

**Oligonucleotide-Directed Alkylation of Double
Stranded DNA: A Product and Kinetic Analysis**

Thesis

by

Matthew Jabez Taylor

In Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

California Institute of Technology
Pasadena, California

1996

(Submitted April 17, 1996)

© 1996

Matthew Jabez Taylor

All Rights Reserved

"The important thing here is that to enrich our relationships we need to make small changes. Big changes generally require some suppression of who we truly are. This is not good."

J. G.

"I have no problem with that."

P. I.

"READY?"

O. M.

Acknowledgments

I would first like to thank my advisor, Professor Peter B. Dervan. I feel that one of the very best decisions I made was to get to Caltech and spend time in Peter's group. The luck of having the examples of both the science and management styles that Peter provides I haven't taken for granted, and I look forward through my career in science to learning all that I learned in the Dervan group.

I thank my committee members, Professors Dennis Dougherty, Andy Myers, and Doug Rees for participating in my education throughout graduate school and teaching me the importance of being clear, concise, and composed.

Professor Paul Hopkins provided me my first example of how research is done. It was a great one, and prepared me for what was to come in graduate school. I appreciated the opportunity to work with the talented people in his group.

Leemor Joshua-Tor, Kyoung-Hee Kim, and Jean-Francois Mouscadet were people I collaborated with in my time at Caltech. You know your at a top institution when working with people such as these.

Thanks to Dr. Jeff Dial for getting me started and everyone else at O. C. for giving me a strong foundation in science, Professor Norman Davidson for his wisdom, and Nancy Cooper for her faith and support.

Thanks to others: Ramesh, whose commitment to science is beyond most any I've seen, and Eldon, who shows that the impossible isn't so hard if you just keep it simple - you guys made this *alot* of fun. Scott and Jen for hope. Will for his love of chemistry. David Liberles for an opportunity to learn about DNA bending. Scott Priestley for discussions and debates about the true value of things. Milan and Joe for having a clue and keeping the group on its toes. Jim Jones for good coffee. To the cool ones - you know who you are.

Thank you Claude (aka C. Patrick), Shelley, Hannah, Jack, and of course Lucky for providing what to me have been times as close to paradise or heaven as I could get. Thank you Angelique and John for providing a haven of a different kind.

Thank you Mom, Joe, Donna, Brad, and Andrew (Jasmine too?). Though it sometimes does seem that I'm working on the wing without a parachute, you've kept things on the ground so I don't have far to fall.

Abstract

This thesis research focuses on providing a better understanding of a methodology developed in the group of Professor Peter B. Dervan for the sequence-specific alkylation of double stranded DNA. It was previously reported from this group that an oligonucleotide equipped with an electrophilic bromoacetyl moiety could be used to cleave yeast chromosome III (340 kilobase pairs) at a single site in high yield. Attachment of the nondiffusible electrophile to the 5-position of a thymine at the 5'-end of a pyrimidine oligodeoxyribonucleotide affords sequence-specific alkylation of a guanine two base pairs to the 5'-side of a local triple-helical complex. Warming and base treatment then produces cleavage at this site. We have identified the products of this reaction. Consistent with an alkylation / depurination mechanism is the formation of phosphate termini in the oligonucleotide products, and identification of the oligonucleotide-guanine conjugate produced in the reaction was achieved, demonstrating that this reaction for cleavage of double helical DNA is atom-specific for the N7-position of guanine in duplex DNA.

Limitations in specificity for these systems will be due, in part, to the rate enhancement achieved upon covalent attachment of the electrophile to the DNA-binding ligand. We now understand this in quantitative terms following a comparison of the rate of bimolecular reaction of 2-bromoacetamide at the N7-position of guanine in duplex DNA to the rate of unimolecular reaction at N7 of guanine mediated by triple helix formation. Employing a polyacrylamide gel assay, these rate constants were determined and provide data indicating a maximum rate enhancement upon localization of bromoacetyl by triple helix formation of ten-million fold. The work described here has given us a detailed understanding of this system, which couples sequence-specific recognition one million times more specific than restriction enzymes with atom-specific reactivity ten million times

faster than its bimolecular counterpart, and represents a foundation for development of future systems of this class.

Part two of this thesis describes our efforts to extend oligonucleotide-directed alkylation to double stranded RNA. Because it has been shown that only RNA single strands bind double-helical RNA by triple helix formation, an N-bromoacetyloligoribonucleotide was synthesized. Described is the synthesis of the phosphoramidite necessary for the synthesis of this novel RNA-electrophile conjugate. In collaboration with Jean-Francois Mouscadet, a postdoctoral fellow in the Dervan group, it was found that triplex helix mediated alkylation of double stranded RNA occurred with only modest efficiency. It is likely that solutions to this problem will be attained only through long-term efforts beyond the scope of this thesis.

Part three describes our attempts to determine an x-ray crystal structure of a triple-helical complex. In collaboration with Professor Doug Rees at Caltech, various triple-helical complexes of the pyrimidine•purine•pyrimidine type were prepared on a ten micromole scale and crystallography experiments were performed. Single crystals were obtained which diffracted to high resolution, and analysis of these crystals indicated a stoichiometry consisted with triple-helical complexes. No solution of any of the complexes giving x-ray quality crystals were obtained. Documented in this work is methodology for the synthesis and purification of 5-iodo-2'-deoxyuridine-containing oligonucleotides. At this time, no high resolution x-ray crystal structure of these complexes has been reported, highlighting the continued importance of efforts in this area.

Table of Contents

Acknowledgments.....	iv
Abstract.....	v
Table of Contents.....	vii
List of Figure and Tables.....	viii
 Introduction	 1
Chapter 1. Product and Kinetic Analysis of Sequence-Specific Alkylation of Double Stranded DNA by Oligonucleotide-Directed Triple Helix Formation.....	 19
<i>Section 1:</i> End-Group Analysis.....	27
<i>Section 2:</i> Released-Product Analysis.....	31
<i>Section 3:</i> Kinetic Analysis of N-Bromoacetyloligonucleotide and 2-Bromacetamide Alkylation of Double Stranded DNA.....	 54
Chapter 2. Oligonucleotide-Directed Alkylation of Double Stranded RNA.....	141
Chapter 3. X-Ray Crystallography of Triple-Helical Complexes of the Pyrimidine•Purine•Pyrimidine Type.....	 157

List of Figures and Tables

Introduction.

Figures.

Figure 1.	The chemical structure of DNA.....	3
Figure 2.	CPK model of B-form DNA.....	5
Figure 3.	Hydrogen bond donors and acceptors on the base pairs in DNA.....	7
Figure 4.	Binding site sizes for DNA binding molecules and frequency of sites.....	8
Figure 5.	CPK model of double, single, and triple stranded DNA.....	11
Figure 6.	Affinity cleavage of DNA by triplex formation ("Moser experiment").....	13
Figure 7.	Dimethyl sulfate reaction at N7 of guanine in DNA.....	14

Chapter 1.

Figures.

Figure 1.1.	N-Bromoacetyldistamycin and a model of side-by side binding of a polyamide.....	20
Figure 1.2.	A model of a pyrimidine-rich oligonucleotide bound in the major groove of double stranded DNA and base triplets formed upon binding..	22
Figure 1.3.	Systems for triplex alkylation of DNA.....	23
Figure 1.4.	Alkylation of duplex DNA by oligonucleotide-directed triple helix formation.....	25
Figure 1.5.	Products of triplex alkylation of DNA.....	28
Figure 1.6.	Characterization of the DNA termini resulting from oligonucleotide- directed alkylation and base-promoted strand cleavage.....	30
Figure 1.7.	Products of enzymatic phosphodiester hydrolysis of DNA.....	33

Figure 1.8.	Synthesis of authentic standards for HPLC analysis of products of oligonucleotide-directed alkylation of duplex DNA.....	36
Figure 1.9.	2D NMR experiment on authentic T-G standard.....	38
Figure 1.10.	Autoradiogram demonstrate alkylation under HPLC conditions.....	40
Figure 1.11.	Effect of denaturation in HPLC analysis of alkylation products.....	43
Figure 1.12.	HPLC analysis of enzymatically digested products of oligonucleotide-directed alkylation of double stranded DNA.....	45
Figure 1.13.	Expansion of an HPLC trace of a time course for triplex alkylation.....	47
Figure 1.14.	Control experiment showing reaction of 5-methyl-2'-deoxycytidine with the activated ester of bromoacetic acid.....	50
Figure 1.15.	Proposal for side product formation.....	51
Figure 1.16.	Snake venom phosphodiesterase digestion of the products of reaction of N-chloro-, N-bromo-, and N-iodo-succinimide with the aminooligonucleotide.....	52
Figure 1.17.	A plot of pH versus time of 2-bromoacetamide in HEPES buffer.....	56
Figure 1.18.	Reversed phase HPLC analysis of the starting materials and potential products in control experiments for 2-bromoacetamide stability.....	58
Figure 1.19.	HPLC time course to determine the stability of 2-bromoacetamide in the presence of HEPES buffer.....	61
Figure 1.20.	HPLC time course to determine the stability of 2-bromoacetamide in the presence of phosphate buffer.....	63
Figure 1.21.	HPLC time course to determine the stability of 2-bromoacetamide in the presence of Bis Tris Acetate.....	65
Figure 1.22.	Plasmid pUCINET.....	68
Figure 1.23.	Polyacrylamide gel used to analyze the products of oligonucleotide-	

	directed alkylation and base promoted strand cleavage within a restriction fragment.....	70
Figure 1.24.	Polyacrylamide gel used to analyze the products of oligonucleotide- directed alkylation and base promoted strand cleavage within a restriction fragment.....	72
Figure 1.25.	Method of quantitation for determination of rates.....	74
	Determination of the rate constant for triplex alkylation.....	75
Figure 1.26.	Scheme for alkylation of duplex DNA with one GC base pair.....	76
Figure 1.27.	Method for determining rates.....	78
Figure 1.28.	Alternate method for determination of rates.....	79
Figure 1.29.	Scheme for alkylation of duplex DNA with two GC base pairs.....	80
Figure 1.30.	Method for determining sums of rates.....	81
Figure 1.31.	Quantitation to obtain sums of rates.....	81
Figure 1.32.	Quantitation to obtain sums of rates.....	82
Figure 1.33.	Quantitation for determination of single rate constants.....	84
Figure 1.34.	Polyacrylamide gel used to analyze the products of intermolecular 2-bromoacetamide alkylation of a restriction fragment.....	87
Figure 1.35.	Method for quantitation of restriction fragment alkylation.....	89
	Determination of observed rate constant for alkylation of guanine in the restriction fragment.....	90
Figure 1.36.	Separation of hairpin and duplex.....	92
Figure 1.37.	Polyacrylamide gel used to analyze the products of 2-bromoacetamide alkylation of duplex DNA.....	94
Figure 1.38.	Method for quantitation of restriction fragment alkylation.....	96

	Determination of rate constant for alkylation of guanine in the restriction fragment.....	97
Figure 1.39.	Polyacrylamide gel used to analyze the products of 2-bromoacetamide alkylation of duplex DNA.....	100
Figure 1.40.	Determination of the rate constant k_2 for the intermolecular alkylation reaction of 2-bromoacetamide of duplex DNA.....	102
Figure 1.41.	Method of quantitation for determination of the effect of background on k_{obs}	104
	Determination of the effect of background on k_{obs}	105
Figure 1.42.	Polyacrylamide gel used to analyze the products of 2-bromoacetamide alkylation of duplex DNA at long reaction times.....	107
Figure 1.43.	Analysis of the data at long alkylation reaction times.....	109
Figure 1.44.	Polyacrylamide gel used to determine the effect of local triple helix formation on the rate of alkylation.....	112
Figure 1.45.	Determination of the effect of triple helix formation on the rate of 2-bromoacetamide alkylation.....	114
Tables.		
Table 1.1	Concentration dependence of the rate of N-bromoacetyloligonucleotide alkylation of duplex DNA.....	76
Table 1.2	Rates of intermolecular 2-bromoacetamide and intramolecular triple helix mediated alkylation and rate enhancement obtained.....	115

Chapter 2.

Figures.

Figure 2.1	Mechanistic classes of RNA cleavage.....	142
Figure 2.2	Ribbon model depicting oligoribonucleotide-directed alkylation of double stranded RNA.....	144
Figure 2.3	Synthesis of the phosphoramidite for triple helix mediated alkylation of RNA.....	147
Figure 2.4	Synthesis of the N-bromoacetyl oligoribonucleotide for alkylation of double stranded RNA.....	148

Chapter 3.

Figures.

Figure 3.1.	Complexes YM1-YM5.....	158
Figure 3.2.	Intramolecular triple-helical complexes A-G.....	158
Figure 3.3.	YM6 provides duplex DNA termini identical to that in drug-DNA co-crystal structures.....	159
Figure 3.4.	YM7 with an overhanging residue.....	159
Figure 3.5.	YM8 from the known crystal structure of the TFIIIA transcription factor binding site.....	159
Figure 3.6.	YM9 with a biotin moiety for co-crystallization trials with avidin and streptavidin proteins.....	160
Figure 3.7.	A "mix-and-match" approach.....	160
Figure 3.8.	Characterization of crystals obtained from sample YM12.....	163
Figure 3.9.	Complex YM18.....	161
Figure 3.10.	Complexes YM19-YM22.....	161
Figure 3.11.	5-Bromo- and 5-Iodo-2'-deoxyuridine-containing triple-helical complexes for x-ray crystallography trials.....	164

Figure 3.12. Purification and characterization of a 5-iodo-2'-deoxyuridine- containing oligonucleotide.....	166
--	-----

Introduction

Double stranded deoxyribonucleic acid (DNA) is the molecule in which genetic information is stored in the cell. Since this was first determined by Avery and coworkers over 50 years ago (1), interest in the properties of DNA has been great. The elucidation of the structure of duplex DNA by Watson and Crick in 1953 provided an opportunity for chemists to study the structure of this molecule (2). Exciting was the potential for chemical solutions to the regulation of biological processes involving the molecular recognition of DNA.

Double-helical DNA is a right-handed ladder of stacked base pairs on the interior of a sugar-phosphate backbone (Figure 1). Specificity between the base pairs is dictated by what have termed Watson-Crick hydrogen bonding where adenine (A) pairs with thymine (T) bases and guanine (G) pairs with cytosine (C) bases. It is this pairing specificity that provides fidelity in the transfer of genetic information to messenger RNA for protein synthesis and to DNA copied for cell division. Though other forms of duplex DNA exist, the most common form has been termed B-form DNA (Figure 2).

The four base pairs are distinguished by the pattern of hydrogen bond donors and acceptors displayed in both grooves of the DNA (Figure 3). It is these differences that provide potential recognition elements for sequence specific DNA binding molecules. In considering general chemical solutions to the recognition of discrete sequences of DNA, an analysis of the number of possible sequences, considering a composition of AT, TA, GC, and CG base pairs in duplex DNA, demonstrates that for the three billion base pairs in the human genome, recognition of *one out of two billion* possible sixteen base pair sequences is necessary for binding at single sites in this molecule (Figure 4) (3). Four classes of

Figure 1. A graphical description of the chemical structure of DNA. Highlighted are the three repeating components of the double helix - the bases, the sugar units, and the phosphate groups.

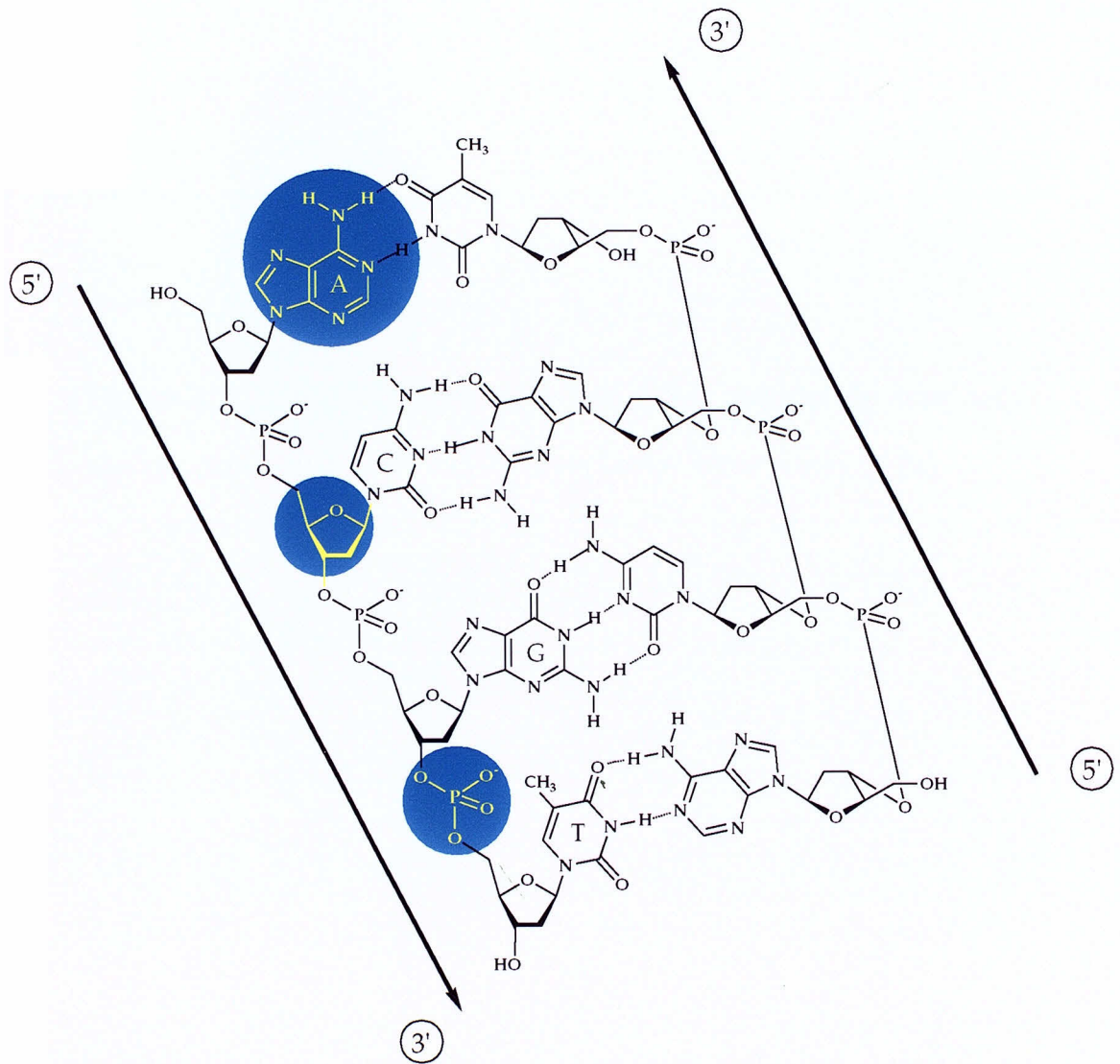


Figure 2. A space-filling model of B-form DNA, showing the wide and shallow major groove (left) and the deep, narrow minor groove (right).

