (s, 1H), 7.57 (s, 1H), 7.45 (d, 2H), 7.20-7.35 (m, 7H), 6.83 (d, 4H), 6.27 (t, J = 6 Hz, 1H, $H_{1'}$), 4.39 (d, J = 6 Hz, 1H), 3.90 (m, 1H), 3.80 (s, 6H), 3.62-3.67 (m, 1H), 3.61 (s, 3H), 3.29 (m, 1H), 2.99 (t, 1H), 2.51-2.62 (m, 4H), 1.80-1.88 (m, 2H, $H_{2'}$). HRMS calculated for $C_{34}H_{37}O_9N_2$ ($M+H^+$): 617.2500. Found 617.2462.

Nucleoside 19. Nucleoside **18** (320 mg, 494 μmol) was stirred in ethylene diamine (5 mL) at room temperature. After 36 hours, TLC (3% MeOH/$MeCl_2$) indicated complete loss of starting material. The reaction was concentrated under reduced pressure, dissolved in dioxane, and lyophilized. The resulting white powder was dissolved in dry dioxane (8 mL). Pyridine (300 μL, 3.71 mmol) was added and the reaction chilled to 0 °C. Fluorenylmethyl chloroformate (159 mg, 0.615 μmol) was added, and the reaction allowed to warm to room temperature. After 30 minutes, additional fluorenylmethyl chloroformate (150 mg) was added (580 μmol, 1.19 mmol total). The reaction was frozen and lyophilized. Flash chromatography (3% MeOH/$MeCl_2$) yielded product in high yield (100%), indicating the possible presence of an impurity. ^1H NMR ($CDCl_3$): δ 7.64-7.75 (m, 4H), 7.56 (d, 2H), 7.20-45 (m, 14H), 6.81 (t, 4H), 6.52 (bm, 1H), 6.33 (m, 1H, $H_{1'}$), 5.32 (bm, 1H), 4.39 (d, 2H, Fmoc), 4.21 (t, 1H, Fmoc), 3.85 (m, 1H), 3.76 (s, 6H), 3.56-3.64 (m, 3H), 3.24 (m, 4H), 2.52-2.61 (m, 2H), 2.36-44 (m, 2H), 1.82 (m, 2H, $H_{2'}$). MS: (positive ion FAB) m/z 867 (<1), 304 (24), 303 (100), 273 (3), 227 (5), 213 (3), 210 (5), 179 (16), 178 (13), 154 (10), 136 (9), 107 (3). HRMS calculated for $C_{50}H_{51}O_{10}N_4$ ($M+H^+$): 867.3607. Found 867.3619.

Nucleoside 20. Nucleoside **19** (368 mgs, 424 μmol), dimethylaminopyridine (13 mgs, 106 μmol), pyridine (400 μL, 5.0 mmol), and succinic anhydride (74.3 mgs, 742 μmol) were refluxed in dry methylene chloride (15 mL). After 12 hours, TLC (9:1 $CHCl_3$:MeOH) indicated complete conversion to a single product. The reaction was diluted with methylene chloride, washed twice with ice cold 10% citric acid, once with an aqueous solution of saturated NaCl, dried (Na_2SO_4), and concentrated under reduced pressure. Flash chromatography (5% MeOH/$MeCl_2$) yielded 244 mg of product (60% yield). The low yield reflects the presence of an impurity in the starting material. MS: (positive ion FAB) m/z 966 (2), 665 (4), 304 (22), 303 (100), 243 (10), 213 (9), 179 (15), 178 (14), 149 (25), 137 (34), 136 (47). HRMS calculated for $C_{54}H_{55}O_{13}N_4$ ($M+H^+$): 967.3767. Found 967.3715..

Control Pore Glass-3'-T^{amine} (21). Nucleoside **20** (220 mg, 227 μmol) and p-nitro-phenol (81.4 mg, 585 μmol) were dissolved in THF (5 mL). Dicyclohexylcarbodiimide (84 mgs, 408 μmol) was added, and the reaction stirred for six hours. Additional p-nitrophenol (30 mg 215 μmol) and dicyclohexylcarbodiimide (30 mg, 145 μmol) were added, and the reaction stirred for an additional four hours. The reaction was diluted with diethyl ether, washed with an aqueous solution of saturated $NaHCO_3$, water, an aqueous solution of saturated NaCl, and concentrated under reduced pressure. Flash chromatography in 80:20 ethyl acetate:methylene chloride afforded 222 mgs of product (90% yield), which was used directly to derivatize the CPG solid support.

Controlled pore glass was treated with 1 M triethylamine in DMF. The support was filtered over medium porosity glass filter, washed extensively with DMF, and dried under vacuum. The para-nitro ester of Alk-2 (60 mg, 55 μmol) was stirred with approximately 500 mgs of support and 5 μL of

triethylamine in DMF (10 mL). Use of excess triethylamine results in cleavage of the Fmoc protecting group. After 24 hours, the support was filtered (medium glass pore) and washed extensively with DMF, methylene chloride and methanol. Loading was determined by treatment with 60% perchloric acid and measurement by UV of the dimethoxytrityl cation released.[67] A yield of 20 μmol/gm support was achieved. Acetylation of the support was achieved using acetic anhydride according to standard procedure.[67]

Synthesis of oligonucleotides. Oligonucleotides were synthesized on an Applied Biosystems 380B automated DNA synthesizer, and deprotected with concentrated NH_4OH at 55 °C for 20 hours. After lyophilization, oligonucleotides were purified by electrophoresis through 20% denaturing polyacrylamide gels. The probe was eluted from the excised band by incubation in 200 mM NaCl, 1 mM EDTA for 24 hours at 37 °C. Oligonucleotides were dialyzed against Millipore Ultra-pure water, and their concentrations determined based on their UV absorbance (extinction coefficients were determined by summing the values for the individual nucleosides, using values of e_{260} = 8800 for T and T^{alk}, e_{260} = 5700 for Me^5C, e_{280} = 6400 for T and e_{280}= 8300 for Me^5C).

Derivatization of oligonucleotides. Ten nanomoles of the amino-oligonucleotide were dissolved in 10 μL of 200 mM borate buffer, pH 8.9. An equal volume of 250 mM *N*-hydroxysuccinimidyl bromoacetate in DMF was added. After five minutes, the solution was injected directly onto a Hewlett-Packard 1090 HPLC.

Purification was accomplished using either reverse phase or anion exchange chromatography. Reverse phase chromatography was

accomplished with a Brownlee Aquapore OD-300 C-18 column (200 mm x 4.6 mm) using a gradient of 10 - 15% acetonitrile in 100 mM ammonium acetate, pH 4.6, over 25 minutes. Retention times for oligonucleotides **4** and **5** under these conditions are 15.2 and 16.2 minutes.

Anion exchange chromatography was accomplished with a Vyadac oligonucleotide anion exchange 4.6 x 200 mm column using 20 mM Na_2HPO_4 (solvent A) and 20 mM Na_2HPO_4, 1 M NaCl (solvent B). The oligonucleotide **5** was eluted using a gradient of 40-80% B over 60 minutes. Elution times were 15.4 minutes for the underivatized oligonucleotide **4**, and 18.2 minutes for the *N*-bromoacetyloligonucleotide **5**. The product peak was collected and extracted with butanol to remove the acetonitrile. Salts were removed using a 3 ml Sephadex G-50-80 spin column, and the eluent precipitated with three volumes of NaOAc/ethanol and washed with 70% ethanol. The probe was diluted in water to determine its concentration via UV analysis, lyophilized, and stored at -20 °C until needed.