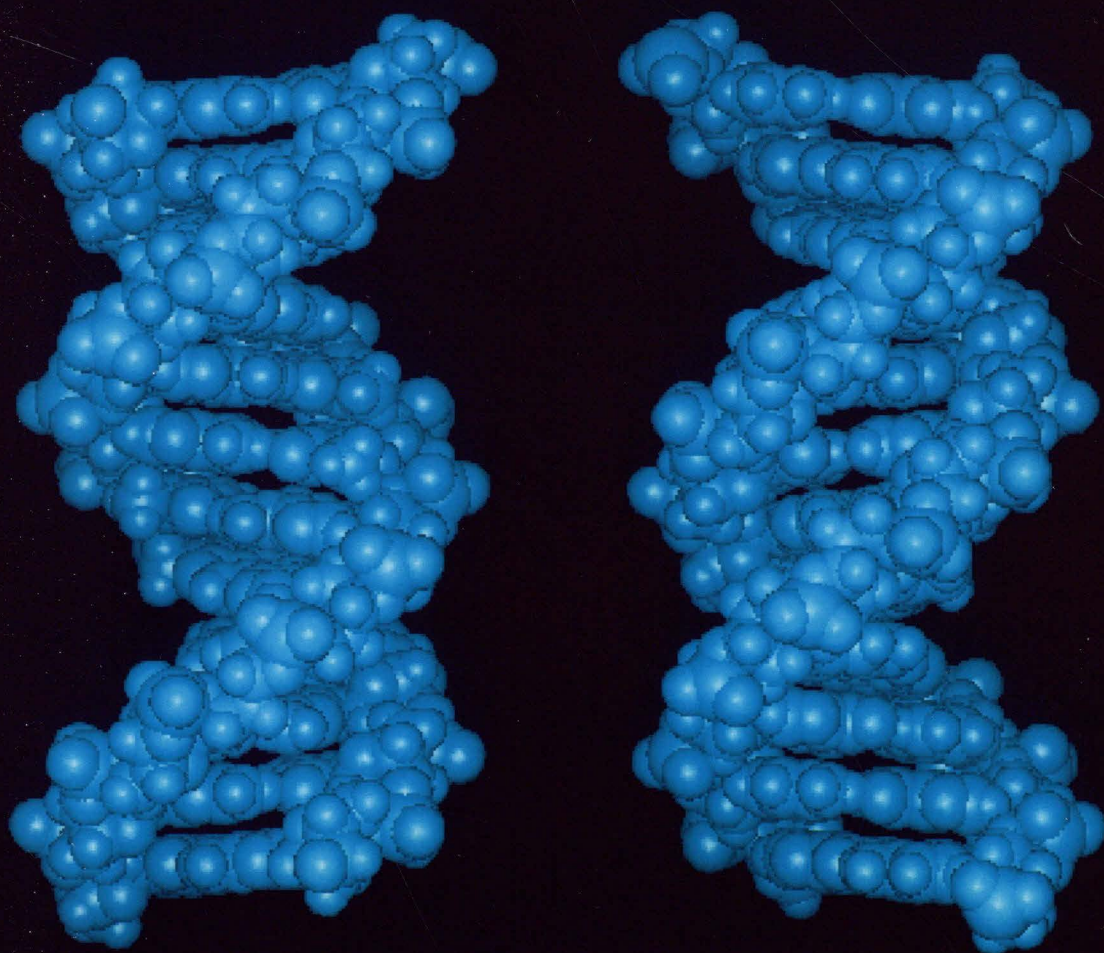
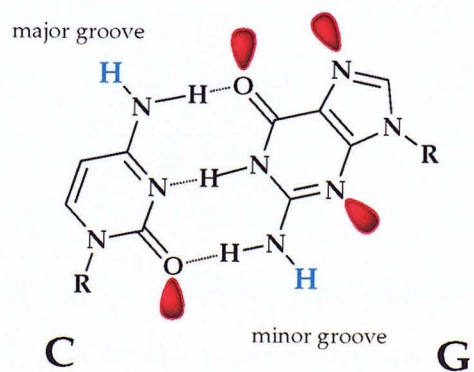
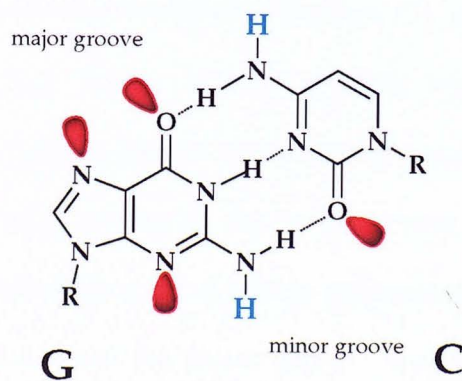
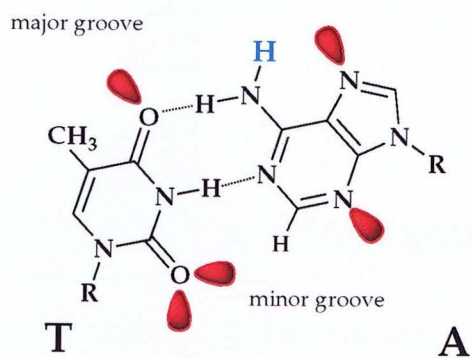
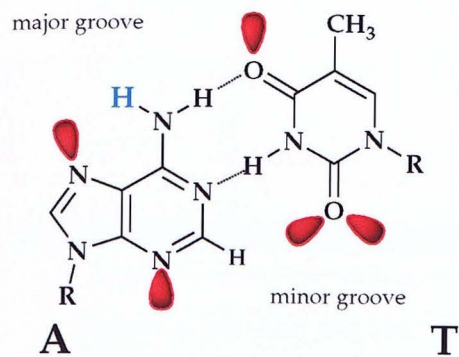


Figure 3. Representations of the four base pairs in DNA. Different base pairs display different combinations of hydrogen bond acceptors (lone pairs in red) and hydrogen bond donors (protons in blue), indicating potential elements used for sequence discrimination by DNA-binding molecules.



	n	Unique Sites
natural products	2	10
	3	32
	4	136
	5	512
restriction enzymes	6	2,080
	7	8,192
	8	32,896
	9	131,072
	10	524,800
	11	2,097,153
	12	8,390,656
	13	33,554,432
	14	~134,225,000
	15	~536,870,000
oligonucleotides	16	~2,147,510,000 Human Genome ~3X10 ⁹ bp

Figure 4. Binding site size and frequency of sequences in double stranded DNA.

DNA binding molecules for potential general *chemical solutions* to the recognition of double stranded DNA that have been investigated in the Dervan Group are proteins (4), RecA•oligonucleotide filaments (5), polyamides (6), and oligonucleotides (7). Proteins, though diverse in terms of their ability to bind duplex DNA with high affinity and specificity, present at this time no general solution for the design of sequence specific DNA binding molecules due to the protein folding problem. RecA•oligonucleotide filaments, created when RecA protein polymerizes on single stranded DNA, insert into double-helical DNA from the minor groove. Specificity is dictated by the bases within the filament. Polyamides, analogs of the natural products distamycin and netropsin, bind as antiparallel side-by-side dimers in the minor groove of DNA. At this time, these molecules represent the best class of molecules for a general chemical solution to the molecular recognition of DNA.

Three-stranded complexes of single strands of DNA have been known to exist for almost forty years. By UV mixing curves it was demonstrated that two strands of poly-U could associate with one strand of poly-A (Figure 5) (8). It was later shown that, at acidic pH, two strands of poly-C could associate with one strand of poly-G (9). In 1987, Moser and Dervan demonstrated for the first time that short oligonucleotides could bind to double-helical DNA by triple helix formation 6(7). An oligonucleotide containing thymine and cytosine bases was shown to bind duplex DNA by affinity cleavage, where an oligonucleotide is tethered to EDTA and upon binding iron can produce a diffusible oxidant which cleaves the DNA (Figure 6). This result implied that, assuming an understanding of the rules for triple helix formation existed, anyone with access to a DNA synthesizer could make a molecule to bind any desired sequence of double stranded DNA and potentially control biological processes. Others showed that nonnatural analogs of thymidine and cytidine could form triple helices, though these materials were less desirable due to nonavailability (10).

Two methods exist for triple helix mediated cleavage of duplex DNA, enzymatic (11) and chemical (12). Within enzymatic methods are two classes. The first involves tethering of a nuclease to an oligonucleotide (11a). Upon binding by triple helix formation, hydrolysis of the phosphodiester bonds in the duplex proximal to the local triplex helix occurs in high yield. Site-specificity of cleavage suffers due to flexibility in the linker between the binding and cleaving domains. A second method of enzymatic genomic cleavage has been termed triplex achilles heel cleavage (11b,c). Specificity obtained by an oligonucleotide bound to DNA overlapping a restriction enzyme site is used to selectively protect one site from a methylase. Upon dissociation of the oligonucleotide, treatment with a restriction enzyme cleaves the DNA only at the unmethylated position. This method was used to site-specifically cleave human chromosome IV.

Figure 4. Oligonucleotides form three stranded complexes under appropriate conditions.

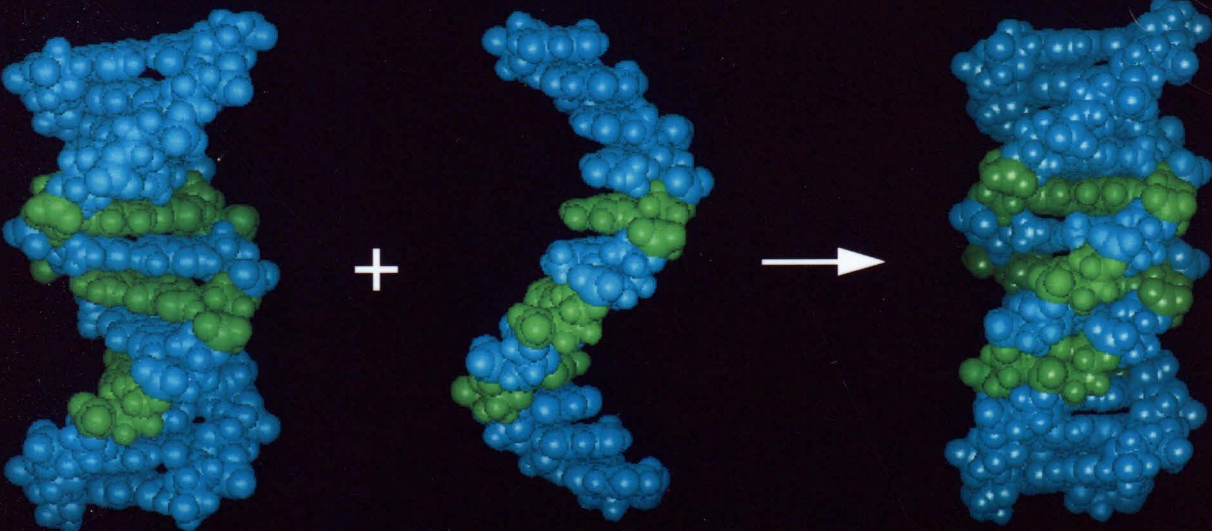
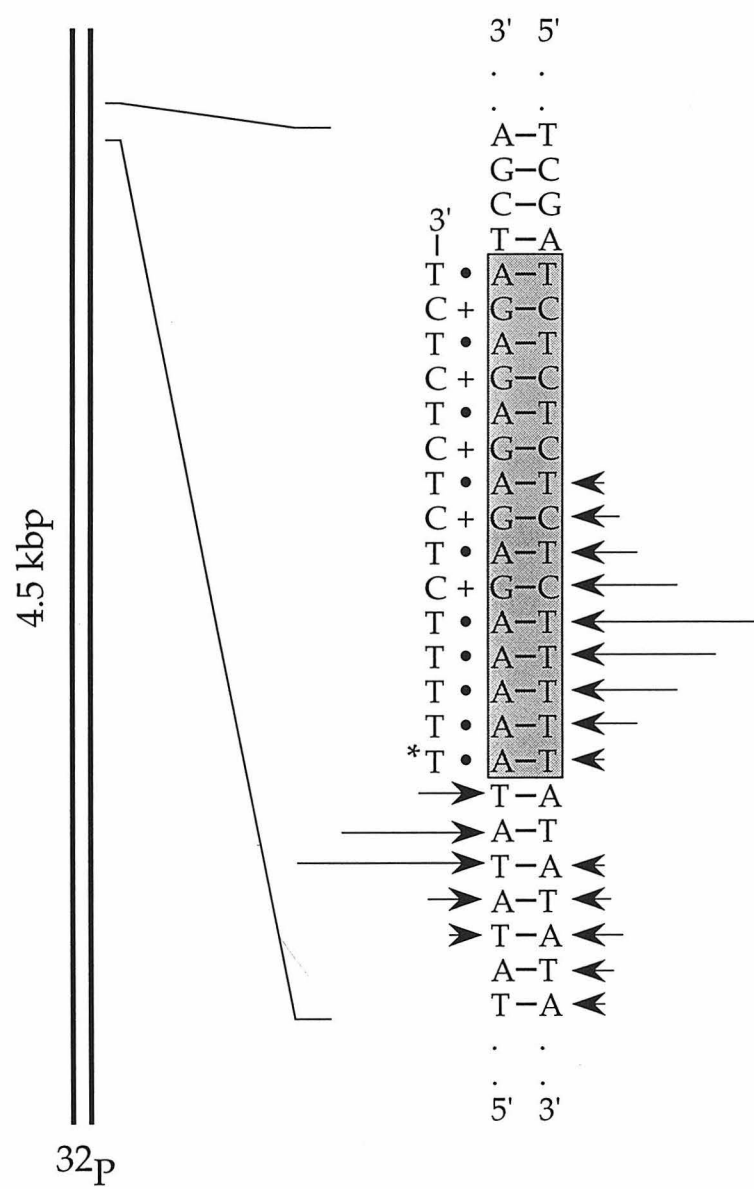


Figure 6. Affinity cleavage demonstrates that oligonucleotides bind to double helical DNA.



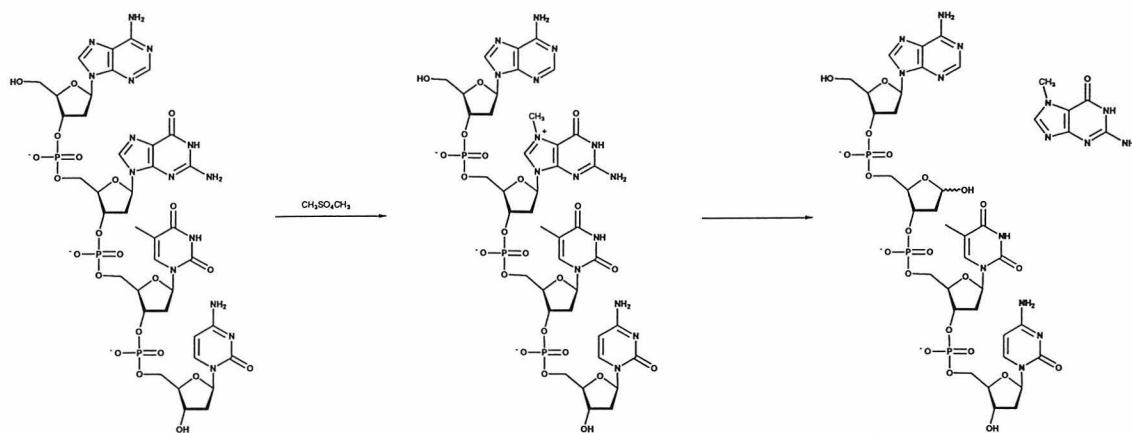


Figure 7. Dimethyl sulfate reacts predominantly at the N7 position of guanine bases in DNA. Alkylation then causes depurination to occur.

Chemical methods for triple helix mediated cleavage of genomic DNA can be grouped into mechanistic classes - oxidative (12a,b) and electrophilic (12c,d). Oxidative cleavage includes affinity cleavage, where an oligonucleotide is attached to EDTA and upon binding iron and exposure to a reductant a diffusible oxidant is generated. While this method has been used for cleavage of genomic DNA in low yield (~7%) (12b), it seems more suited for studies to answer questions of molecular recognition such as binding location and groove location. Atom-specific, high yielding methods seemed more suitable for genomic DNA cleavage (12c,d). These methods have the advantage that, where selectivity is sufficient, single products are obtained as opposed to mixtures for non-selective cleaving agents. Atom-specific chemistry, based conceptually on the dimethyl sulfate alkylation of DNA studied by Brooks and Lawley (Figure 7) (13), has been used in the Dervan group for cleavage of yeast chromosome III in 85-90% yield to nucleotide resolution (12d). The future of the field of atom-specific chemistry on duplex DNA will be to direct specific chemical reactions to inherently unreactive sites on duplex DNA, generating a class of artificial enzymes. Further design of these systems will undoubtedly

benefit from an understanding in quantitative terms of what advantage has been gained in both reaction rate and specificity upon covalent attachment of electrophiles to oligonucleotides for alkylation of double-helical DNA.

The work described in this thesis is divided into three parts. Chapter one represents the majority of this thesis work and describes a detailed analysis of a system for oligonucleotide-directed alkylation of double stranded DNA mediated by triple helix formation. The products resulting from alkylation are identified, and the kinetics of the system relative to that by free reactive molecule are determined. It was found that localization of bromoacetyl by triple helix formation results in an almost ten million-fold rate enhancement. Chapter two describes an attempt to extend this system to triple helices containing RNA strands, for alkylation of double stranded RNA. Success was modest, and further work will be necessary in this area for significant progress to be made. Chapter three describes attempts to obtain an x-ray crystal structure of a triple-helical complex. This work was done in collaboration with the Rees group at Caltech. Attempts, though promising, were unsuccessful. At this time, no solution to this elusive problem has been reported.

References

- (1) Avery, O. T.; MacLeod, C. M.; McCarty, M. *J. Exp. Med.*, **79**, 137 (1944).
- (2) (a) Watson, J. D.; Crick, F. H. C. *Nature*, **171**, 737 (1953). (b) Watson, J. D.; Crick, F. H. C. *Nature*, **171**, 964 (1953).
- (3) Dervan, P. B. *Science*, **232**, 464 (1986).
- (4) (a) Sluka, J. P.; Horvath, S. J.; Bruist, M. F.; Simon, M. I.; Dervan, P. B. *Science*, **238**, 1129 (1987). (b) Mack, D. P.; Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.*, **110**, 7572 (1988). (c) Oakley, M. G.; Dervan, P. B. *Science*, **248**, 847 (1990).
- (5) Baliga, R.; Singleton, J. W.; Dervan, P. B. *Proc. Natl. Acad. Sci. USA*, **92**, 10393 (1995).
- (6) (a) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **114**, 8783 (1992). (b) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci., U.S.A.* **89**, 7586 (1992).
- (7) Moser, H. E.; Dervan, P. B. *Science*, **238**, 645 (1987).
- (8) Felsenfeld, G.; Davies, D. R.; Rich, A. *J. Am. Chem. Soc.* **79**, 2023 (1957).
- (9) Lipsett, M. N. *J. Biol. Chem.*, **239**, 1256 (1964).
- (10) LeDoan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucleic Acids Res.* **15**, 7749 (1987).
- (11) (a) Pei, D. H.; Corey, D. R.; Schultz, P. G. *Proc. Natl. Acad. Sci. USA*, **87**, 9858 (1990). (b) Strobel, S. A.; Dervan, P. B. *Nature* **350**, 172 (1991). (c) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B. *Science* **254**, 1639 (1991).
- (12) (a) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **110**, 7927 (1988). (b) Strobel, S. A.; Dervan, P. B. *Science* **249**, 73 (1990). (c) Povsic, T. J.; Dervan,

P. B. *J. Am. Chem. Soc.* **112**, 9428 (1990). (d) Povsic, T. J.; Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **114**, 5934 (1992).

(13) Lawley, P. D.; Brookes, P. *Biochem. J.*, **89**, 127 (1963).

Chapter 1

Product and Kinetic Analysis of Sequence Specific Alkylation of Double Stranded DNA Mediated by Oligonucleotide-Directed Triple Helix Formation.

Introduction

A majority of the material described in this Chapter was submitted for publication to the Journal of the American Chemical Society.

The design of sequence-specific DNA-cleaving molecules requires the integration of recognition and cleavage in a single molecule. Three mechanistic approaches to the bifunctional design of these molecules include attachment of functional groups to DNA-binding molecules that can achieve oxidation of the deoxyribose backbone (1), electrophilic modification of the bases (2), or hydrolysis of the phosphodiester bond. Of these classes, DNA-binding molecules that perform electrophilic modification of the bases can provide additional specificity due to the differential nucleophilicity of positions on the bases in DNA (e. g. N7 of guanine or N3 of adenine).

N-Bromoacetyldistamycin, a designed molecule with two domains consisting of sequence-specific binding and alkylation of double-helical DNA, binds four five-base pair sites in the minor groove of DNA and can form a covalent bond at the N3 position of a single adenine within a 167 base pair restriction fragment of DNA (2). More recently, models for 2:1 binding modes of these molecules for the recognition of mixed sequences

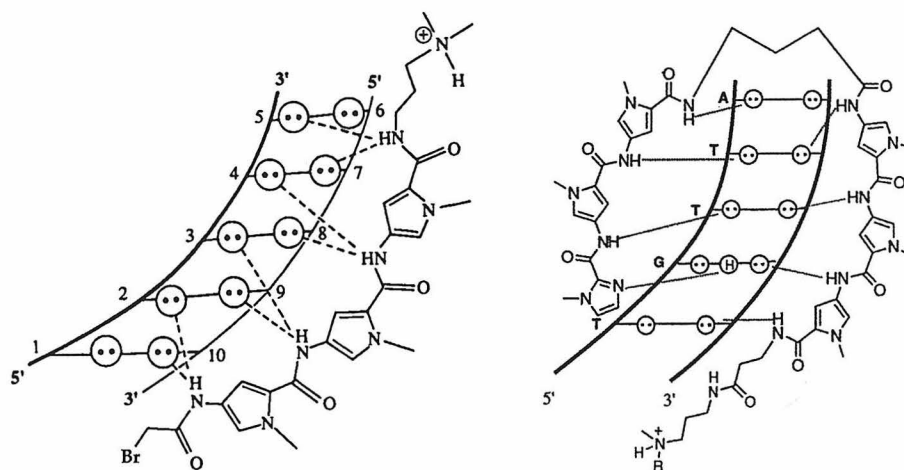
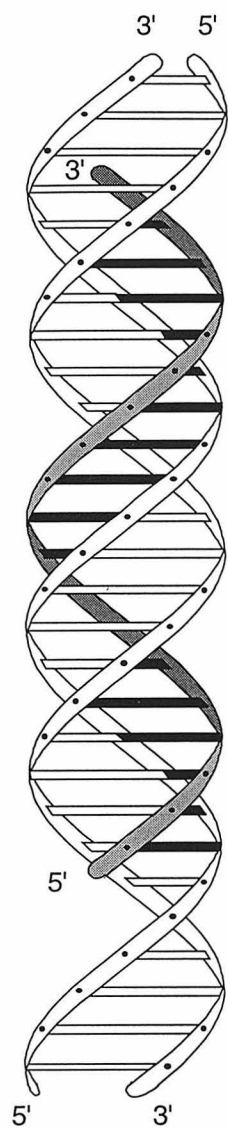


Figure 1.1. N-Bromoacetyldistamycin binds double-helical DNA and alkylates the N3-position of adenine in the minor groove (left). Side-by-side antiparallel binding of polyamides (right).

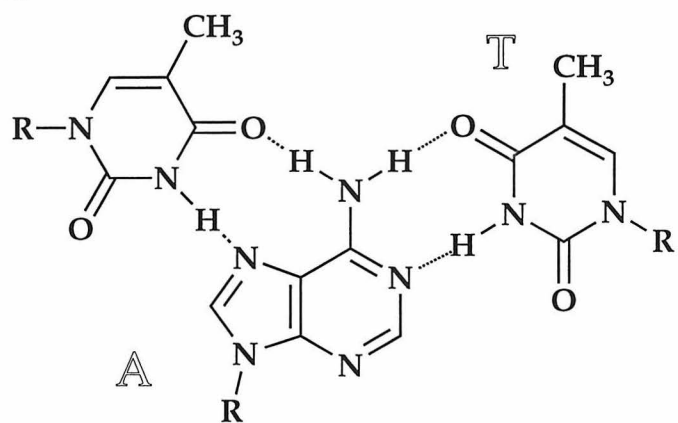
exist (3), opening new possibilities for the design of these bifunctional molecules (Figure 1.1).

Oligonucleotide-directed triple helix formation is one of the most powerful methods for the sequence-specific recognition of single sites within double-helical DNA (4,5). Pyrimidine oligodeoxyribonucleotides bind purine tracts in the major groove of DNA parallel to the purine Watson-Crick strand, through formation of specific Hoogsteen hydrogen bonds to the purine Watson-Crick bases (Figure 1.2). Specificity is derived from thymine (T) recognition of adenine-thymine (AT) base pairs (TAT triplet) and N3-protonated cytosine (C⁺) recognition of guanine-cytosine (GC) base pairs (C⁺GC triplet). The high stability and sensitivity to single-base mismatches (3fij,4cd) make this a powerful method for binding single sites in large double helical DNA (3abdgh).

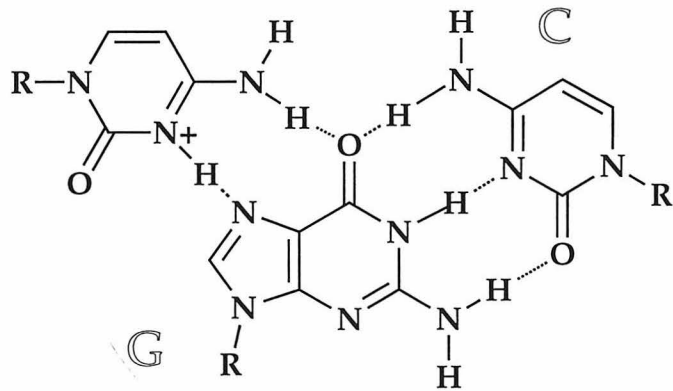
Figure 1.2. A model representing a pyrimidine-rich oligonucleotide bound in the major groove of double stranded DNA and base triplets formed upon binding.



T



C+



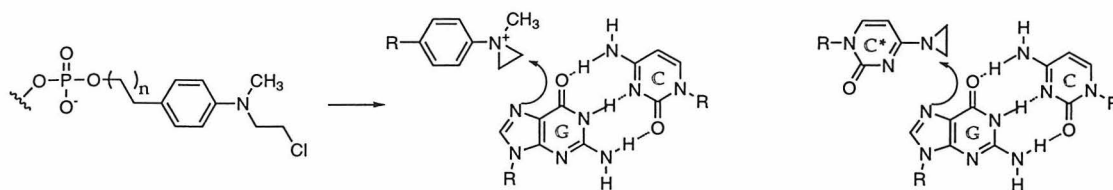
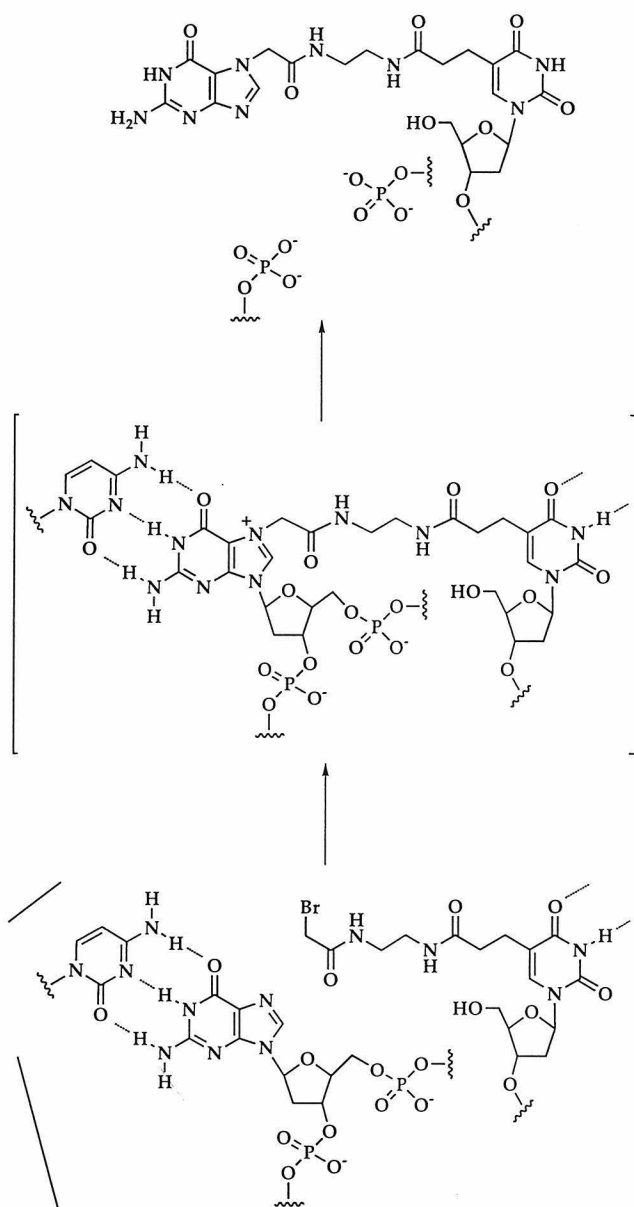
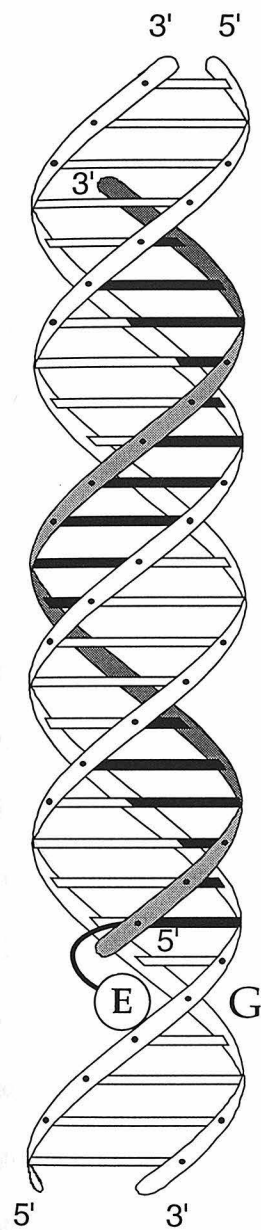


Figure 1.3. A small number of systems have been developed for oligonucleotide-directed alkylation of double stranded DNA.

Efforts have been successful in the use of oligonucleotides for directing electrophiles to react on one strand of duplex DNA at the N7-position of guanine in the major groove via triple helix formation (6). Early examples utilizing 5'-aromatic chloroethylamine-modified pyrimidine oligonucleotides demonstrated alkylation at adjacent guanine bases with modest yield (6ab). Ethano-5-methyl-2'-deoxycytidine residues at the 3'-end of oligonucleotides (Figure 1.3) (6de) and oligonucleotides equipped with N-bromoacetyl at the 5'-end (6c) react with one strand of a duplex DNA target site in very high yield (Figure 1.4). Pyrimidine oligodeoxyribonucleotides bound in the major groove and equipped with a bromoacetyl moiety at the 5'-end position the electrophile proximal to a guanine two base pairs to the 5'-side of the target sequence. Reaction of the electrophilic carbon of N-bromoacetyl with N7 of the guanine base adjacent to the local triple helix results in covalent attachment of the oligonucleotide to the target sequence. Upon warming in the presence of base, depurination at the position of alkylation occurs and cleavage of one strand of the DNA backbone is observed.

More recently, the utility of this approach to achieve alkylation of both strands of double-helical DNA has been demonstrated (7). Pyrimidine oligonucleotides modified at both the 5'-end (for same strand alkylation) and 3'-end (for opposite strand alkylation) with aromatic chloroethylamines cross-link short double-helical DNA at guanine bases

Figure 1.4. The bromoacetyl electrophile is localized in the major groove by triple helix formation proximal to a GC base pair (G) in the Watson-Crick duplex target site. Alkylation at N7 of guanine followed by depurination results in cleavage of the deoxyribose backbone.



separated by 22-26 base pairs in 80% yield (7b). N-Bromoacetyloligonucleotides bind adjacent inverted purine tracts on double-helical DNA by triple helix formation and alkylate single guanine positions on opposite strands (7a). Double strand cleavage at a single site within yeast chromosome III occurs in 85-90% yield, demonstrating the utility of this nonenzymatic approach for localization of atom-specific chemistry and cleavage of large double-helical DNA. Further design of these systems will undoubtedly benefit from an understanding in quantitative terms of what advantage has been gained in both reaction rate and specificity upon covalent attachment of electrophiles to oligonucleotides for alkylation of double-helical DNA.

Two criteria for the design of molecules that perform atom-specific reactions at single sites on double stranded DNA are i) the incorporation of a specific reaction at a designated atom within the bound ligand-DNA complex and ii) intramolecular rates of reaction that are on a reasonable time scale under physiological conditions. Success in high yield modification requires that rates of intermolecular reactions of cleaving functionalities employed with components in aqueous media be slow relative to intramolecular reaction within the bound complex. Limitations in specificity for these systems will, in part, be due to the rate enhancement achieved upon covalent attachment of the reactive functionality to the DNA-binding ligand. Advantages in specificity can be gained in cases where a slow intermolecular reaction on DNA by unbound functionality is transformed into a fast intramolecular reaction within the bound complex. This criterion can be addressed qualitatively in the context of oligonucleotide-directed alkylation of double stranded DNA by an analysis of background rates of alkylation resulting from electrophilic modification by free bifunctional molecule. In cases where background reaction is not observable under conditions where the directed reaction is fast and high yielding, rate enhancement is

acceptable (5,6). As nucleotide-specific reactivity is crucial for these approaches, the maximum rate enhancement attainable at specific sites will be limited by the affinity of the ligand for its target site in duplex DNA. High concentrations necessary for binding of bifunctional molecules having low affinities for double stranded DNA will accelerate intermolecular reactions on DNA and limit attainable rate enhancements.

To address these issues, we report a product and kinetic analysis of oligonucleotide-directed bromoacetyl alkylation of double stranded DNA. Products of N7 alkylation of guanine in i) the target duplex DNA and ii) the pyrimidine-rich Hoogsteen strand of the bound triple-helical complex are determined. Rate constants for the alkylation reaction at N7 of guanine within the triple helical complex and the reaction of 2-bromoacetamide with N7 of guanine within duplex DNA are determined, allowing for an analysis of the rate enhancement of alkylation obtained at a single position upon covalent attachment of the bromoacetyl electrophile to the Hoogsteen strand of a local triple-helical complex.