Derivatization of Alk-14, Alk-12, Alk-9 and Alk-7 oligonucleotides was accomplished as described above. The oligonucleotides were not purified by polyacrylamide gel electrophoresis, but were simply deprotected, lyophilized, derivatized, and purified using HPLC. Purification of the oligonucleotides using anion exchange HPLC was accomplished using the solvents described above, with the following gradients:

Alk-7	10 - 50 % B over 40 minutes
Alk-9	20 - 60 % B over 40 minutes
Alk-12	25 - 65 % B over 40 minutes
Alk-14	30 - 70 % B over 40 minutes.

Separation of the amino-nucleoside, the derivatized nucleoside, and the N-1 length unmodified oligonucleotide was possible for oligonucleotides of length ≤ 14 bases. Retention times under these conditions were:

	amino-oligonucleotide	*N*-bromoacetyloligonucleotide
Alk-7	16.1 minutes	21.3 minutes
Alk-9	14.4 minutes	18.2 minutes
Alk-12	16.5 minutes	18.9 minutes
Alk-14	15.5 minutes	18.5 minutes.

The dinucleosideT-T^E (8) was synthesized by coupling of phosphoramidite **3** to T-support (Cruachem) using automated synthesis. The dinucleoside was deprotected by treatment with concentrated ammonia (55 °C, 24 hrs), and lyophilized to dryness. Derivatization to the *N*-bromoacetyldinucleoside **9** was accomplished using *N*-hydroxysuccinimidyl bromoacetate. HPLC Purification was accomplished using reverse-phase chromatography in 100 mM ammonium acetate (pH 4.7), with a gradient of 0 - 8% acetonitrile from 0 to 24 minutes, followed by 8 - 20% acetonitrile from 24-50 minutes. Synthesis of the expected hydrolyzed product was accomplished by derivatization of **8** with the *N*-hydroxysuccinimidyl ester of glycolic acid. Retention times were:

8	21.5 minutes
9	28.8 minutes
10	23.7 minutes.

Construction of plasmids.

pUCALK. The plasmid pUCALK was constructed by ligation of oligonucleotides of sequence 5'-

GATCTAAGGGGAAAAGAAAAGGAAAGAAAAAGCTTTCTTCTTCCCTA-3' and the complementary oligonucleotide 5'-TCGATAGGGAAGAAGAAAGCTTTTTCTTTCCTTTTCTTTTCCCTTA-3' into pUC19[114] linearized with BamH I and Sal I. Linearized plasmid (0.5 µgm, 0.28 pmol) was mixed with ligase buffer, ATP (1 mM final concentration), and hybridized oligonucleotides in various ratios, and reacted with ligase for 24 hours at 6 °C. Competent *E. Coli* cells were transformed with ligated plasmid by mixing 15 µl of the ligation reaction with 100 µl of cells and 1.7 µl of ß-mercaptoethanol. The reactions were placed on ice for 30 minutes, heat shocked at 42 °C for 45 seconds, and chilled on ice for 2 minutes. L-B[77] media (900 µl) was added and the cells incubated at 37 °C for 1 hour with shaking. The cells were plated and selected on an ampicillin media using an α-complementation assay.[77] Colonies were picked, and mini-preps of transformants sequenced to insure that the appropriate sequence was present. Large scale plasmid preparation was accomplished by growing transformed cells overnight in 500 mL of ampicillin containing culture. A Quiagen plasmid Maxi purification kit was used according to the manufacturer's instructions to obtain purified DNA. Yields were excellent (1.5 mgs/500 ml of culture). DNA contained a small amount of material which ran as distinct bands with slower mobility in agarose gels.

pUCLEU2C. The plasmid pUCLEU2C was constructed by Scott Strobel by ligation of oligonucleotides of sequence 5'-TCGACTTTTCTTTCCTTTTCTTTTACTAGTAAAAGAAAAGGAAAGAAA AG-3' into the unique Xho I site of plasmid pUCLEU2,[31] destroying the Xho I site. Competent *E. Coli* cells were transformed with plasmid Xho I digested plasmid pUCLEU2C, and selected on an ampicillin media using an α-complementation assay. Transformants were sequenced to insure that an

appropriate sequence was present. Plasmids were purified from 500 mL of culture using a Quiagen plasmid Maxi purification kit according to the manufacturer's instructions as described above.

pUCLEU2D. Plasmid pUCLEU2D was constructed by Scott Strobel using oligonucleotides containing the sequence 5'-TCGACTAGTAAAGAAAAAGCTTTCTTCTTACTAG-3' with the procedure described above. Large scale preparation followed the procedure previously described.

pTPTR. Plasmid pTPTR was constructed by Jim Maher by ligation of oligonucleotides containing the sequence 5'-GTGACGGAAGAAAGAAGAAAAGAAAAAGCTTTCTTCTTA into the site of the plasmid pLJM1. Plasmids were purified as indicated above.

pUCSfi. The plasmid pUCSfi was constructed by Scott Strobel via insertion of an oligonucleotide containing the sequence 5'-GATCGGCCTCTAGGGCCCTGCA-3' into the polylinker region of the BamH I and Pst I linearized parent plasmid, pUC19.[31] Because the insertion does not contain a frame shift in the lacZ gene, and the two amino acid substitutions are conservative, the resulting plasmid continues to yield blue colonies in α-complementation assays and contains a unique Sfi I site. Subsequent insertion of DNA fragments into the Sfi I site can thus be screened via blue/white α-complementation assay.

Production of alternate yeast strain. Yeast strain SEY6210 *leu2-* was grown in a 200 ml liquid culture to an OD600 of 0.7, harvested by centrifugation at 5,000 rpm for five minutes at 5° C, washed in sterile H_2O, centrifuged, resuspended in 40 mL 0.1 M LiOAc, 10 mM Tris-HCl pH 8.0, 1 mM EDTA and incubated at 30 °C for 1 hour. Cells were again harvested by centrifugation, resuspended in 2 mL of LiOAc solution and 50 mL of

competent cells aliquoted into sterile microcentrifuge tubes. Five μg of PstI digested pUCLEU2C was added to the yeast solution and incubated at 30 °C for one hour. The cells were heat shocked for five minutes at 37 °C and harvested by brief (3 sec.) centrifugation. The supernatant was removed, the cells resuspended in 125 ml of sterile H2O and plated on yeast minimal media plates lacking leucine. Recombinants were detected after two days at 30 °C, selected colonies restreaked on a second minimal media plate, grown an additional two days at 30 °C, and screened for proper insertion of the oligonucleotide target sites adjacent to the LEU2 gene by PCR amplification (using a *LEU2* specific primer and a copy of the inserted oligonucleotide as primers for amplification) and DNA hybridization.