Discussion

Section 1. The production of phosphate termini is consistent with a mechanism involving base-promoted DNA cleavage at the abasic site within the target duplex resulting from alkylation at the N7 position of guanine and depurination (Figure 1.5). The 29 base pair duplex of sequence composition 5'-GCTAGTAAAAGAAAAGGAAAGAAAAGTCG-3' possessing a GC base pair two base pairs to the 5'-side of the homopurine tract was synthesized and 5'- or 3'-³²P-end labeled for 3'- and 5'-termini analysis, respectively. Radiolabeled duplexes were reacted with N-bromoacetyloligonucleotide of sequence composition 5'-

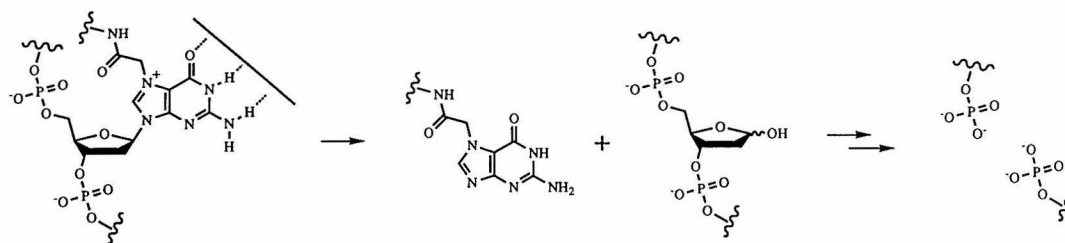
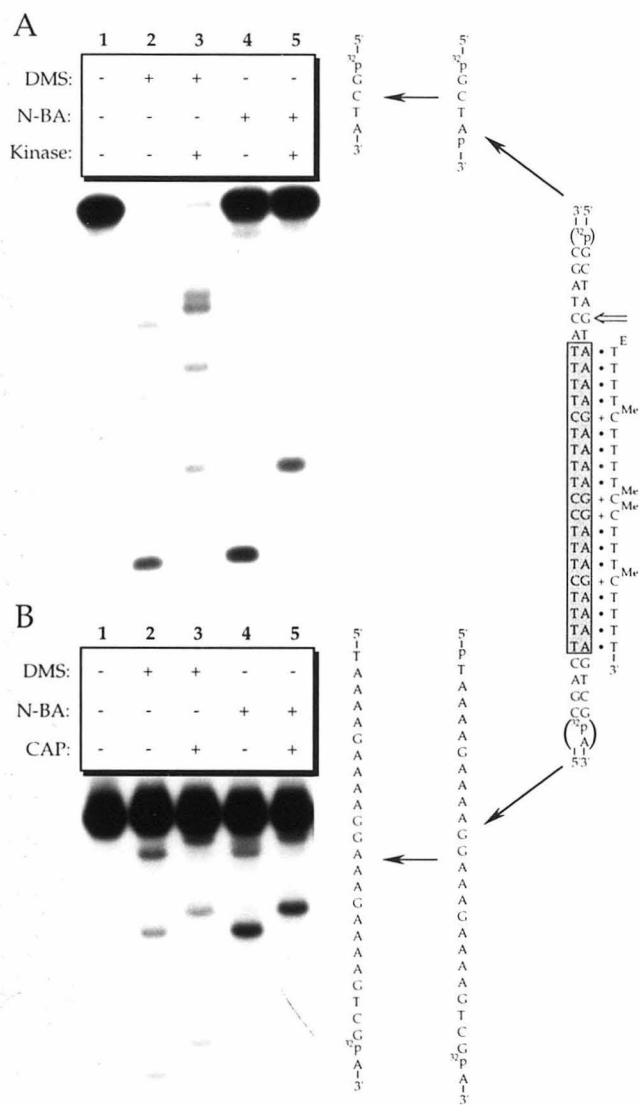


Figure 1.5. Products of oligonucleotide-directed alkylation of duplex DNA.

${}^E\text{TTTT}^{\text{Me}}\text{CTTTT}^{\text{Me}}\text{C}^{\text{Me}}\text{CTTT}^{\text{Me}}\text{CTTTT}-3'$ (8,9) at 2.5 μM concentration (20 mM HEPES, pH 6.8, 1 mM cobalt hexammine trichloride) for 24 hours at 37° C, and following piperidine treatment (90° C, 30 min) the reactions were analyzed by high resolution polyacrylamide gel electrophoresis. Treatment of the radiolabeled duplex with dimethyl sulfate followed by piperidine affords authentic standards of the expected phosphate-terminal oligonucleotides (10). Alkylation by the N-bromoacetyl oligonucleotide occurs at the targeted guanine base proximal to the local triple helix and oligonucleotide products comigrate by gel electrophoresis with products of dimethyl sulfate-treated DNA (Figure 1.6).

For further evidence of terminal phosphate functionality at the site of strand cleavage, oligonucleotide products were treated with the appropriate enzyme to achieve hydrolysis of the 5'- and 3'-phosphate groups. The 5'-end labeled oligonucleotide fragment possessing the putative 3'-phosphate produced in the alkylation reaction was subjected to treatment with T4-polynucleotide kinase, an enzyme which can catalyze the hydrolysis of 3'-phosphate groups (11). Following this treatment, a retardation of gel mobility is observed (Figure 1.6, A). The observed shift in gel mobility is identical to that observed for the products of dimethyl sulfate alkylation of DNA subjected to kinase treatment. The analysis was repeated for the 3'- ^{32}P -end labeled product possessing the putative 5'-

Figure 1.6. Autoradiograms of 20% denaturing polyacrylamide gels used for characterization of the DNA termini resulting from oligonucleotide-directed alkylation and base-promoted strand cleavage at the targeted guanine base: (A) 3'-end analysis using a 5'-end labeled 29 base pair target duplex and T4 polynucleotide kinase; (B) 5'-end analysis using a 3'-end labeled 29 base pair target duplex and calf alkaline phosphatase. Lane 1, controls - 24 hours with no N-bromoacetyloligonucleotide; lane 2, duplex treated with dimethyl sulfate; lane 3, enzymatically treated dimethyl sulfate products; lane 4, duplex treated with 2.5 μ M N-bromoacetyloligonucleotide for 24 hours; lane 5, enzymatically treated products of oligonucleotide-directed alkylation.

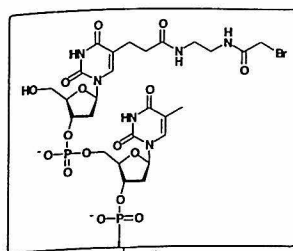


phosphate using calf alkaline phosphatase, an enzyme which can catalyze the hydrolysis of 5'-phosphate groups. Again, a product of slower mobility is produced and the observed shift in gel mobility for the phosphatase-treated products of dimethyl sulfate alkylation of DNA is equivalent (Figure 1.6, B). These data are consistent with the formation of phosphate termini at the site of strand cleavage.

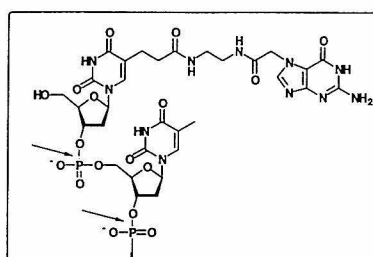
Section 2. Alkylation at the N7 atom of guanine within the Watson-Crick duplex is achieved by the bromoacetyl electrophile appended to the 5'-end of the Hoogsteen strand bound in the local triple-helical complex. Identification of the guanine-(linker)-oligonucleotide conjugate formed after depurination would provide direct evidence for this covalent bond formation. To address this issue, reversed phase HPLC analysis can be performed on the reaction products enzymatically digested by a 3'-exonuclease. Products of this hydrolysis reaction are nucleotides (resulting from internal and 3'-terminal residues) and nucleosides (resulting from 5'-terminal residues) (Figure 1.7) (12). Nucleosides resulting from modifications of the linker arm at the 5-position of the 5'-terminal thymidine in the pyrimidine-rich Hoogsteen strand can then be separated by standard reversed phase chromatographic techniques from a potentially large background of nucleotides.

Authentic standards of products expected from enzymatic digestion of unreacted N-bromoacetyloligonucleotide, as well as products of N7-alkylation of guanine and hydrolysis of starting material, were synthesized from the thymidine methyl ester **1** (Figure 1.8) (13). Amidation of the methyl ester with ethylenediamine afforded **T-NH₂** in 96% yield. Acylation of **T-NH₂** with N-hydroxysuccinimidyl 2-bromoacetate (23) provided the bromo-(linker)-thymidine compound **T-Br** in 24% yield. Acylation of **T-NH₂** with N-hydroxysuccinimidyl 2-hydroxyacetate provided the hydroxy-(linker)-

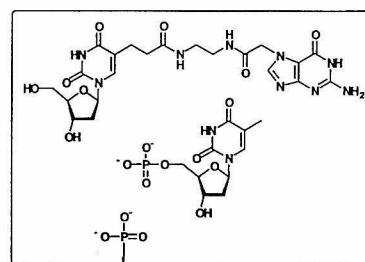
Figure 1.7. Enzymatic phosphodiester hydrolysis of the modified pyrimidine-rich Hoogsteen strand by the 3'-exonuclease snake venom phosphodiesterase produces thymidine nucleosides with linker arm modifications at the 5 position, thymidine 5'-monophosphate, and 5-methyl-2'-deoxycytidine 5'-monophosphate. These products can then be separated and identified by reversed-phase HPLC.



1) Alkylation
2) Depurination



Snake
Venom
Phosphodiesterase



thymidine compound **T-OH** in 14% yield. Synthesis of N7-carboxymethyl guanine was performed as previously described (24), affording the product in 26% yield. Acylation of **T-NH₂** with N7-carboxymethyl guanine activated *in situ* with 4-nitrophenol and 1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDCI) afforded the guanine-(linker)-thymidine standard **T-G** in milligram quantities. To assure that the authentic adduct was fully characterized, 2-dimensional NMR experiments were performed. Both COESY and NOESY 2D NMR spectral data was consistent with the structure of **T-G** (Figure 1.9).

N-bromoacetyloligonucleotide (20 nmol, 2.7 μ M) of sequence composition 5'-^ETTTT^{Me}CTTTT^{Me}C^{Me}CTTT^{Me}CTTTT-3' was bound to the 60 base pair duplex (20 nmol, 2.7 μ M) of sequence composition 5'-TCGGTACCCGGGGATCTAAAGTAAAAGAAAAGGAAAGAAAAAGCTTTCTTC TTCCCTATC-3' and a time course performed (20 mM HEPES pH 7.0, 0.2 mM cobalt hexammine trichloride, 37° C). Three nmol aliquots were removed over time and heated to 90° C to depurinate N7-alkylated guanine bases (14). Cleavage products were then digested with snake venom phosphodiesterase and subjected to HPLC analysis.

To demonstrate that alkylation occurs at the targeted guanine base in the duplex, an experiment was performed where a small amount of 5'-³²P-end labeled duplex was added to the reaction mixture. Base workup and denaturing polyacrylamide gel electrophoresis was performed and demonstrated that alkylation occurs at the guanine base proximal to the triple-helical complex (Figure 1.10).

Optimization of the HPLC conditions for analysis of these reaction products was required. The most problematic result initially was the appearance of "noise" in the HPLC trace overlapping the region where elution of the authentic guanine-(linker)-thymidine

Figure 1.8. Scheme for the synthesis of authentic standards for HPLC analysis of products of oligonucleotide-directed alkylation of duplex DNA from the known thymidine methyl ester **1**.

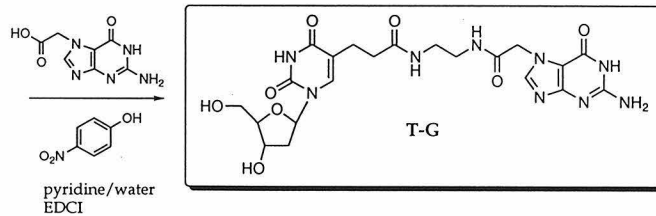
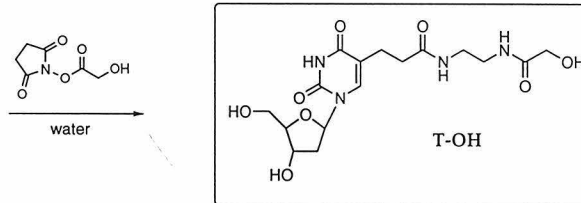
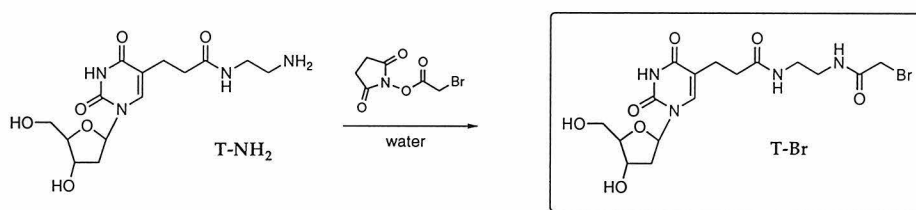
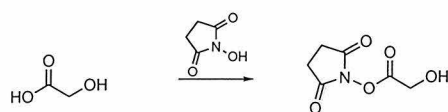
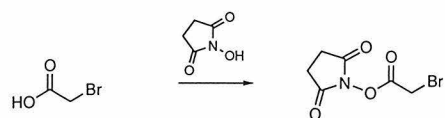
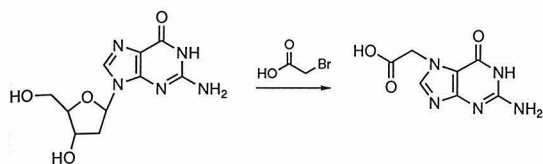
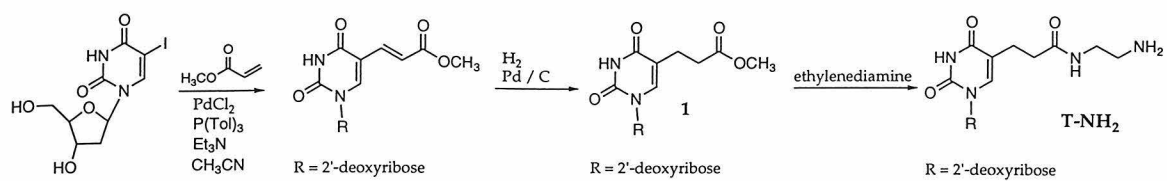


Figure 1.9. The authentic standard T-G was characterized by TOCSY and NOESY 2D-NMR experiments. Experiments were performed at 500 MHz.

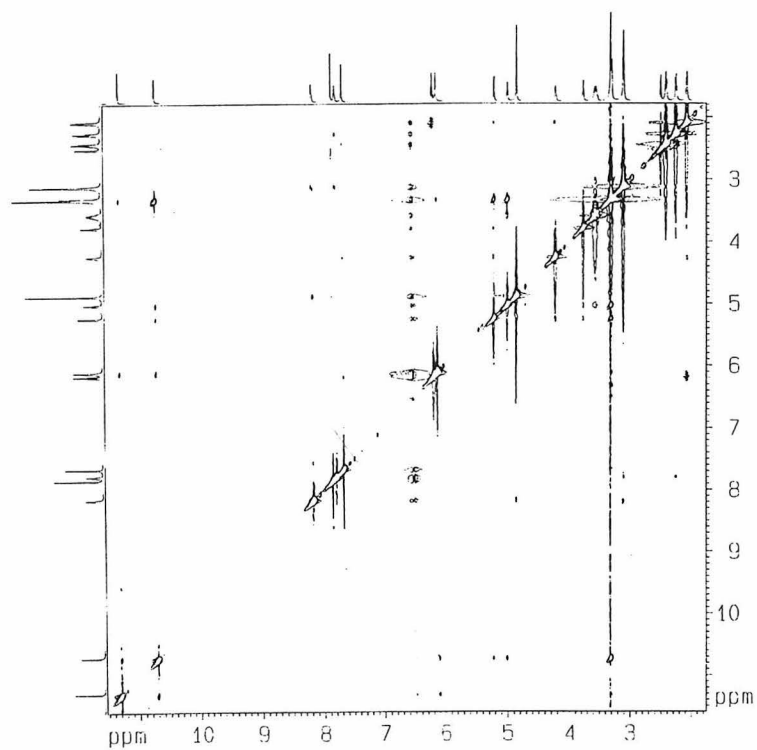
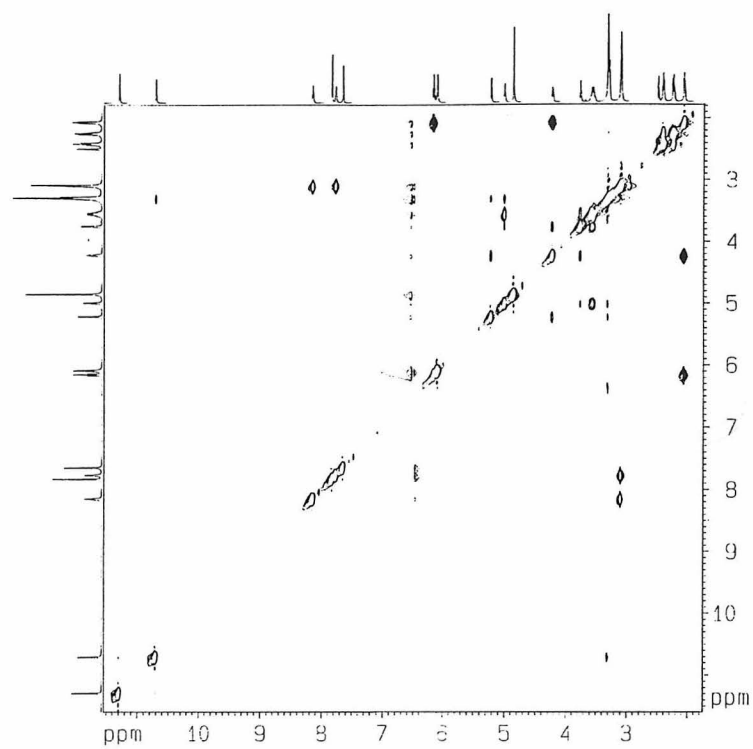
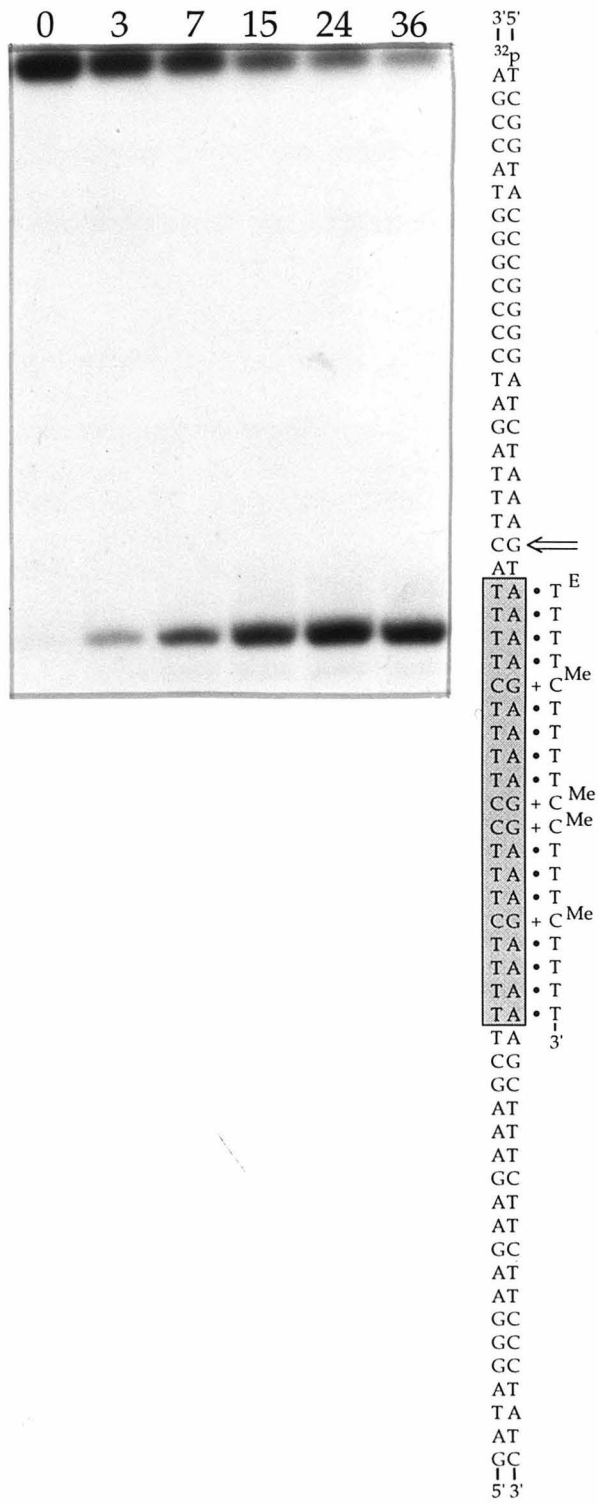


Figure 1.10. Autoradiogram of a 15% denaturing polyacrylamide gel to demonstrate the specificity of alkylation in released-product analysis experiments. N-bromoacetyloligonucleotide (2 μ M) incubated with the 5'-³²P-end labeled duplex with unlabeled duplex (3 μ M), 20 mM HEPES pH 7.0, 0.2 mM cobalt hexammine trichloride at 37° C alkylates the targeted guanine base in the duplex by triple helix formation. Lanes are annotated with the time points used for the experiment.



conjugate (T-G) occurs (Figure 1.11). Control experiments with undigested alkylation reaction products showed this not to be due to products of enzyme digestion. Based on the hypothesis that this was due to structure formation, HPLC conditions were changed to favor denaturation. Upon raising the column temperature to 50°C, this resulted in a clean HPLC trace showing the presence of the guanine-(linker)-thymidine compound in the enzyme digest.

A chromatogram of the reversed phase HPLC analysis shows two classes of products resulting from enzyme treatment (Figure 1.12). Nucleotides, which elute early in the chromatogram, and nucleoside, which elute later and represent a small proportion of the total amount of products. Oligonucleotides comprising the duplex target DNA were only partially digested under the conditions of the experiment. The peak at 27 min is 2'-deoxyadenosine-5'-monophosphate resulting from partial enzymatic hydrolysis of the duplex DNA and the amount of this product remains approximately constant throughout the experiment.

An expansion of the "nucleoside" region of the reversed phase HPLC analysis of the reaction products is shown in Figure 1.13. The authentic guanine-(linker)-thymidine standard (**T-G**) expected from the alkylation reaction elutes at 48.5 minutes under these HPLC conditions (panel a, Figure 1.13). Featured in the time course (panels b-d, Figure 1.13) is the presence of a compound with a retention time of 47.8 minutes (**T-G**, Figure 1.13). The relative amount of this product increases during the course of the reaction and comigrates upon coinjection with the authentic standard **T-G** (data not shown). The amount of 5-methyl-2'-deoxycytidine 5'-monophosphate (retention time 11 min) remains constant over time and provides an internal standard for relative peak area comparison. Upon superimposition and normalization of the UV spectra of peak **T-G** and the authentic

Figure 1.11. Conditions favoring denaturation are necessary for HPLC analysis for alkylation products. (A) Reversed phase HPLC analysis at room temperature of a three nmol sample of digested alkylation products. (B) HPLC analysis of a three nmol sample of undigested alkylation products gives a number of peaks at 45-60 min, obscuring products in the resulting digest. (C) HPLC analysis at 50°C of digested alkylation products removes these peaks, consistent with the denaturation of structure formed by undigested oligonucleotides.

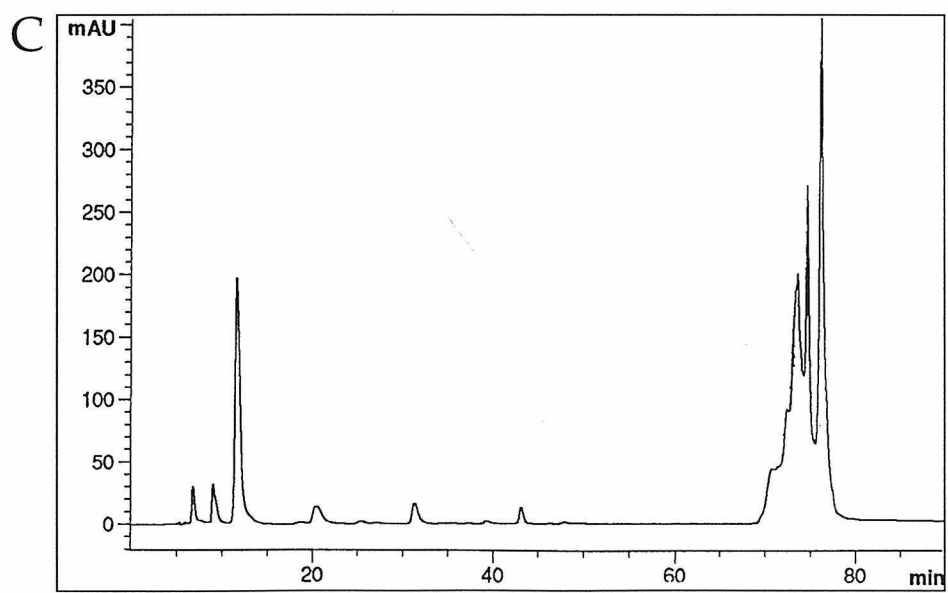
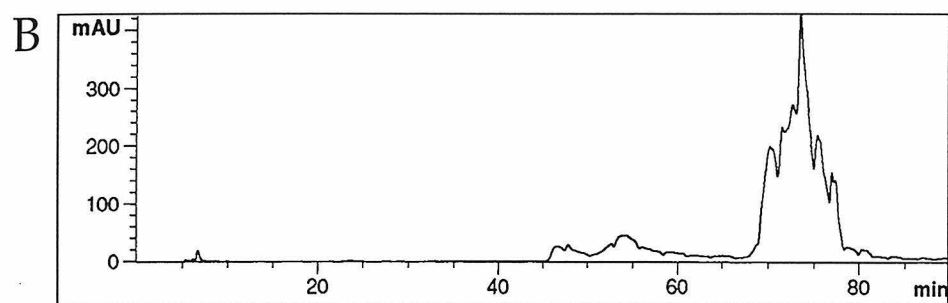
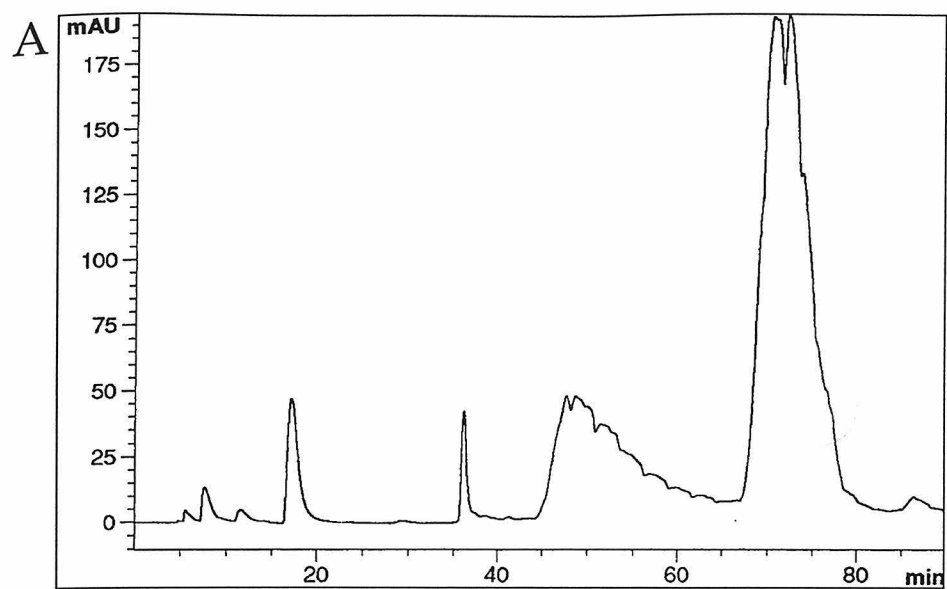


Figure 1.12. HPLC analysis of enzymatically digested products of oligonucleotide-directed alkylation give three sets of products - nucleotides, nucleosides, and undigested products from the duplex substrate. 5-methyl-2'-deoxycytidine elutes at 11 min and provides a well resolved internal standard for the comparison of peak area ratios.

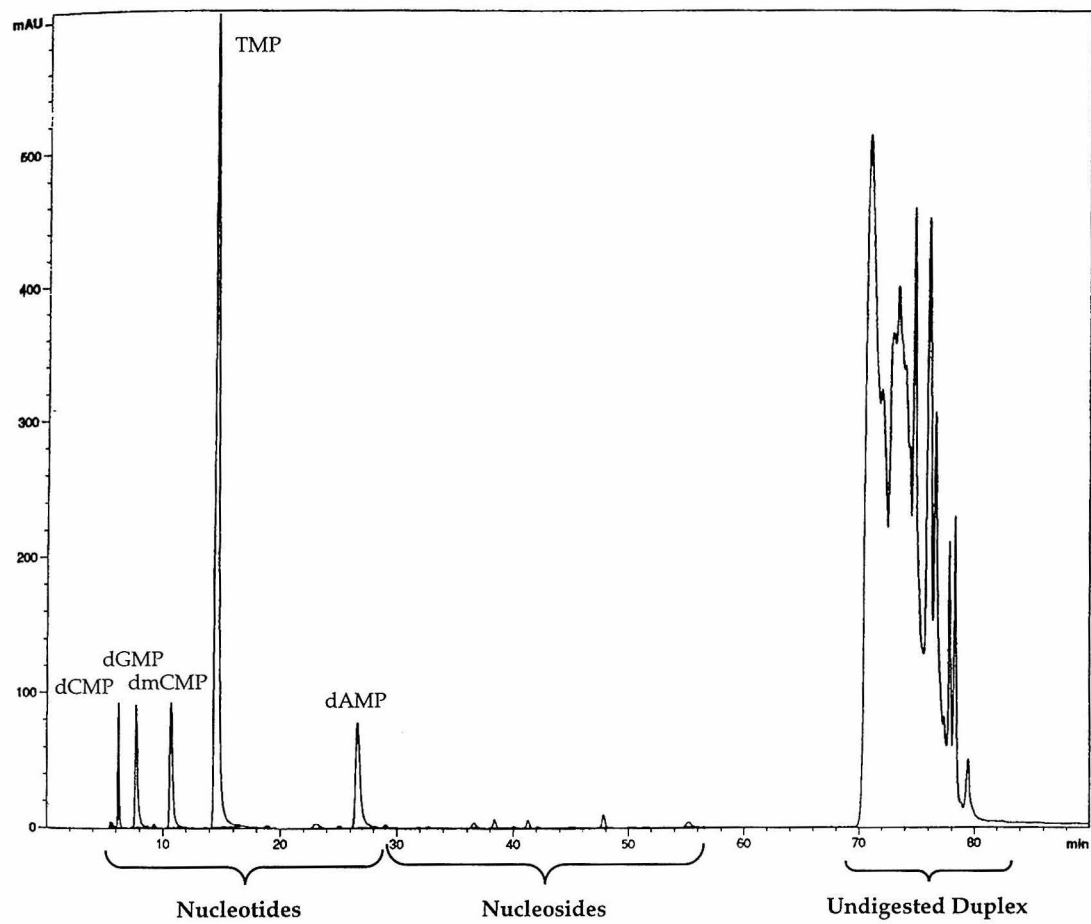
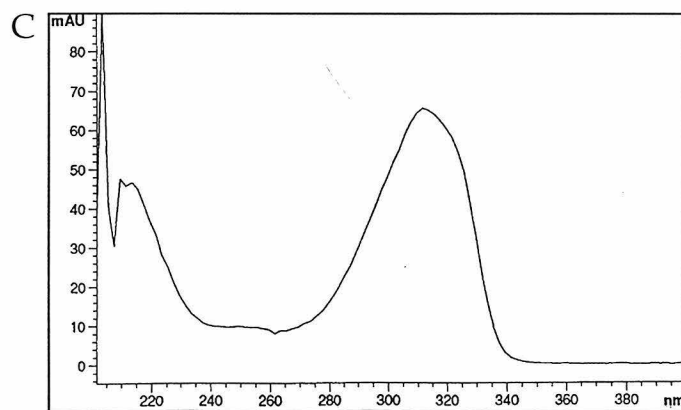
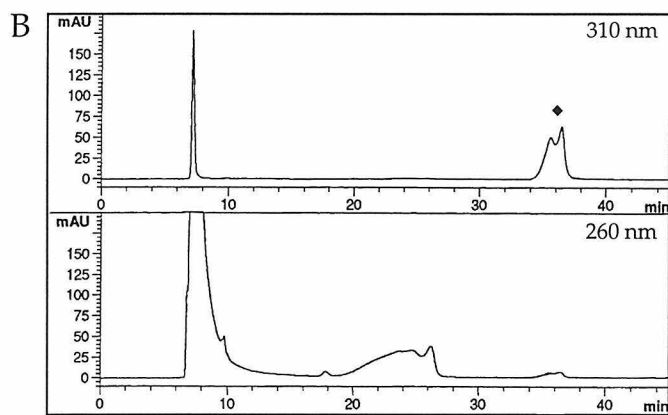
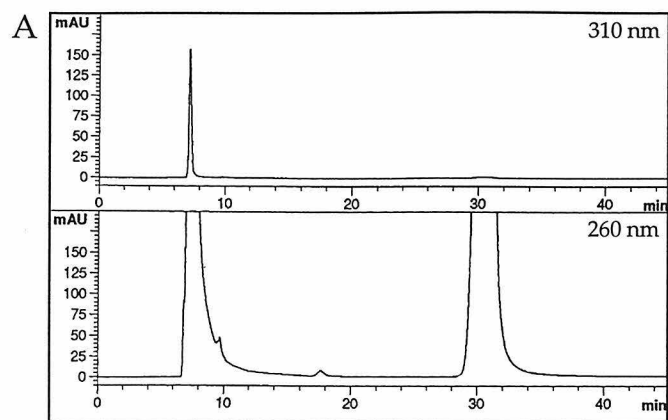


Figure 1.13. HPLC analysis of the enzymatically digested products of oligonucleotide-directed alkylation of double stranded DNA. Peak heights for the time course are normalized according to peak area ratio with 5-methyl-2'-deoxycytidine 5'-monophosphate and chromatograms were monitored at 260 nm: panel a, HPLC trace of authentic guanine-(linker)-thymidine standard T-G; panels b-d, HPLC traces of enzyme digests of oligonucleotides (3 nmol) from an alkylation reaction at times 3, 10.5, and 29 hours; panel e, UV spectrum of the peak at 47.8 minutes in panel d superimposed and normalized with the UV spectrum of authentic standard T-G in panel a. Labels for peak assignments are: T-OH for hydroxy-(linker)-thymidine, T-G for guanine-(linker)-thymidine, and T-Br for bromo-(linker)-thymidine nucleoside.

standard, they are shown to be identical (panel e, Figure 1.13). These data indicate that the covalent bond formation occurs at the N7 position of guanine proximal to the triple-helical complex. Two additional features of the time course are the presence of peaks consistent with retention times of the starting material (55.2 min, **T-Br**, Figure 1.13) and the nucleoside formed by hydrolysis (38.4 min, **T-OH**, Figure 1.13). The relative amounts of **T-Br** and **T-OH** decrease over time, consistent with conversion in the alkylation reaction of N-bromoacetyloligonucleotide to guanine-(linker)-oligonucleotide conjugate. As the amount of **T-OH** follows the disappearance of **T-Br**, the presence of the former product likely results from hydrolysis during the enzymatic digest (2 hours, pH 8.5) of starting material remaining from the alkylation reaction. This reaction is also seen to occur in control digests of the N-bromoacetyloligonucleotide.

An analysis of the data at alternate wavelengths (>300 nm) indicated the presence of a peak in the HPLC spectrum with a maximum absorbance of 310 nm. Control experiments showed that this product arises from treatment of the aminooligonucleotide with the activated ester of bromoacetic acid, specifically from reaction at 5-methyl-2'-deoxycytidine residues within the oligonucleotide (Figure 1.14). A survey of the literature revealed that reactions of this type are precededented (15). Nucleic acids treated with alpha-haloacetaldehydes results in the formation of fluorescent products by reaction of these bifunctional reagents with cytosine and adenine. In analogy to this, a side product could be formed as in Figure 1.15. UV-visible absorbance characteristics are consistent. Based on the molar extinction coefficient, it was estimated that this side product represent *ca.* 20% of the total material. Simply reducing the amount and the reaction time in the acylation of the aminooligonucleotide eliminated the formation of this product.

Figure 1.14. Control experiments with thymidine and 5-methyl-2'-deoxycytidine and the N-hydroxysuccinimidyl ester of 2-bromoacetic acid demonstrate that reaction of the latter nucleoside with the bifunctionalized ester results in formation of a product with λ_{max} at 310 nm.