PCR amplification was performed using two 20 base oligonucleotides which hybridize approximately 250 bp on either side of the oligonucleotide insertion site. Chromosomal DNA was prepared from 0.5 mL of yeast liquid culture by glass bead lysis followed by PCR amplification with Taq polymerase (Perkin Elmer Cetus) according to manufacturer's protocol. The 500 bp amplified products were digested with either Xho I or Hind III. Proper insertion of the oligonucleotide target sites yields a PCR product cut only with Hind III, whereas an improperly modified strain is cut by Xho I. Intact chromosomal DNA from Hind III cut transformants was prepared in an agarose matrix, separated on a pulsed-field gel and screened for unique insertion of the LEU2 gene within chromosome III by Southern blotting with the random primer labeled EcoR I/Kpn I fragment from pUCLEU2C. A yeast strain meeting these criteria was then used in alkylation reactions with *N*-bromoacetyloligonucleotides.

Analysis of alkylation/cleavage reactions to nucleotide resolution.

Radio-labeling. Plasmid pUCALK was linearized with Pst I, 3'-end-labeled using terminal transferase and ^{32}P-ddATP[115] in the buffer provided by BRL.[115] After incubation at 37 °C for 45 minutes, unincorporated nucleotides were removed by phenol extraction (2x), precipitation with NaOAc (2x), and washing with 70% ethanol. Digestion with Ssp I produced bands 659 and 2064 bp in length. The 659 bp restriction fragment was isolated after electrophoresis through 5% polyacrylamide (non-denaturing) by elution of the product band at 37 °C using a solution of 200 mM NaCl, 1 mM EDTA. The gel pieces were removed by filtration through a 0.45 μM (Centrex) filter, and the DNA purified by butanol extraction and a series of ethanol precipitations.

DNA for end-product analysis was obtained by digestion with Pst I (3' end-label) or EcoR I (5' end-label). After digestion and precipitation of the DNA, 3' end-labeling was accomplished using terminal transferase,[115] using the buffer provided (BRL) and ^{32}P-2',3'-dideoxy-ATP.[115] After digestion with Eco0109, a 783 bp fragment was isolated after 5% polyacrylamide gel electrophoresis as described above. Labeling of the 5' end was accomplished by linearization with EcoR I in 100 μL, phenol extraction to remove the restriction endonuclease, addition of 100 μL of 1x CAP buffer and removal of the terminal 5'-phosphate with CAP (2 units). After phenol extraction (twice), washing with diethyl ether (three times), and NaOAc/ethanol precipitation, radio-labeling was accomplished using ^{32}P-γ-ATP and polynucleotide kinase (37 °C, 1 hr.). Unincorporated nucleotides were removed by phenol extraction (twice), precipitation with NaOAc (twice), and washing with 70% ethanol. After digestion with AlWNI, the 863 bp fragment was isolated after 5% polyacrylamide gel electrophoresis as described previously.

The plasmids pUCLEU2D and pUCLEU2C were radio-labeled by first digesting with Nar I, resulting in the production of linear pieces of DNA 5597 and 1134 base pairs in length. The DNA was labeled at the 5' end using polynucleotide kinase and γ-^{32}P ATP, and at the 3' end using the Klenow fragment to incorporate α-^{32}P-dCTP and α-^{32}P-dGTP. After labeling, unincorporated nucleotides were removed using a 1 ml Sephadex G-50-80 spin column. The DNA was digested with EcoR I, producing bands 4.5, 1.1, 0.9, and 0.2 kilobase pairs in size. The 0.9 kbp band, containing the target sequence, was isolated from a 5% polyacrylamide gel and eluted at 37 °C by incubation of the ground gel slice in a solution of 200 mM NaCl, 1 mM EDTA. The ground gel was removed by filtration through a 0.45 μM (Centrex) filter, and the DNA purified by butanol extraction and a series of ethanol precipitations.

Cleavage reactions. Approximately 10,000 cpm of DNA was incubated in 20 μL containing 0.8 mM $Co(NH_3)_6^{+3}$, 10 mM Hepes pH 7.4 and 100 nM *N*-bromoacetyloligonucleotide. After 24 - 36 hours, the reactions were precipitated with the addition of NaOAc and EtOH, washed, and treated with 1.0% piperidine at 90° C for 30 minutes The samples were lyophilized, resuspended in formamide loading buffer, and loaded onto an 8% (studies on pUCALK plasmid DNA) or 6% (studies on pUCLEU2C and pUCLEU2D plasmid DNA) denaturing polyacrylamide gel for analysis. Cleavage efficiencies were determined by phosphorimaging (*vide infra*).

Rate constants for the reaction of *N*-bromoacetyloligonucleotide **5** with the modified strand in pUCALK and with each strand in pUCLEU2C were determined by running several parallel reactions. A reaction mixture containing 0.8 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes, pH 7.4, and 1 μM *N*-bromoacetyloligonucleotide in a final volume of 150 μL was divided into 10

reactions (15 μL per reaction). The tubes were incubated at 37 °C, and individual tubes were quenched at appropriate time intervals by addition of 0.1 volume of 3 M NaOAc and 3 volumes of ethanol, and storage at -78 °C. The reactions were precipitated, washed, and treated with piperidine (1.0% aqueous piperidine, 90 °C, 30 minutes). Phosphorimaging of the 6 % gel used to separate the cleavage products was used to determine the fraction of intact and cleaved DNA. Plotting ln(fraction of intact DNA) vs. time yielded pseudo first-order kinetics from which rate constants were derived.

Phosphorimaging. Phosphorimaging analysis was accomplished using a Molecular Dynamics PhosphorImager 400S and IQ software. Typically, a gel was exposed to a phosphor-storage screen overnight. The intensity of bands corresponding to alkylation/cleavage and uncut DNA was determined for both lanes containing DNA incubated with and without *N*-bromoacetyloligonucleotides. Band intensity was determined using boxes of equal size for each band, with a zero value obtained using a box of equal size in a non-exposed part of the gel.

Double strand modification and cleavage of plasmid DNA.

Double strand cleavage in solution. Plasmid pUCLEU2C was linearized with Hind III, resulting in a linear piece of DNA approximately 6.7 kilobase pairs in length. Approximately 1 μg of DNA was incubated in 0.8 mM $Co(NH_3)_6^{+3}$, 10 mM Hepes pH 7.4, and 100 nM *N*-bromoacetyloligonucleotide (final volume 20 μL) for 24 hours at 37 °C. The DNA was precipitated with NaOAc/ethanol, washed in 70% ethanol, dried, and treated with 100 μL of 0.1% piperidine, 100 mM NaCl, 10 mM Tris pH 8.0, 10 mM EDTA at 55 °C for 12 hours. The DNA was loaded onto a 0.8% ethidium containing agarose gel and electrophoresed for 5 hours at 100 V. Size standards were obtained by

digesting circular pUCLEU2C with Dra I (yielding fragments 0.2, 0.7, 1.0, and 4.8 kbp in size) and Ssp I (yielding fragments 0.1, 0.5, 1.2, 1.5, and 3.4 kbp in size).