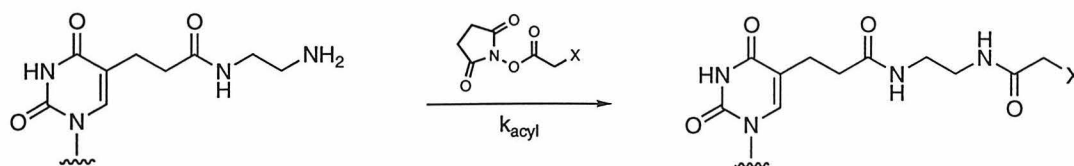
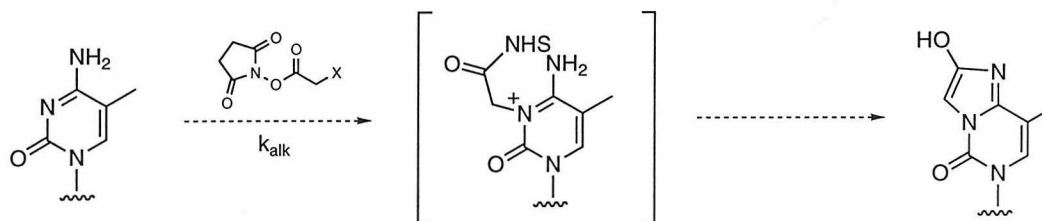
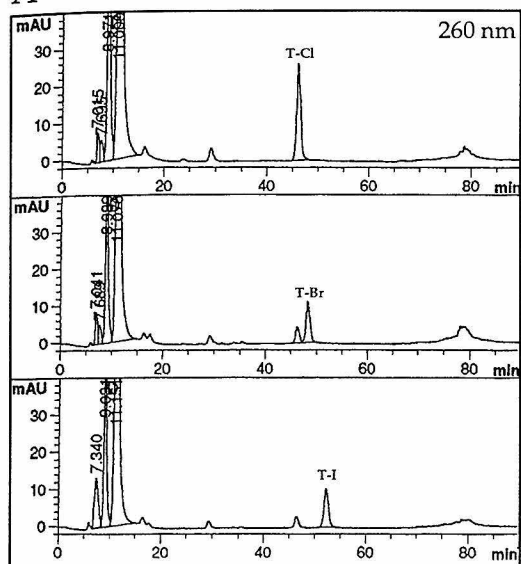
Desired reaction at 5'-terminal residue**Possible side reaction at internal 5-methyl-dC residues**

Figure 1.15. Reaction of the activated ester of bromoacetic acid with the aminooligonucleotide results in formation of a small amount of a side product.

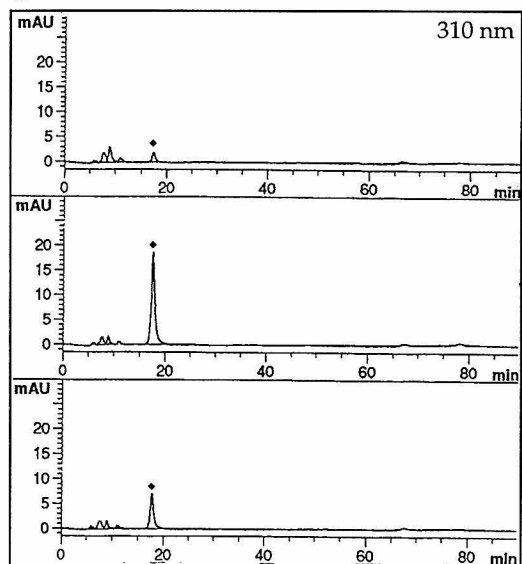
A result from the work of Povsic and Dervan in this alkylation system demonstrating the rate differences for chloro-, bromo-, and iodo-acetyl as the electrophile directed by triple helix formation remained unexplained. The relative rates were 0.02, 1.0, and 0.2, respectively. According to the normal rules for nucleophilic substitution reactions, this trend is anomalous (16). The presence of this side product, if present in varying amounts due to treatment of the aminooligonucleotide with the different activated esters, could explain this trend as a variation in purity of starting material. Aminooligonucleotide was reacted with chloro-, bromo-, and iodo-acetate activated esters, digested with snake venom phosphodiesterase, and analyzed by HPLC (Figure 1.16). The relative ratios of the side product corresponded roughly to that seen for triple helix mediated alkylation. Though no literature is available on the intrinsic reactivity of alpha-haloamides to nucleophilic

Figure 1.16. Snake venom phosphodiesterase digestion and HPLC analysis of the products of reaction of N-hydroxysuccinimidyl 2-chloro-, 2-bromo-, and 2-iodoacetate with the oligonucleotide $\text{NH}_2\text{-TTTT}^{\text{Me}}\text{CTTTT}^{\text{Me}}\text{CTTT}^{\text{Me}}\text{C}^{\text{Me}}\text{CTTTT-3'}$. (A) Monitoring at 260 nm reveals the presence of each 5'-end modification in the oligonucleotide. (B) Monitoring at 310 nm reveals the presence of the putative reaction product of activated ester with 5-methyl-2'-deoxycytidine in the oligonucleotide.

A



B



substitution, it seems likely that the observed trend reflects the inherent relative reactivity of these electrophiles to nucleophilic substitution.

Section 3. It has been demonstrated that localization of bromoacetyl by triple helix formation on duplex DNA achieves alkylation at one guanine base. It is yet to be determined what effect localization of bromoacetyl has on the rate of the alkylation reaction at N7 of guanine in duplex DNA. To determine the rate enhancement obtained upon tethering of the bromoacetyl moiety to the triple-helical complex, the rate of the intramolecular reaction by the N-bromoacetyloligonucleotide at a single guanine base in duplex DNA is compared to the rate of intermolecular alkylation by 2-bromoacetamide at the same guanine base.

In preparation for experiments involving relatively large concentrations of 2-bromoacetamide for purposes of measuring intermolecular rate constants, control experiments were performed to determine the stability of high concentrations (*ca.* 50 mM) 2-bromoacetamide in the presence of all components in the reaction mixture except DNA.

A preliminary experiment involved the examination of pH versus time was performed (Figure 1.17). Initial conditions for the alkylation reaction were based on previously published conditions (20 mM HEPES pH 7.4, 0.8 mM cobalt hexammine trichloride, 37° C), and under these conditions the release of hydrobromic acid causes a drop in pH at relevant time ranges and concentration 2-bromoacetamide.

A more detailed analysis in these control experiments involved an HPLC analysis of 2-bromoacetamide stability. Reversed phase HPLC chromatography, monitored at 214 nm, proved to be effective in following the stability of the electrophile under the reaction conditions. All components and potential products are resolved by this technique (Figure 1.18). Experiments demonstrated that substantial degradation of 2-bromoacetamide occurs,

Figure 1.17. A plot of pH versus time demonstrates a drop in pH in the time range of the kinetics experiments performed in HEPES buffer (20 mM HEPES pH 7.4, 0.8 mM cobalt hexammine trichloride).

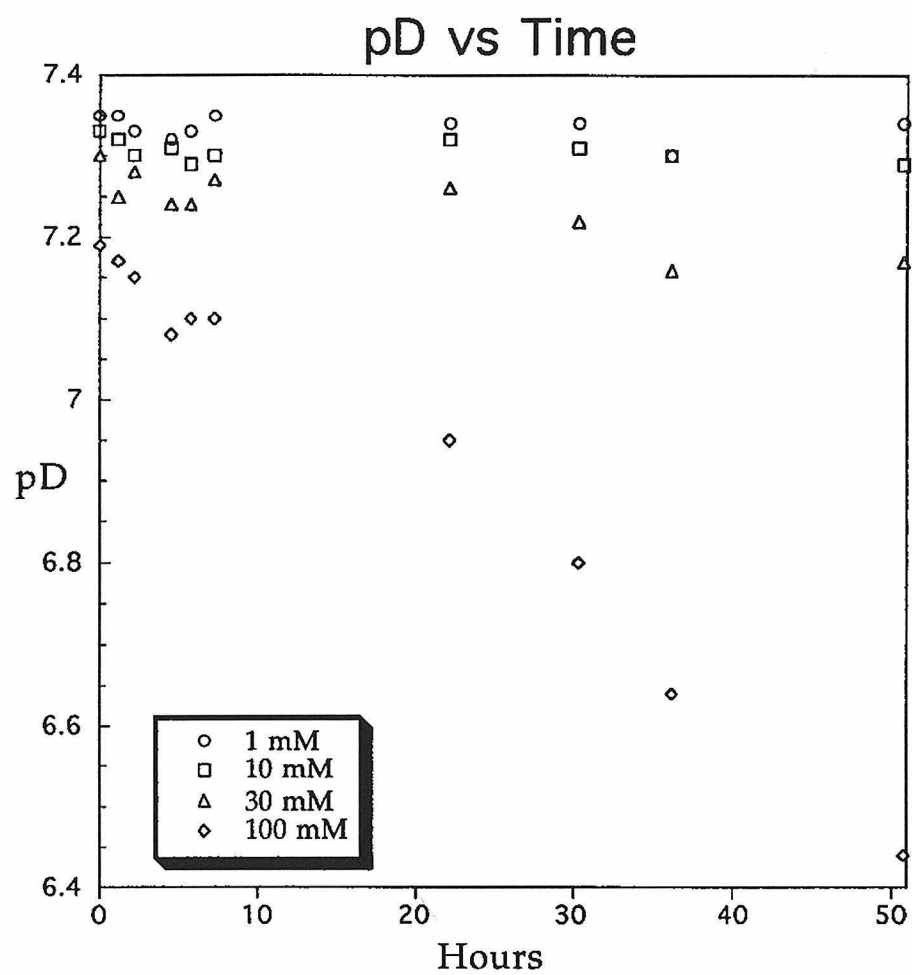
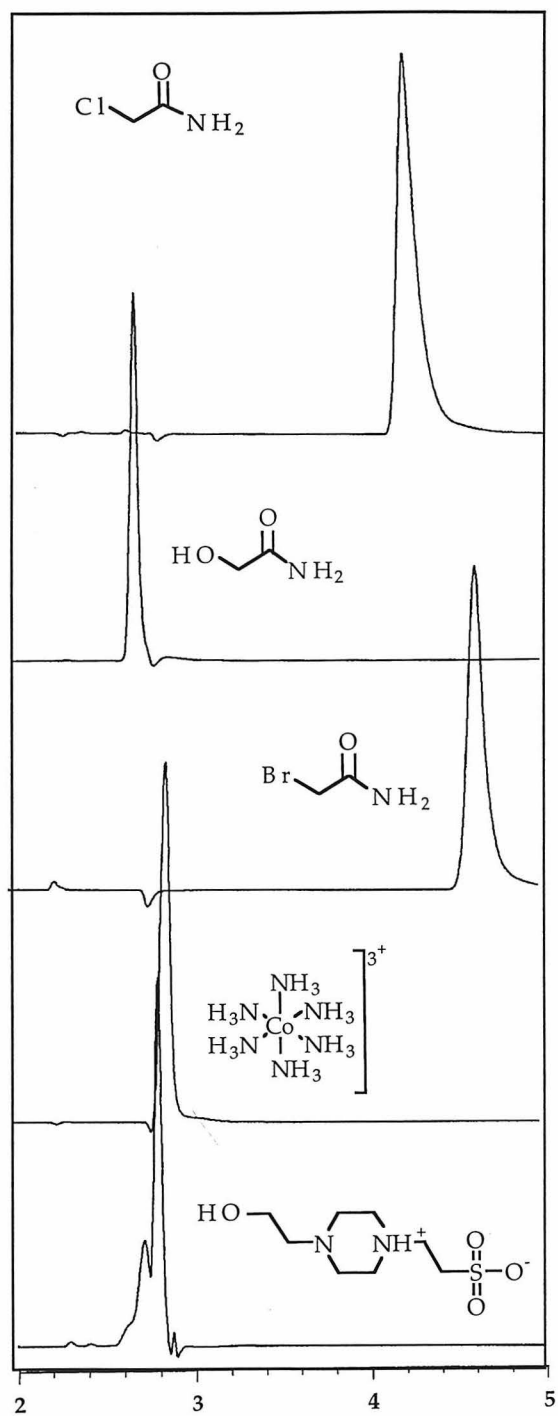


Figure 1.18. Reversed phase HPLC analysis of the starting materials and potential products in control experiments for 2-bromoacetamide stability.



consistent with the pH drop observed previously (Figure 1.19). Interestingly, no hydrolysis (resulting in the formation of glycolamide) occurs in the reaction - further controls demonstrated that reaction with HEPES is the sole degradation pathway within the limit of detection by HPLC analysis. To find conditions where the electrophile would have suitable stability, other buffers were tried. Sodium phosphate, like HEPES, was also found to be unacceptable (Figure 1.20), but Bis Tris Acetate exhibited very little reaction with 2-bromoacetamide (Figure 1.21). Because a small drop in pH was observed with this buffering system at pH 7.4, the pH was lowered to 7.0 where control experiments showed that pH was stable through long reaction times.

The equilibrium complexation of an oligonucleotide appended with an electrophile (O^E) with duplex DNA (D) containing a target site for local triple helix formation, followed by alkylation of the bound duplex to form alkylated product (P) is described by



where K represents the equilibrium association constant and k_1 the first-order rate constant for alkylation at the N7 position of one guanine base within the bound complex. The rate of disappearance of duplex DNA is described by the following first-order rate expression: $-\ln([D]_t/[D]_0) = k_1 t$, and will be at a maximum when the duplex target site is saturated with N-bromoacetyl oligonucleotide. Values for these parameters can be determined experimentally by quantitation of bands from a polyacrylamide gel of alkylation products.

The N-bromoacetyl oligonucleotide of sequence composition 5'- E TTTT^{Me}CTTTT^{Me}C^{Me}CTTT^{Me}CTTTT-3' is capable of site-specific alkylation of a guanine

Figure 1.19. HPLC traces representing a time course to determine the stability of 10 mM 2-bromoacetamide in the presence of 180 mM HEPES buffer (pH 7.4) and 0.8 mM cobalt hexammine trichloride. 2-Bromoacetamide is annotated by a diamond and chromatograms were monitored at 214 nm.

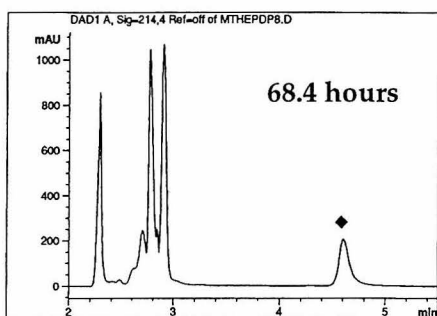
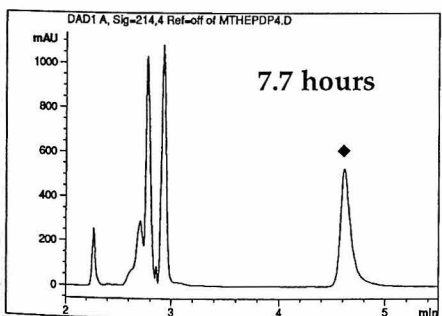
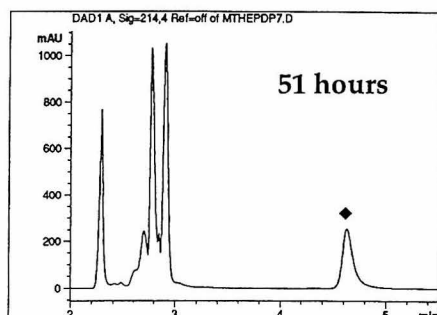
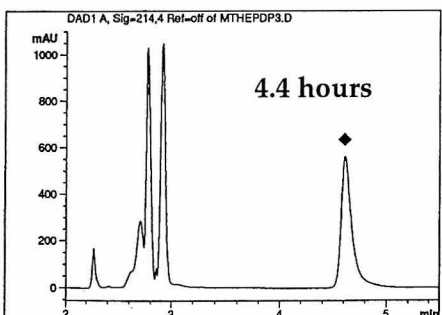
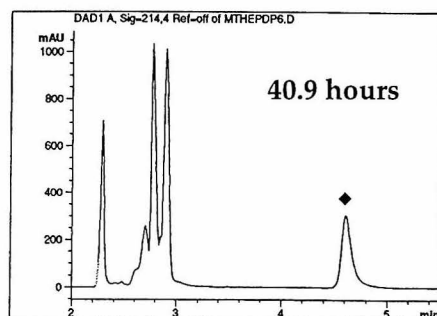
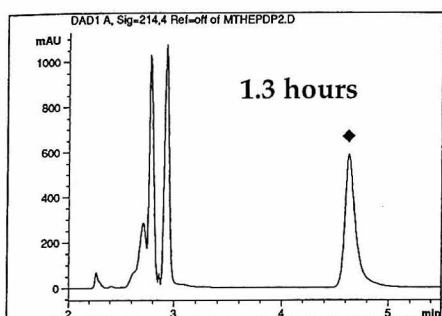
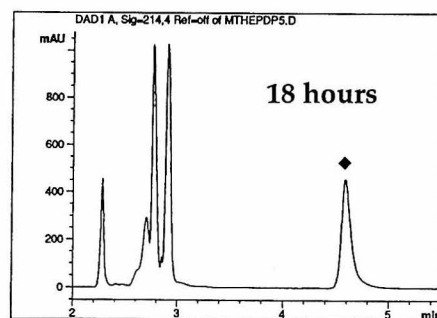
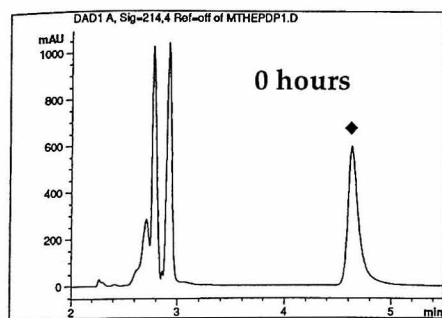


Figure 1.20. HPLC traces representing a time course to determine the stability of 10 mM 2-bromoacetamide in the presence of 180 mM phosphate buffer (pH 7.4) and 0.8 mM cobalt hexammine trichloride. 2-Bromoacetamide is annotated by a diamond and chromatograms were monitored at 214 nm.