Cleavage efficiencies were determined from experiments using radiolabeled DNA and phosphorimaging. Hind III linearized DNA was labeled at both ends using Klenow and α-^{32}P-dATP. Unincorporated radioactive nucleotides were removed using a 1 ml Sephadex G-50-80 spin column. The labeled DNA was incubated with *N*-bromoacetyloligonucleotide **5** (0.8 mM $Co(NH_3)_6^{+3}$, 10 mM Hepes pH 7.4, and 100 nM *N*-bromoacetyloligonucleotide **5**), treated with 0.1 % piperidine (100 mM NaCl, 10 mM EDTA, 10 mM Tris, pH 8.0, 55 °C, 12 hrs.), and separation of the cleaved fragments achieved using 1% agarose gel electrophoresis. Labeled DNA was also digested with EcoR I, Nar I, and Sal I to produce labeled markers approximately 4.8, 3.4, 3.1, 2.6, 2.0, and 1.1 kbp in size. After electrophoresis, the gel was dried and exposed. Quantitation was accomplished using phosphorimaging as above.

Cleavage of plasmid DNA in agarose plugs. Cleavage of double helical DNA within agarose plugs was effected by mixing Hind III linearized pUCLEU2C DNA doubly ^{32}P end-labeled at the 3' end in 1.6 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.0 to a final volume of 15 µL. An equal volume of 2.0% agarose solution (50 °C) in water was added, the solution vortexed and chilled immediately on ice. The resulting plug was incubated in approximately 100 µL of 0.8 mM $Co(NH_3)_6^{+3}$, 10 mM Hepes pH 7.0 and *N*-bromoacetyloligonucleotide **5** added to a concentration of 500 nM. The plug was incubated at 37 °C for 36 hours, washed 3x with 900 µL of 0.1% piperidine, 100 mM NaCl, 10 mM Tris pH 7.4, 10 mM EDTA. To effect strand scission, the plug was soaked in 100 µL of piperidine solution and heated to 60 °C (12 hrs). The DNA plug was melted, loading buffer added to the solution,

and the solution was loaded directly. After separation of the products by electrophoresis (100 V, 4 hrs.), cleavage efficiencies were determined by phosphorimaging.

Agarose may contain nucleophilic sites capable of reaction with bromoacetyl moieties. To protect these sites, DNA was incubated in agarose pre-treated by incubation bromoacetamide. Agarose (1.0%) was poured into 6 x 6 x 100 mm molds to afford a high surface:volume ratio. The resulting plugs were incubated with 10 mM bromoacetamide at 37 °C for 36 hours, after which they were washed extensively with water. After washing, the plugs were melted, and plasmid DNA imbedded within them as described above. The final agarose concentration in these studies was 0.8%. Reactions with these plugs were run as described above.

Ligation of double strand cleavage products with Sfi I cut ends.

The plasmid pUCSfi was derived by insertion of the oligonucleotide sequence 5'-GATCGGCCTCTAGGGCCCTGCA-3' into the polylinker region of the BamH I and PstI linearized parent plasmid, pUC19. Because the insertion does not contain a frame shift in the lacZ gene and the two amino acid substitutions are conservative, the resulting plasmid continues to yield blue colonies in α-complementation assays and contains a unique Sfi I site. Subsequent insertion of DNA fragments into the Sfi I site can thus be screened via blue/white α-complementation assay.

Production of ligatable ends. pUCLEU2C DNA was digested with EcoR I to completion and cleaved in high yield using *N*-bromoacetyloligonucleotide **5**. The resulting 450 and 600 bp fragments were gel purified and cleaned with Prep-a-Gene matrix (Bio-Rad) according to manufacture's protocol. Five μg of pUCSfiI DNA was cut with Sfi I and EcoR I, gel purified and desalted as above.

Ten ng of pUCSfiI DNA in ten mL T4 DNA ligase buffer was added to two molar equivalents of either of the oligonucleotide/EcoR I fragments from pUCLEU2C, heated to 55 °C for 5 minutes to denature all sticky ends, cooled to 16 °C, and allowed to react with 10 units T4 DNA ligase for four hours. The ligation mixture was used to transform competent *E. Coli* (XL1-Blue, Stratagene) according to manufacturer's protocol. Transformed cells were spread on TYE plates containing the antibiotic ampicillin (50 mg/ml) and X-gal and IPTG for α-complementation. Growth overnight at 37 °C produced white colonies which were miniprepped by SDS/NaOH lysis and sequenced by dideoxynucleotide chain termination using the USB Sequenase 2.0 kit.

Yeast experiments.

Isolation of yeast genomic DNA and incorporation into agarose. Yeast strains containing modified chromosome III were cultured overnight in 100 ml of media. The yeast cells were spun in 50 ml tubes at 5000 rpm, 4 °C, for five minutes. The pellet was resuspended in 0.9 M sorbitol, 50 mM EDTA, 20 mM Tris, pH 7.5, and spun again as above. The pellet was resuspended in 16 mL of the sorbitol solution and 1/100 volume of 2 mgs/ml zymolyase in 20 mM phosphate pH 7.0 buffer was added. The solution was incubated at 37 °C for one hour, and the cells pelleted as above. After removal of the supernatant, the pellet was resuspended in 6 mL of sorbitol solution and warmed to 37 °C. A solution of 1.8% low melt agarose, 0.9 M sorbitol, 20 mM Tris pH 7.5, 50 mM EDTA was prepared and cooled to approximately 45 °C. Equal volumes of the agarose solution and yeast suspension were mixed and rapidly poured into 6 x 6 mm plastic molds, which were chilled immediately on ice. After the agarose plugs had solidified, they were removed from the

molds and placed in 50 ml sterile polypropylene tubes. Approximately 25 mL of a solution containing 0.5 M EDTA, 20 mM Tris, pH 8.0, 1% lauryl sarcosyl, and 1 mg/ml Proteinase K was added to each tube, and the tubes incubated at 50 °C with gentle agitation overnight. The plugs were incubated with a second aliquot of Proteinase K solution at a concentration of 0.5 mg/ml for approximately two hours at 50 °C. The plugs were dialyzed extensively, first with a solution of 0.5 M EDTA, and then with a solution of 10 mM EDTA, in which they were stored at 4 °C.

Alkylation of yeast genomic DNA. A typical procedure for alkylation/cleavage of yeast chromosomal DNA is described. Changes were made to study effect of pH, oligonucleotide concentration, and number of cycles on cleavage efficiency. Agarose plugs were cut to approximately 1 mm thickness (volume approximately 40 μL) and placed in 2 ml Eppendorf tubes. The plugs were washed four times with 900 μL of 1 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.2, and 5 mM EDTA (triplex/alkylation buffer). After removal of the last wash, the plugs were incubated with 120 μL of triplex/alkylation buffer, and *N*-bromoacetyloligonucleotide **5** was added to a concentration of 1 μM. After 20 hours at 37 °C, the plugs were washed 4x in 900 μL of 10 mM Tris, pH 9.5, 10 mM EDTA for 30 minutes each, allowing the *N*-bromoacetyloligonucleotide to fully diffuse from the DNA and out of the agarose. The plugs were washed 4 times for 15 minutes each with 900 μL of triplex/alkylation buffer, and incubated with 120 μL triplex/alkylation buffer and *N*-bromoacetyloligonucleotide as before. After this second 20 hour incubation, the plugs were washed three times (15 minutes) in 900 μL of 0.1% piperidine, 100 mM NaCl, 10 mM Tris, pH 8.0, 10 mM EDTA. After removal of the final wash, the plugs were heated in 900 μL of the above piperidine

solution at 55 °C for 12 hours. The plugs were washed twice with 0.5x TBE buffer before casting into a 1.0% agarose gel and electrophoresis using a BioRad Chef-DRTM II Electrophoresis control station and cell. Electrophoresis was effected at 200 V for 24 hours, with switch times ramped from 10 to 40 seconds over the first 18 hours, and from 60 to 90 seconds over the last six hours. The gel was stained with ethidium bromide and photographed.

Hybridization with radio-labeled probes. Due to the large size of the DNA, efficient transfer to membranes requires fragmentation of the DNA, accomplished by exposure of each side of the gel to 254 nm UV transilluminator light for 40 seconds. The gel was soaked in 1.0 M NaCl, 0.5 M NaOH for 30 minutes to denature the DNA, and equilibrated in 1.5 M NaCl, 0.5 M Tris, pH 7.4 to neutralize the gel. The DNA was transferred to a Schleicher and Schuell Nytran membrane using a Stratagene Pressure Control Station (1 hour, 75 psi). Cross-linking of the membrane was accomplished using a Stratagene Stratolinker UV apparatus. The HIS4 probe was obtained from Scott Strobel, and was originally obtained from clones containing the HIS4 gene.[26,116] The probe was labeled using random probe hybridization with degenerate 6mers and α-^{32}P-dCTP.[77] Briefly, 3 µL of degenerate 6mer (90 OD units/ml), 29 µL of water, and 1 µL of DNA template were mixed, heated to 98 °C for two minutes, chilled on ice (two minutes), and incubated at 37 °C for 20 minutes. OLB buffer (10 µL) (containing dGTP, dATP, and dTTP), 5 µL of ^{32}P-α-dCTP, and 1 µL (2 units) of Klenow fragment were added and the solution incubated for ≥ four hours.[77] Enzyme and unincorporated α-^{32}P-dCTP were removed by phenol extraction and a 1 ml Sephadex G-50-80 column. Typical yields were 60 x 10^6 cpm.

The membrane was pre-hybridized with 5 mL of a solution containing 6x SSPE, 10x Denhardt's, 1% SDS, and 50 μg/ml Salmon sperm DNA at 42 °C for 2-4 hours. After removal of this solution, the membrane was washed with 2 mL of 6x SSPE, 50% formamide, and 1% SDS, 50 μg/ml salmon sperm DNA solution at 42 °C for 30-60 minutes The ^{32}P-labeled probe was denatured by incubation at 37 °C for 5 minutes with 1/10 volume of 0.1 M NaOH, and incubated with the membrane in 2 mL of 6x SSPE, 50% formamide, 1% SDS, 50 μg/ml salmon sperm DNA solution at 42 °C overnight. The blot was washed four times in 25 mL of 1x SSPE buffer, 1% SDS at 42 °C, exposed to film, and cleavage efficiencies quantitated using phosphorimaging.

Determination of oligonucleotide length requirements, cooperativity experiments, and cleavage by degenerate oligonucleotides.

The Nar I/EcoR I ^{32}P 3'-end-labeled fragment from pUCLEU2C (10,000 cpm) was incubated in 1.0 mM $Co(NH_3)_6^{+3}$, 20 mM Hepes pH 7.0 with oligonucleotides Alk-7 - Alk-19 (length studies) or Alk-9 with/without Coop-10 (cooperativity studies) at desired concentrations for ~ 36 hours at 37 °C. Exact reaction times and conditions are listed in the figure legend associated with the individual autoradiogram. The DNA was precipitated, washed, and strand scission effected using 1.0% piperidine (90 °C, 30 min.). After lyophilization and resuspension in formamide loading buffer, DNA products were separated by electrophoresis using a 6% polyacrylamide gel. Cleavage efficiencies were determined by phosphorimaging.

Determination of rate of reaction between bromoacetamide and DNA.

The Nar I/EcoR I ^{32}P 3'-end-labeled fragment from pUCLEU2C (10,000 cpm) was incubated in 50 mM NaCl, 40 mM Hepes pH 7.4. An equal volume of 2-bromoacetic acid in water was added. The DNA and bromoacetic acid were mixed to afford a final volume of 110 μL, which was aliquoted into 5 tubes containing 20 μL each. One tube was immediately quenched by addition of three volumes of 0.1 M NaOAc/EtOH, and stored at -78 °C. The remainder were incubated at 37 °C and, at appropriate time intervals, individual reactions were removed from the incubator and quenched as above. After all the reactions were quenched, the DNA reactions were precipitated, washed (70% ethanol), and treated with piperidine (1% aqueous piperidine, 90 °C, 30 min.). After lyophilization, the DNA was resuspended in formamide buffer and loaded onto a 6% polyacrylamide gel for electrophoresis and analysis by phosphorimaging.

Transcriptional Studies.

The plasmid pTPTR was constructed by Jim Maher via ligation of oligonucleotides containing the sequence 5'-GTAAGAAAGAAGAAAAGAAAAAGCTTTCTTCTTAG-3' into the plasmid pLJM1. This plasmid contains promoters for the T3 and T7 polymerases on opposite strands, allowing the study of either T3 or T7 RNA polymerases. The cloned site contains binding sites for the 19 base oligonucleotide Alk-19 and the cross-over *N*-bromoacetyloligonucleotide **23**.

Plasmid pTPTR was digested with Sal I or BamH I for studies with T7 polymerases or T3 polymerases respectively. Transcription with T7 or T3 polymerases to the end of the DNA fragment (run off transcription) results in the production of transcripts approximately 400 base pairs in length. The

triplex site is located approximately 300 base pairs from the T7 promoter, and 100 base pairs from the T3 promoter. Thus, premature termination of transcription at the triple helix site results in the production of 300 and 100 base transcripts by the polymerases T7 and T3 respectively.

Linearized plasmid pTPTR (200 ngms) was incubated in a volume of 20 μL with *N*-bromoacetyloligonucleotide (2 μM), $Co(NH_3)_6^{+3}$ (1 mM), Hepes pH 7.2 (20 mM) at 37 °C for 24 hours. The reactions were precipitated and washed, and the DNA dissolved in 10 μL of a solution containing polymerase buffer (final concentrations: 25 mM Tris pH 8.0, 70 mM NaCl, 20 mM $MgCl_2$, and 0.4 mM spermidine), CTP (400 μM), GTP (400 μM), ATP (400 μM), DTT (10 mM), 3.2 units of RNase inhibitor, two units of polymerase, and 0.04 μL of α-^{32}P-UTP per reaction. Reactions were incubated at 37 °C for one hour, whereupon 15 μL of formamide containing 0.05% xylene cyanol and bromophenol blue was added directly to the reaction. Approximately 2 μL were loaded onto a 6% polyacrylamide gel for electrophoretic analysis. The gel was run for one hour at 1500 V, not allowing unincorporated nucleotides to run into the buffer. After transfer to a 3mm Whatman filter paper, the bottom of the gel was discarded. Analysis was accomplished using phosphorimaging. For analysis of T3 polymerase inhibition, the strength of the signal for the truncated transcriptional sequences is multiplied by 4, since this fragment is only 0.25 the length of the full length transcript, thus containing only 1/4 as many radio-labeled residues.