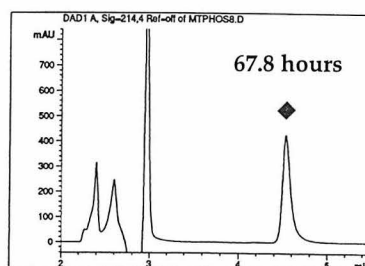
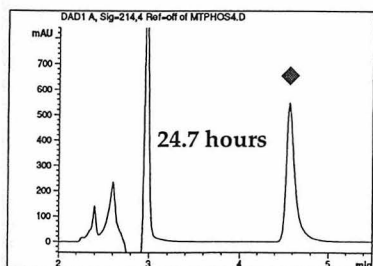
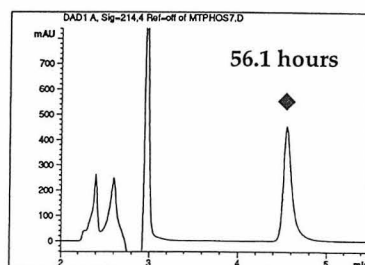
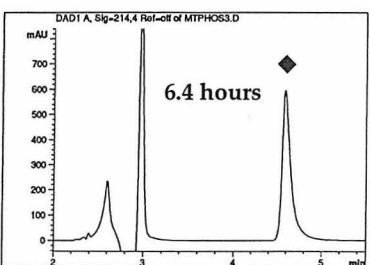
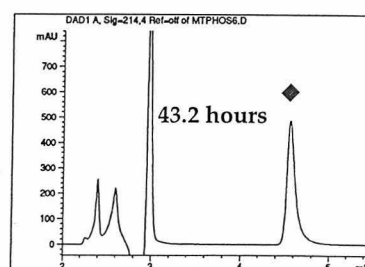
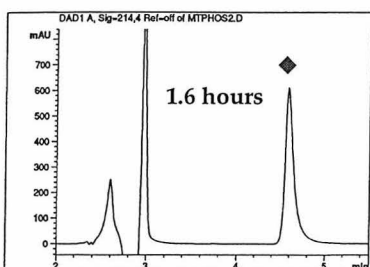
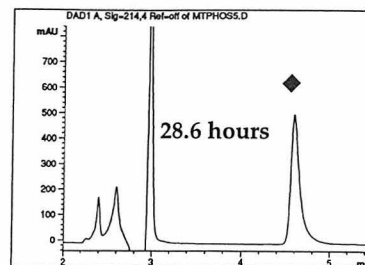
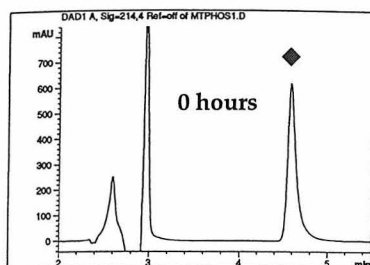
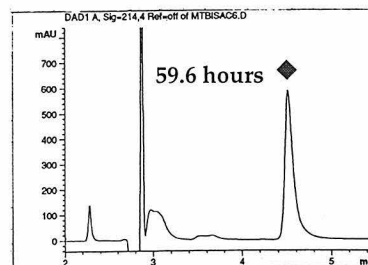
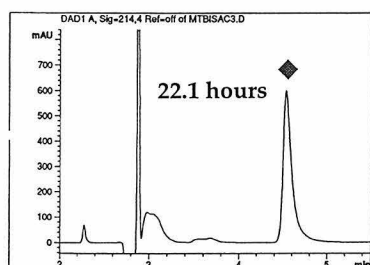
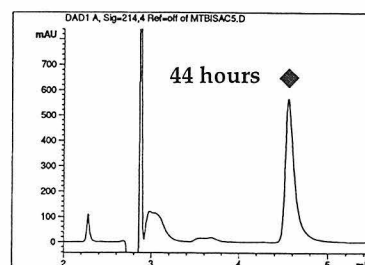
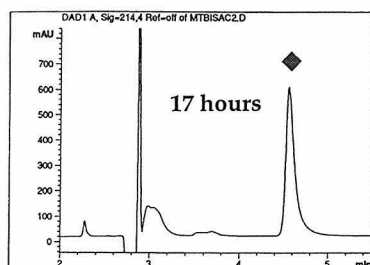
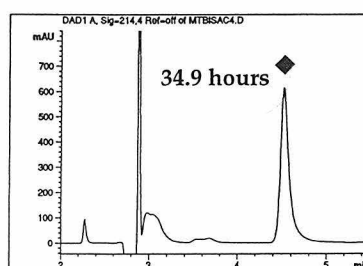
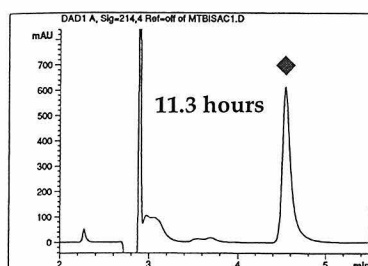
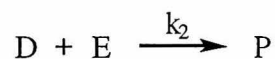


Figure 1.21. HPLC traces representing a time course to determine the stability of 10 mM 2-bromoacetamide in the presence of 180 mM Bis Tris Acetate buffer (pH 7.4) and 0.8 mM cobalt hexammine trichloride. 2-Bromoacetamide is annotated by a diamond and chromatograms were monitored at 214 nm.



base two base pairs to the 5'-side of the homopurine target site within the 666 base pair fragment of double stranded DNA derived from plasmid pUCINET (Figure 1.22). A titration shows that under these conditions (20 mM Bis-Tris Acetate, pH 7.0, 0.8 mM cobalt hexammine trichloride, 37° C) the rate is maximal down to 200 nM concentration N-bromoacetyloligonucleotide (Figures 1.23 and 1.24, Table 1.1). A plot of $\ln ([D]_i/[D]_f)$ versus time is linear, indicating that the reaction is first-order in DNA target concentration, and provides a rate constant k_1 at 200 nM of $2.7 (\pm 0.5) \times 10^{-5} \text{ s}^{-1}$ ($t_{1/2} = 7.2$ hours) for the alkylation reaction within the bound complex (Figure 1.25) (17). The plot deviates from linearity beyond two half-lives, likely due to the similarity of the rate constants for alkylation and dissociation of the Hoogsteen strand of the triplex (4e).

For determination of the rate constant of 2-bromoacetamide alkylation at the N7 position of guanine within duplex DNA an alkylation reaction on double stranded DNA (D) by 2-bromoacetamide (E) to form product (P) can be described by



where k_2 represents the second-order rate constant for alkylation at the N7 position of one guanine base G in the duplex. The rate of disappearance of duplex containing intact guanine base G is then described by the following second-order rate expression: $-d[D]/dt = k_2[E][D]$. Formally, the rate of disappearance of fully intact duplex DNA is a sum of the rates of reaction at all guanine bases in the duplex. What is desired in this analysis, however, is the rate of alkylation at guanine base G. For a determination of the rate constant

Figure 1.22. Plasmid pUCINET was constructed by ligation of a synthetic duplex into pUC19 previously digested with Bam HI and Sal I, from which is produced a 3'-³²P-end labeled Hind III/Ssp I fragment.

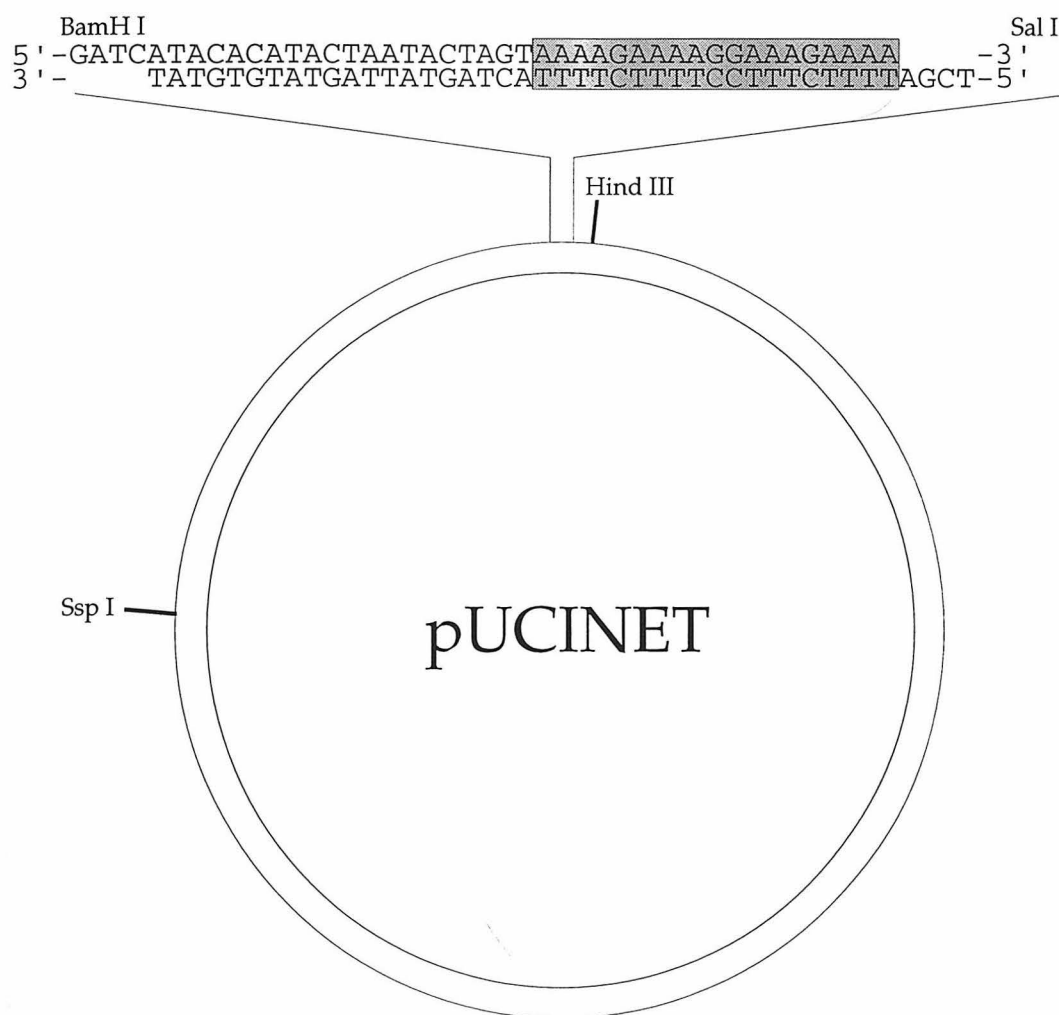


Figure 1.23. Gray-scale representation of an 8% polyacrylamide gel used to analyze the rates of different concentrations of N-bromoacetyloligonucleotide alkylation of the 3'-³²P-end labeled 666 base pair fragment of duplex DNA. Lanes are annotated with concentrations N-bromoacetyloligonucleotide of 1 μ M, 200 nM, and 100 nM N-bromoacetyloligonucleotide. Lanes 1, 8, and 15 - 0 hour; lanes 2, 9, and 16 - 1 hour; lanes 3, 10, and 17 - 3 hours; lanes 4, 11, and 18 - 5 hours; lanes 5, 12, and 19 - 8.8 hours; lanes 6, 13, and 20 - 12.7 hours; lanes 7, 14, and 21 - 27 hours.

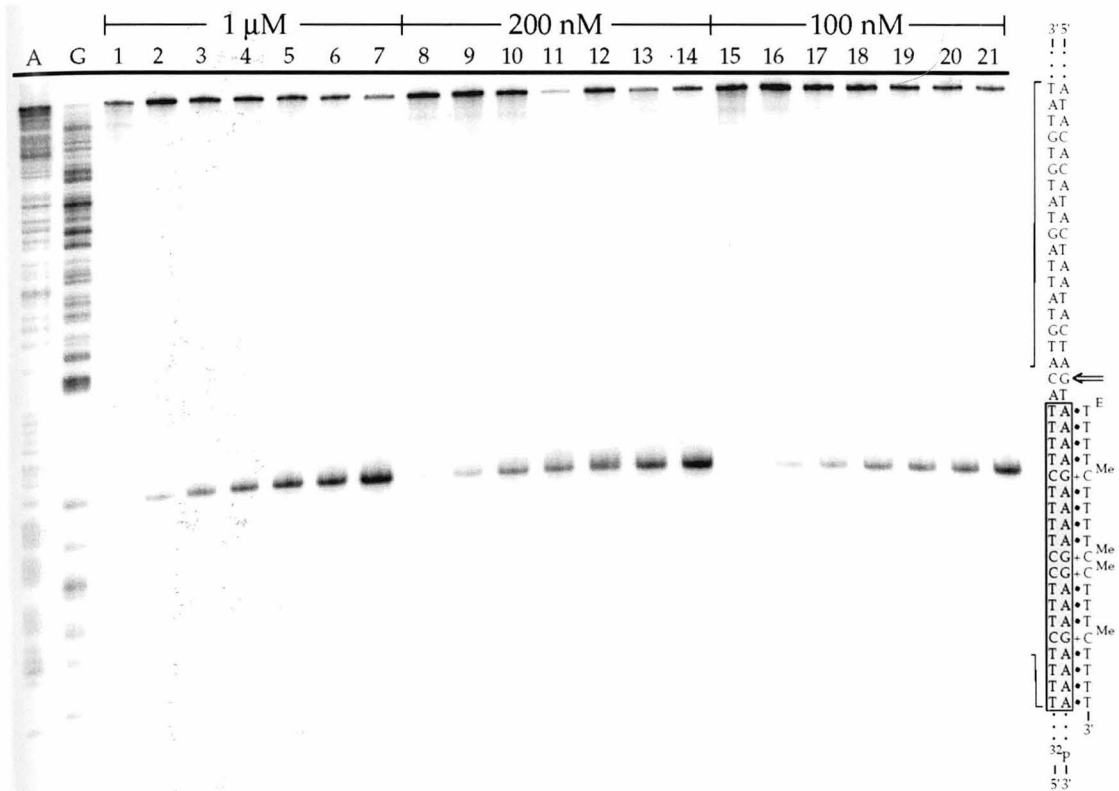


Figure 1.24. Autoradiogram of an 8% denaturing polyacrylamide gel used to analyze the products of oligonucleotide-directed alkylation and base promoted strand cleavage within the 3'-³²P-end labeled 666 base pair restriction fragment: lane 1, control lane - 24 hours with no N-bromoacetyloligonucleotide; lane 2, A-specific sequencing reaction; lane 3, G-specific sequencing reaction; lanes 4-10, reaction of 200 nM N-bromoacetyloligonucleotide with duplex DNA at 37° C at times 0, 1, 3, 5, 8.8, 12.7, and 27 hours.

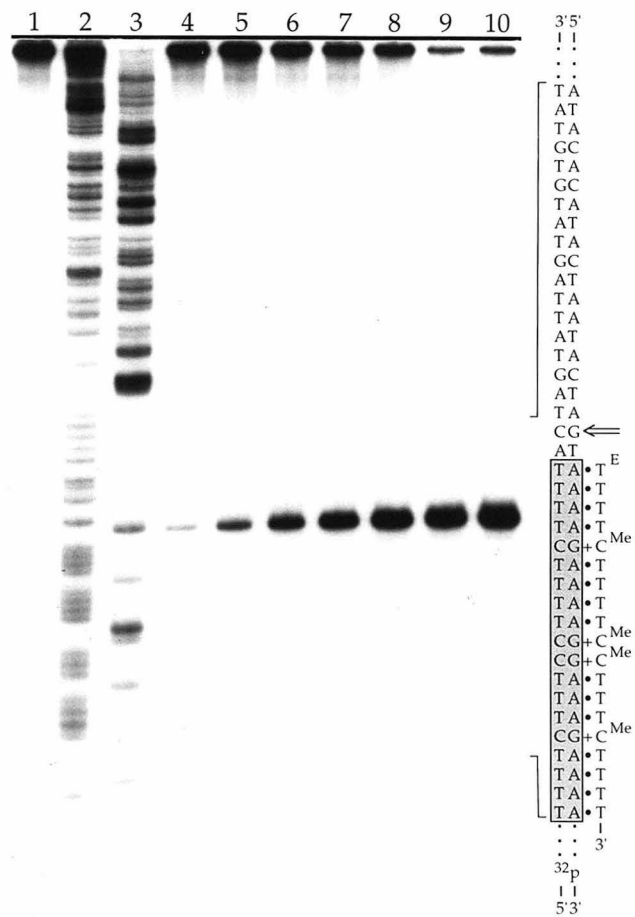
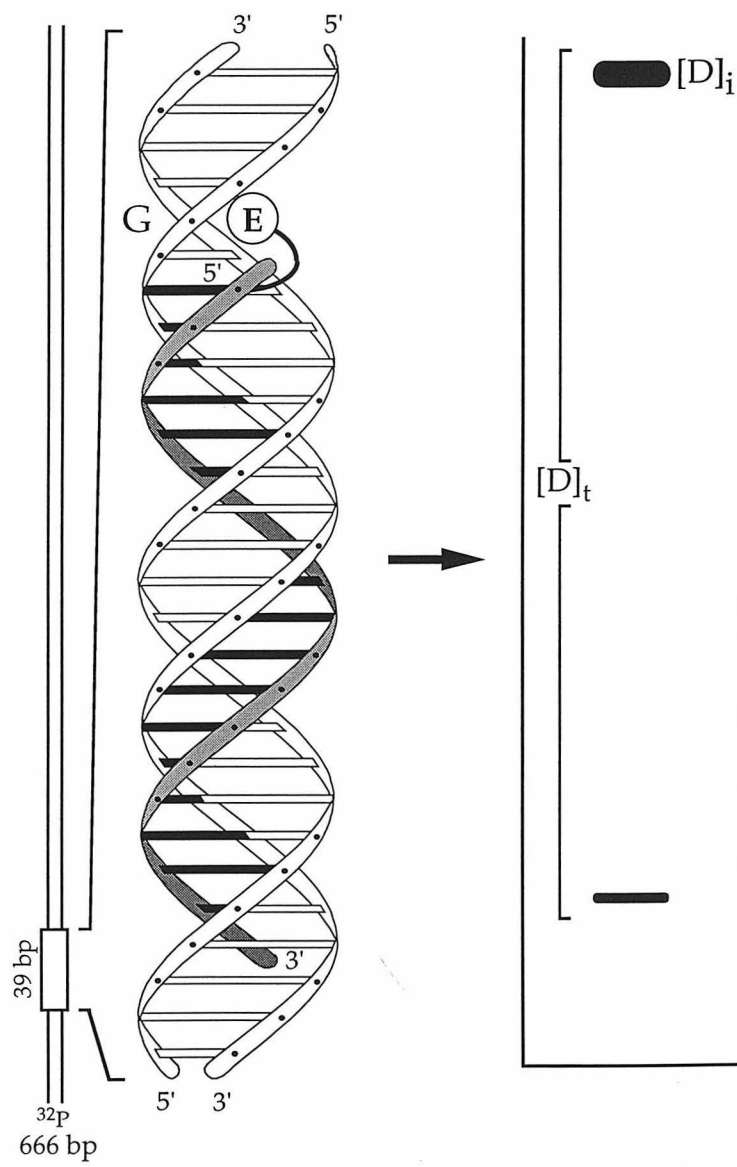


Figure 1.25. Determination of the rate constant k_1 for oligonucleotide-directed alkylation at the N7 position of guanine within duplex DNA. Data is quantitated as shown in the accompanying graphic. Shown is a plot of $\ln ([D]_i/[D]_t)$ versus time at 200 nM N-bromoacetyloligonucleotide concentration, which provides from an average of four determinations from the slope of the straight line a rate constant k_1 of $2.7 (\pm 0.5) \times 10^{-5} \text{ s}^{-1}$.