In a second set of experiments, *N*-bromoacetyloligonucleotide TRSALK-19 was incubated with BamH I cut target DNA in three sets of six reactions. In each set, the six reactions contained: (i) no salts; (ii) 20 mM Hepes, pH 7.1; (iii) 20 mM Hepes and 6 mM Mg^{+2}; (iv) 20 mM Hepes and 2 mM spermine; (v) 20 mM Hepes, 6 mM Mg^{+2} and 2 mM spermine; (vi) 1 mM $Co(NH_3)_6^{+3}$ and 20

mM Hepes pH 7.1. After 24 hours at 37 °C, the reactions were precipitated. The first set of reactions were transcribed directly. The second set was first treated with 0.3 M NaOAc at 90 °C for 30 minutes to effect depurination, and precipitated. The third set was treated with aqueous 1% piperidine, 90 °C, 30 minutes, to effect strand scission, followed by precipitation. The reactions were transcribed by dissolution of the DNA in transcription buffer (2 mM spermine, 6 mM Mg^{+2}, 20 mM Tris pH 8.0) and addition of triphosphates, RNase inhibitor, and T3 polymerase as described above.

Studies for the analysis of site of alkylation reaction.

Analysis of DNA Termini. Radio-labeled restriction fragments from the plasmid pUCALK were obtained by digestion with Pst I for the 3' end-label or EcoR I followed by treatment with calk alkaline phosphatase for the 5' end-label. Radio-labeling with 32P was accomplished using terminal transferase[115] (3' end-label) or polynucleotide kinase (5' end label), followed by digestion with a second restriction endonuclease to produce 3' end-labeled (Pst I/Eco0109) or 5' end-labeled (EcoR I/AlWNI) restriction fragments. Fragments were purified and isolated using 5% polyacrylamide gels as described previously.

5' end-group analysis was conducted by incubating N-bromoacetyloligonucleotide 5 with three aliquots of 3' end-labeled DNA in the presence of 1 mM Co(NH3)6+3 and 20 mM Hepes, pH 7.2. After 72 hours, the DNA was precipitated, strand scission effected using piperidine (1% piperidine, 90 °C, 30 min.), and the solution lyophilized. Two aliquots of DNA were dissolved in dephosphorylation buffer, and one unit of calf alkaline phosphatase added to each reaction. After 30 minutes at 37 °C, two additional units of CAP were added, and the solution incubated for an

additional 30 minutes. After phenol extraction to remove enzyme, the DNA was precipitated and washed. One aliquot of the DNA was treated with kinase in the presence of ATP (37 °C, 25 minutes), effecting addition of a terminal 5' phosphate.

3' end-group analysis was conducted by incubating N-bromoacetyloligonucleotide 5 with two aliquots of 5' end-labeled DNA in the presence of 1 mM Co(NH3)6+3 and 20 mM Hepes, pH 7.2. After 72 hours, the DNA was precipitated, strand scission effected using piperidine (1% piperidine, 90 °C, 30 min.), and the solution lyophilized. One aliquot was dissolved in buffer containing 10 mM Tris, pH 6.6, 10 mM Mg+2 and 10 mM 2-mercaptoethanol, and incubated with 10 units of polynucleotide kinase (37 °C, 30 minutes) After 30 minutes, an additional 10 units of kinase were added (37 °C, 30 minutes). The DNA was then precipitated.

Precipitated DNA from each of the above reactions was dissolved in formamide buffer and loaded onto a 15% polyacrylamide denaturing gel.

Analysis of product produced in alkylation reaction. Equimolar mixtures of oligonucleotides PA-3 and PA-4 (2 nmol) were mixed with *N*-bromoacetyloligonucleotide **5** (2 nmol) in 1 mM $Co(NH_3)_6^{+3}$, 10 mM Hepes pH 7.0, 30% ethanol. The solution was heated to 55 °C and allowed to cool to room temperature, then incubated at 4 °C for several hours. The reaction was transferred to a 37 °C incubator. After 36 hours, the reaction was precipitated with ethanol, and depurination was effected by incubating the DNA in a solution of 0.3 M NaOAc, pH 5.2, at 90 °C for 20 minutes. The DNA was again precipitated, washed, and digested with mononucleosides using phosphodiesterase and calf alkaline phosphatase. The DNA was suspended in a buffer of 100 mM $MgCl_2$, 50 mM Tris, pH 8.1. Digestion with phosphodiesterase and calf alkaline phosphatase was accomplished for 2

hours at 37 °C, after which the reaction was lyophilized, suspended in 10 µl of water, and injected for anlaysis on HPLC. Analysis was conducted on a reverse phase C-18 column (Brownlee Aquapore OD-300 C-18 column (200 mm x 4.6 mm) on a Hewlett-Packard HP 1090 using the following solvents and gradient:

Solvent A: 10 mM ammonium acetate, pH 5.2
Solvent B: 10 mM ammonium acetate, pH 5.2, 80% methanol.

time	%B
0-10	10%
10-20	10-20%
20-30	20-30%

Retention times under these conditions are as follows:

amino-thymidine	7.4 min.
N-bromoacetylthymidine	28.7 min.
product from alkylation rxn.	25.4 min.

Synthesis of expected product from DNA alkylation reaction.

Nucleoside 25. Nucleoside 24[64] was stirred in freshly distilled ethylene diamine for 2 days. The ethylene diamine was removed under reduced pressure and the product was dried under vacuum.

Nucleoside 26. Nucleoside 25 (90 mg, 263 µmol) was dissolved in 2 ml of water, and *N*-hydroxysuccinimidyl bromoacetate (90 mg, 342 µmol) was added. After 10 min., analysis by analytical HPLC (gradient as above) indicated that the reaction had proceeded to approximately 50%. After addition of an additional 90 mg of *N*-hydroxysuccinimidyl bromoacetate, the reaction was complete. The solution was extracted twice with ethyl acetate, the the aqueous layer frozen and lyophilized to a slightly brown solid.

Nucleoside 27. Nucleoside 26 (product from above reaction, ~ 263 μmol) was dissolved in DMF. 5'-O-DMT-2'-deoxyguanosine (300 mg, 500 μmol) was added, and the reaction stirred at 40 °C for 2 days. The solvent was removed under reduced pressure, and the reaction mixture was dissolved in water. Non-soluble product and 5'-O-DMT-2'-deoxyguanosine was removed by filtration, and the product purified on a preparative HPLC (Hewlett-Packard Series 1050), using a Vydac Protein and Peptide 218TP210 column. The solvents used to effect separation were:

A: 10 mM ammonium acetate, pH 5.2.
B: 10 mM ammonium acetate, pH 5.2; 80% methanol.

The following gradient was employed with a flow rate of 2.0 ml/min:

time	gradient
0-10 min	10% B
10-30 min.	10-30% B

The column was washed with 100% acetonitrile at 4 ml/min following each preparative run, to insure removal of 5'-O-DMT-2'-deoxyguanosine.

Retention times observed were:

N-bromoacetylthymidine	26.1
Isolated product	21.9

A major impurity formed during the reaction with a retention time of 25.0 was also observed, but was not isolated. The product was lyophilized. NMR is consistent with the formation of a guanine-thymidine adduct. ^{1}HNMR (D_2O, referenced to HDO): d 7.90 (s, 1H), 7.53 (s, 1H), 6.18 (t, J = 7.0 Hz, 1H), 4.93 (s, 2H), 4.81 (s, 1H), 4.36 (m, 1H), 3.93 (m, 1H), 3.65-3.75 (m, 2H), 3.13-3.26 (m, 4H), 2.40 (m, 2H), 2.25-2.29 (m, 4H). UV showed maxima at 252 and 273 nm.

References

(1) Dervan, P. B. *Science* **1986**, *232*, 464-471.

(2) Dervan, P. B. In *Nucleic Acids and Molecular Biology*; F. Eckstein and D. M. J. Lilley, Ed.; Springer-Verlag: Heidelberg, 1988; Vol. 2; pp 49-64.

(3) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861-6863.

(4) Schultz, P. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1983**, *80*, 6834-6837.

(5) Schultz, P. G.; Dervan, P. B. *J. Biomolec. Struct. Dyn.* **1984**, *1*, 1133-1147.

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

(7) Wade, W. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 1574-1575.

(8) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1986**, *108*, 5008-5009.

(9) Griffin, J. H.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 6840-6842.

(10) Youngquist, R. S.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 2565-2569.

(11) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 5528-5529.

(12) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 7564-7566.

(13) Sluka, J. P.; Bruist, M.; Horvath, S. J.; Simon, M. I.; Dervan, P. B. *Science* **1987**, *238*, 1129-1132.

(14) Mack, D. P.; Sluka, J. P.; Shin, J. A.; Griffin, J. H.; Simon, M. L.; Dervan, P. B. *Biochemistry* **1990**, *29*, 6561-6567.

(15) Sluka, J. P.; Griffin, J. H.; Mack, J. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 6369-6374.

(16) Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847-850.

(17) Graham, K. S.; Dervan, P. B. *J. Biol. Chem.* **1990**, *265*, 16534-16540.

(18) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645-650.

(19) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059-3061.

(20) Griffin, L. C.; Dervan, P. B. *Science* **1989**, *245*, 967-970.

(21) Horne, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 2435-2437.

(22) Haner, R.; Dervan, P. B. *Biochemistry* **1990**, *29*, 9761-9765.

(23) Distefano, M. D.; Shin, J. A.; Dervan, P. B. *J. Am. Chem. Soc.* submitted.

(24) Beal, P. A.; Dervan, P. B. *Science* **1991**, *251*, 1360-1363.

(25) Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 7286-7287.

(26) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73-75.

(27) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172-174.

(28) Morgan, A. R.; Wells, R. D. *J. Mol. Biol.* **1968**, *37*, 63-80.

(29) Hsieh, P.; Camerini-Otero, C. S.; Camerini-Otero, R. D. *Genes and Development* **1990**, *4*, 1951-1963.

(30) Rao, B. J.; Dutreix, M.; Radding, C. M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2984-2988.

(31) Strobel, S. A. Ph. D. Thesis, California Institute of Technology, 1992.

(32) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927-7929.

(33) Beasty, A. M.; Behe, M. J. *Nucl. Acids Res.* **1988**, *16*, 1517-1528.

(34) Behe, M. J. *Biochemistry* **1987**, *26*, 7870-7875.

(35) Lesser, D. R.; Kurpiewski, M. R.; Jen-Jacobson, L. *Science* **1990**, *250*, 776-786.

(36) Dervan, P. B.; Baker, B. F. *Ann. N. Y. Acad. Sci.* **1986**, *471*, 51.

(37) Baker, B. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 8266-8268.

(38) Baker, B. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 2700-2712.

(39) Mack, D. P.; Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7572-7574.

(40) Mack, D. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 4604-4606.

(41) Lawley, P. D.; Brooks, P. *J. Mol. Biol.* **1962**, *4*, 216-219.

(42) Lawley, P. D.; Brookes, P. *Biochem. J.* **1963**, *89*, 127-138.

(43) Lawley, P. D.; Orr, D. J.; Jarman, M. *Biochem. J.* **1975**, *145*, 73-84.

(44) Kriek, E.; Emmelot, P. *Biochim. Biophys. Acta* **1964**, *91*, 59-66.

(45) Mattes, W. B.; Hartley, J. A.; Kohn, K. W. *Nucl. Acids Res.* **1986**, *14*, 2971-2987.

(46) Belikova, A. M.; Zarytova, V. F.; Grineva, N. I. *Tet. Lett.* **1967**, *1967*, 3557-3562.

(47) Vlassov, V. V.; Gorn, V. V.; Ivanova, E. M.; Kazakov, S. A.; Mamaev, S. V. *FEBS Lett.* **1983**, *162*, 286-289.

(48) Vlassov, V. V.; Gaidamakov, S. A.; Gorn, V. V.; Grachev, S. A. *FEBS Lett.* **1985**, *182*, 415-418.

(49) Vlassov, V. V.; Zarytova, V. F.; Kutiavin, I. V.; Mamaev, S. V.; Podyminogin, M. A. *Nucl. Acids Res.* **1986**, *14*, 4065-4076.

(50) Vlassov, V. V.; Zarytova, V. F.; Kutiavin, I. V.; Mamaev, S. V. *FEBS Lett.* **1988**, *231*, 352-354.

(51) Boutorin, A. S.; Gus'kove, L. V.; Ivanova, E. M.; Kobetz, N. D.; Zarytova, V. F.; Ryte, A. S.; Yurchenko, L. V.; Vlassov, V. V. *FEBS Lett.* **1989**, *254*, 129-132.

(52) Knorre, D. G.; Vlassov, V. V. *Prog. Nucl. Acid Res. and Mol. Biol.* **1985**, *32*, 291-320.

(53) Summerton, J.; Bartlett, P. A. *J. Mol. Biol.* **1978**, *122*, 145-162.

(54) Meyer, R. B., Jr.; Tabone, J. C.; Hurst, G. D.; Smith, T. M.; Gamper, H. *J. Am. Chem. Soc.* **1989**, *111*, 8517-8519.

(55) Webb, T. R.; Matteucci, M. D. *Nucl. Acids Res.* **1986**, *14*, 7661-7674.

(56) Webb, T. R.; Matteucci, M. D. *J. Am. Chem. Soc.* **1986**, *108*, 2764-2765.

(57) Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 1241-1243.

(58) Iverson, B. L.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4615-4619.

(59) Vlassov, V. V.; Gaidamakov, S. A.; Zarytova, V. F.; Knorre, D. G.; Levina, A. S.; Nekona, A. A.; Podust, L. M.; Fedorova, O. A. *Gene* **1988**, *72*, 313-322.

(60) Fedorova, O. S.; Knorre, D. G.; Podust, L. M.; Zarytove, V. F. *FEBS Lett.* **1988**, *228*, 273-276.

(61) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 9428-9430.

(62) Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 560-564.

(63) Maxam, A. M.; Gilbert, W. *Methods in Enzymology* **1980**, *65*, 499.

(64) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 968-972.

(65) Agrawal, S.; Christodoulou, C.; Gait, M. J. *Nucl. Acids Res.* **1986**, *14*, 6227-6245.

(66) Haralambidis, J.; Duncan, L.; Angus, K.; Treagear, G. W. *Nucl. Acids Res.* **1990**, *18*, 493.

(67) *Oligonucleotide Synthesis: A Practical Approach*; Gait, M. J., Ed.; IRL Press: Oxford, 1984.

(68) Pelligrini, M.; Oen, H.; Cantor, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1972**, *69*, 837-841.

(69) Hartmann, F. C.; Suh, B.; Welch, M. H.; Barker, R. *J. Biol. Chem.* **1973**, *248*, 8233-8239.

(70) de Groot, N.; Lapidot, Y.; Panet, A.; Wolman, Y. *Biochem. Biophys. Res. Comm.* **1966**, *25*, 17-22.

(71) Fritz, H.-J.; Belagaje, R.; Brown, E. L.; Fritz, R. H.; Jones, R. A.; Lees, R. G.; Khorana, H. G. *Biochemistry* **1978**, *17*, 1257-1267.

(72) Greenhut, J.; Rudloph, F. B. *J. Chromatography* **1985**, *319*, 461-466.

(73) Zon, G.; Thompson, J. A. *BioChromatography* **1986**, *1*, 22-32.

(74) Iverson, B. L.; Dervan, P. B. *Nucl. Acids Res.* **1987**, *15*, 7823-7830.

(75) Zakrzewska, K.; Pullman, B. *J. Biomeol. Struct. Dyn.* **1985**, *3*, 437-444.

(76) Arnott, S.; Bond, P. J.; Selsing, E.; Smith, P. J. C. *Nucl. Acids Res.* **1976**, *3*, 2459-2470.

(77) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, N. Y., 1989.

(78) Eritja, R.; Walker, P. A.; Randall, S. K.; Goodman, M. F.; Kaplan, B. E. *Nucleosides and Nucleotides* **1987**, *6*, 803-814.

(79) Takeshita, M.; Chang, C.-N.; Johnson, F.; Will, S.; Grollman, A. P. *J. Biol. Chem.* **1987**, *262*, 10171-10179.

(80) Robinson, J. S. *Mol. Cell. Biol.* **1988**, *8*, 4936-4948.

(81) Kahne, D.; Still, W. C. *J. Am. Chem. Soc.* **1988**, *110*, 7529-7534.

(82) Jencks, W. P. *Catalysis in Chemistry and Enzymology*; McGraw-Hill: New York, 1969.

(83) Jencks, W. P. *Adv. Enzymol. Relat. Areas Mol. Biol.* **1975**, *43*, 219.

(84) Fersht, A. *Enzyme Structure and Mechanism*; W. H. Freeman and Company: New York, N. Y., 1977.

(85) Bruice, T. C.; Brown, A.; Harris, D. O. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 658.

(86) Bruice, T. C. *Annu. Rev. Biochem.* **1976**, *45*, 331.

(87) Dafforn, A.; Koshland, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 2463.

(88) Kirby, A. J. *Adv. Phys. Org. Chem.* **1980**, *17*, 183.

(89) Menger, F. M. *Acc. Chem. Res.* **1985**, *18*, 128.

(90) Breslow, R. A. *Pure and Appl. Chem.* **1990**, *62*, 1859.

(91) Cram, D. J.; Katz, H. E.; Dicker, I. B. *J. Am. Chem. Soc.* **1984**, *106*, 4987-5000.

(92) Van Dyke, M. W.; Dervan, P. B. *Nucl. Acids Res.* **1983**, *11*, 5555-5567.

(93) Van Dyke, M. W.; Dervan, P. B. *Biochemistry* **1983**, *22*, 2373-2377.

(94) Gordon, A. J.; Ford, R. A. *The Chemist's Companion: A Handbook of Practical Data, Techniques, and References*; John Wiley & Sons: New York, N. Y., 1972.

(95) Loke, S. L.; Stein, C. A.; Zhang, X. H.; Mori, K.; Nakanishi, M. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3474-3478.

(96) Griffin, L. C. Ph. D. Thesis, California Institute of Technology, 1990.

(97) Christophe, D.; Cabrer, B.; Bacolla, A.; Targovnik, H.; Pohl, V.; Vassart, G. *Nucl. Acids Res.* **1985**, *13*, 5127-5144.

(98) Hanvey, J. C.; Shimizu, M.; Wells, R. D. *Nucl. Acids Res.* **1989**, *18*, 157-161.

(99) Cantor, C. R.; Schimmel, P. R. *Biophysical Chemistry Part III: The behavior of biological macromolecules*; W. H. Freeman and Co.: San Francisco, CA, 1980.

(100) Hill, T. L. *Cooperativity Theory in Biochemistry: Steady State and Equilibrium Systems*; Springer-Verlag: New York, N. Y., 1985.

(101) Ptashne, M. *A Genetic Switch*; Blackwell Scientific Publications and Cell Press: Palo Alto, CA, 1986.

(102) Johnson, A. D.; Poteete, A. R.; Lauer, G.; Sauer, R. T.; Ackers, G. K.; Ptashne, M. *Nature* **1981**, *294*, 217.

(103) Adhya, S. *Ann. Rev. Genet.* **1989**, *23*, 227.

(104) Maher, L. J., III; Dolnick, B. J. *Arch. Biochem. Biophys.* **1987**, *253*, 214-220.

(105) Maher, L. J., III; Dolnick, B. J. *Nucl. Acids Res.* **1988**, *16*, 3341-3358.

(106) Maher, L. J., III; Dervan, P. B.; Wold, B. *Biochemistry* submitted.

(107) Maher, L. J., III; Wold, B.; Dervan, P. B. *Science* **1989**, *245*, 725-727.

(108) Cameron, V.; Uhlenbeck, O. C. *Biochemistry* **1977**, *16*, 5120-5126.

(109) Herzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23*, 3934-3945.

(110) van de Sande, J. H.; Ramsing, N. B.; Germann, M. W.; Elhorst, W.; Kalisch, B. W.; Kitzing, v. E.; Pon, R. T.; Clegg, R. C.; Jovin, T. M. *Science* **1988**, *241*, 551-557.

(111) Ogilvie, K. K. *Can. J. Chem.* **1973**, *51*, 3799-3807.

(112) Ogilvie, K. K.; Thompson, E. A.; Quilliam, M. A.; Westmore, J. B. *Tet. Lett.* **1974**, 2865-2868.

(113) Smith, M.; Rammier, D. H.; Goldberg, I. H.; Khorana, H. G. *J. Am. Chem. Soc.* **1962**, *84*, 430-440.

(114) Yanisch-Perron, C.; Vieira, J.; Messing, J. *Gene* **1985**, *33*, 103-119.

(115) Yousaf, S. I.; Carroll, A. R.; Clarke, B. E. *Gene* **1984**, *27*, 309-313.

(116) Deshaies, R. J.; Schekman, R. *J. Cell. Biol.* **1987**, *105*, 633.