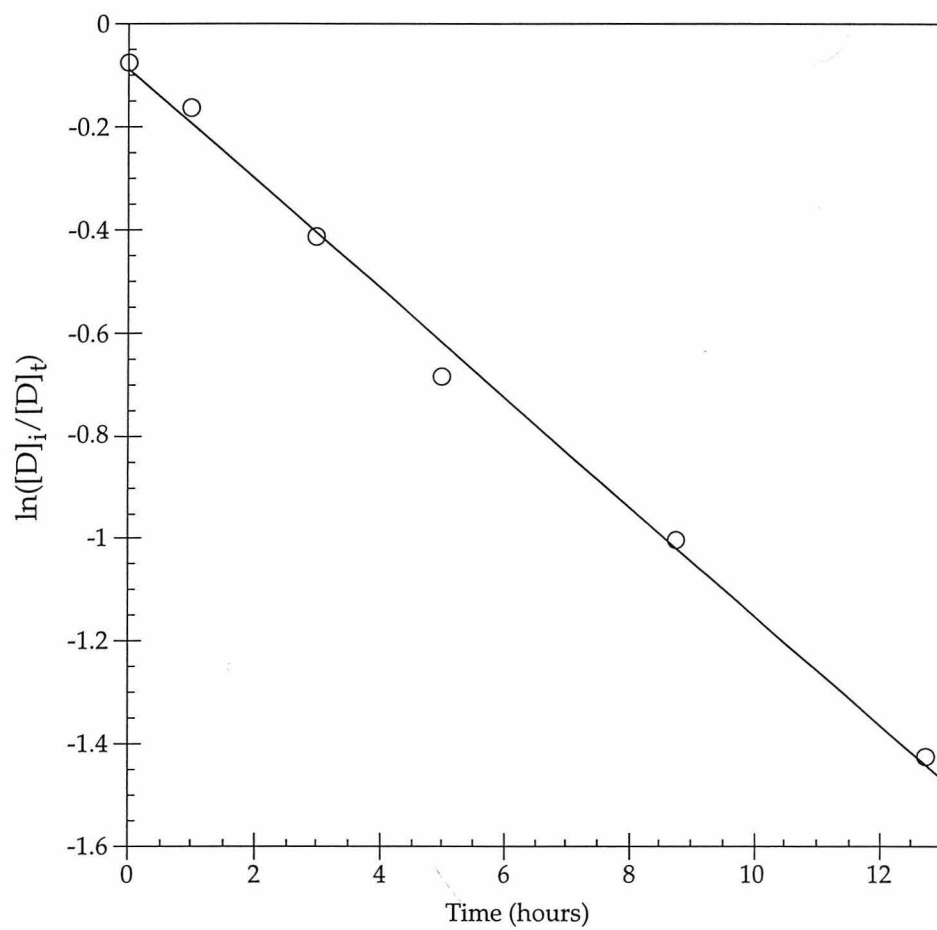


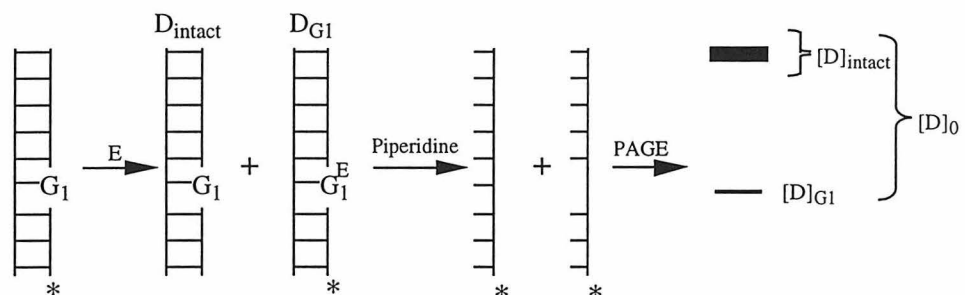
Table 1.1. Rates of Alkylation^a

Concentration (μM)	k_1 (s^{-1}) ^b	$t_{1/2}$ (hr)
1	$2.4 (\pm 0.3) \times 10^{-5}$	$8.0 (\pm 0.8)$
0.2	$2.7 (\pm 0.5) \times 10^{-5}$	$7.2 (\pm 1.5)$
0.1	$1.6 (\pm 0.4) \times 10^{-5}$	$12.0 (\pm 3.4)$

^aError is reported as standard deviation about the mean.^bValue reported for 100 nM concentration is not a true first-order rate constant.

at one position in duplex DNA for a reaction occurring simultaneously at multiple positions, an analysis of the system is required to extract the desired rate constant.

Consider an alkylation reaction by an electrophile on a radiolabeled duplex containing one guanine-cytosine base pair (Figure 1.26).

**Figure 1.26.** Schematic of alkylation of a duplex with one GC base pair.

The rate of disappearance of the intact duplex still containing the guanine base G₁ (D_{intact}) is proportional to the concentration of electrophile and the concentration of intact duplex remaining:

$$-\frac{d[D]_{\text{intact}}}{dt} \propto [E][D]_{\text{intact}}$$

If a proportionality constant k is included,

$$-\frac{d[D]_{\text{intact}}}{dt} = k[E][D]_{\text{intact}} \quad (1)$$

where k is then the rate constant for reaction of the electrophile at the guanine base in the duplex. Assuming the concentration of the electrophile ($[E]$) remains constant throughout the experiment (when $[E] \gg [D]$ and the electrophile does not undergo side reactions), this differential equation can be solved to obtain an equation relating the concentration of intact DNA and the time of reaction:

$$-\frac{d[D]_{\text{intact}}}{dt} = k[E][D]_{\text{intact}}$$

$$\int \frac{d[D]_{\text{intact}}}{[D]_{\text{intact}}} = -\int k[E]dt, \quad \int \frac{dx}{x} = \ln x$$

$$\ln [D]_{\text{intact}} = -k[E]t - C, \quad \text{when } t=0, C = -\ln [D]_0$$

$$\ln [D]_{\text{intact}} = -k[E]t + \ln [D]_0$$

$$\ln [D]_{\text{intact}} - \ln [D]_0 = -k[E]t, \quad \ln x - \ln y = \ln \frac{x}{y}$$

$$\ln \frac{[D]_{\text{intact}}}{[D]_0} = -k[E]t \quad (2)$$

Quantitation of values for $[D]_{\text{intact}}$ and $[D]_0$ from a polyacrylamide gel and a plot of $-\ln ([D]_{\text{intact}}/[D]_0)$ versus time under pseudo first-order conditions ($[E] \gg [D]$) will then provide the value of $k[E]$ from the slope of the resulting straight line (Figure 1.27). If the concentration of the electrophile is known, the value for the bimolecular rate constant k can then be determined.

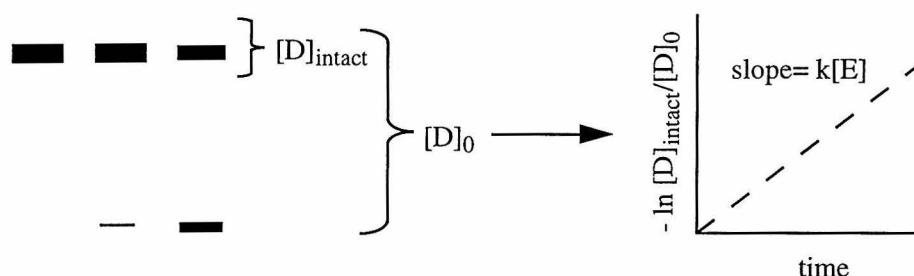


Figure 1.27. Quantitation provides a rate constant for alkylation.

For a longer fragment of DNA containing many GC base pairs, " $[D]_{\text{intact}}$ " can be redefined as *a set of all duplex fragments still containing the intact guanine base G₁* (Figure 1.28). Assuming the rate constant for alkylation at G₁ is not affected by alkylation at other guanine bases in the duplex, treatment of the data using the new value for $[D]_{\text{intact}}$ provides the rate constant k .

If the reaction is performed under "single-hit" conditions, i. e. one alkylation event per duplex, intact DNA is then considered to be full length duplex as each monoalkylated product on the gel originated from this starting material. This simplifies the determination of rate constants at each site, and avoids the need to assume that reaction at one site does not affect the rate at other sites. In cases of where many sites of reactivity on duplex DNA

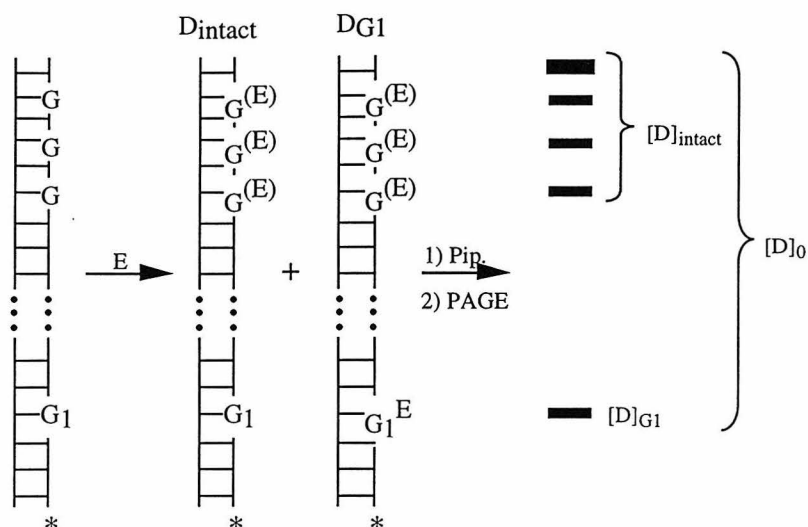


Figure 1.28. Experimentally distinguishable categories of reaction products also provide single rate constants for duplexes with multiple reactive sites.

exist, however, these "single-hit" experiments are not practical as the extent of reaction at each site will be very small at low total conversion.

Consider a duplex containing two GC base pairs (Figure 1.29). The reaction at any guanine x in the duplex is described by

$$-\frac{d[D]_{\text{intact}}}{dt} = k_x[E][D]_{\text{intact}}$$

where $[D]_{\text{intact}}$ is the concentration of duplex containing an intact guanine x and k_x is the rate constant for alkylation at guanine x . The rate of disappearance of the fully intact duplex (fragment a) is then

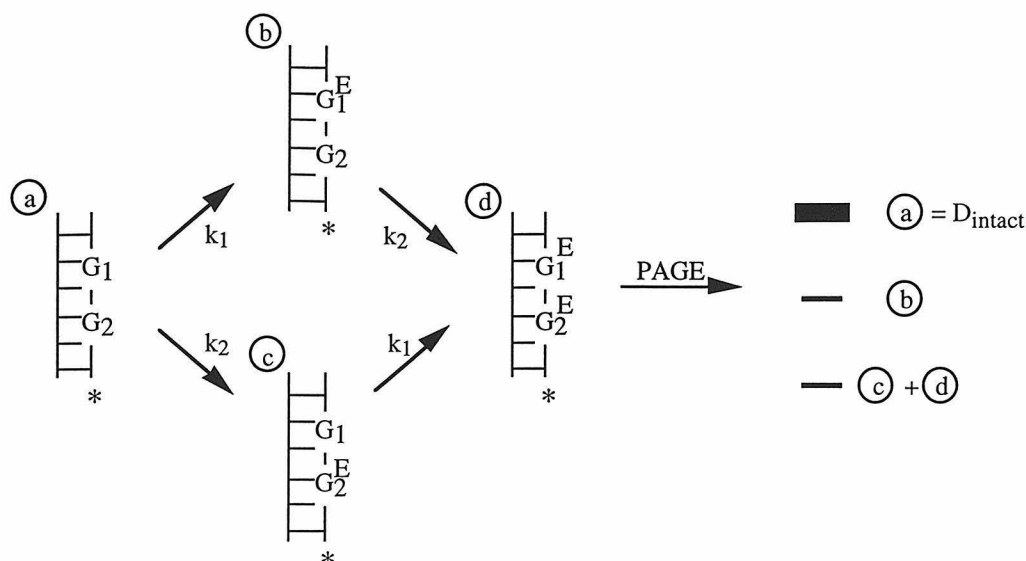


Figure 1.29. Schematic of alkylation of a duplex with two GC base pair.

$$-\frac{d[D]_{\text{intact}}}{dt} = k_1[E][D]_{\text{intact}} + k_2[E][D]_{\text{intact}}$$

By comparison to eq 1, this equation indicates that the duplex will disappear faster when there are two reactive sites than with one reactive site. Assuming that $[E]$, k_1 , and k_2 are constant throughout the experiment, this equation can be solved. These assumptions are true if $[E] \gg [D]$ (pseudo first-order conditions) and reaction at one site does not affect the rate of reaction at other sites in the duplex.

Solving as for equation 2 provides

$$\ln \frac{[D]_{\text{intact}}}{[D]_0} = -(k_1 + k_2)[E]t \quad (3)$$

Quantitation from a gel and plotting the data as in the above examples then provides the sum of rate constants k_1 and k_2 (Figure 1.30).

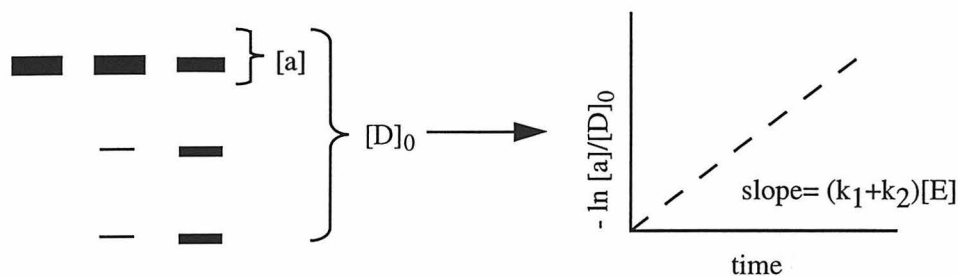


Figure 1.30. Quantitation provides a sum of rate constants for alkylation.

As discussed previously, intact duplex can be redefined as a set of longer fragments still containing G_1 and G_2 . In this example, eq 3 can then be used to provide the value $k_1 + k_2$ (Figure 1.31).

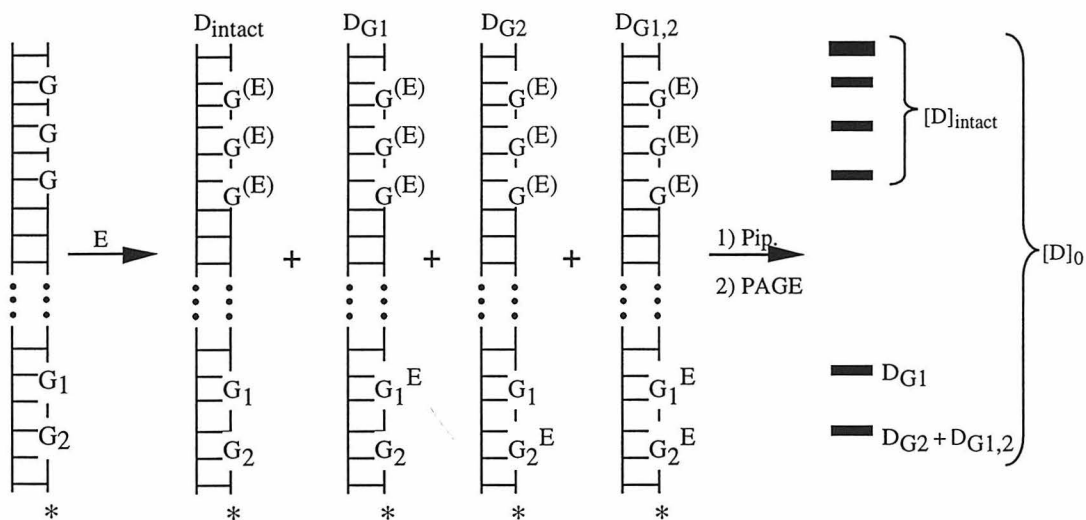


Figure 1.31. Quantitation of the data from a polyacrylamide gel gives a sum of rate constants for alkylation at guanines G_1 and G_2 .

This can then be extended and generalized for a duplex containing n guanine bases. In this case, intact DNA is defined as all bands G_1 to G_{x-1} . From eq 3, quantitation of the data in this way then provides *a sum of rate constants for reaction at all sites G_x to G_n* .

$$\ln \frac{[D]_{\text{intact}(1:x-1)}}{[D]_0} = -(k_x + k_{x+1} + \dots + k_n)[E]t$$

$$\ln \frac{[D]_{\text{intact}(1:x-1)}}{[D]_0} = - \sum_{i=x}^n k_i [E]t \quad (4)$$

The value for this sum of rate constants can be obtained from a plot (Figure 1.32).

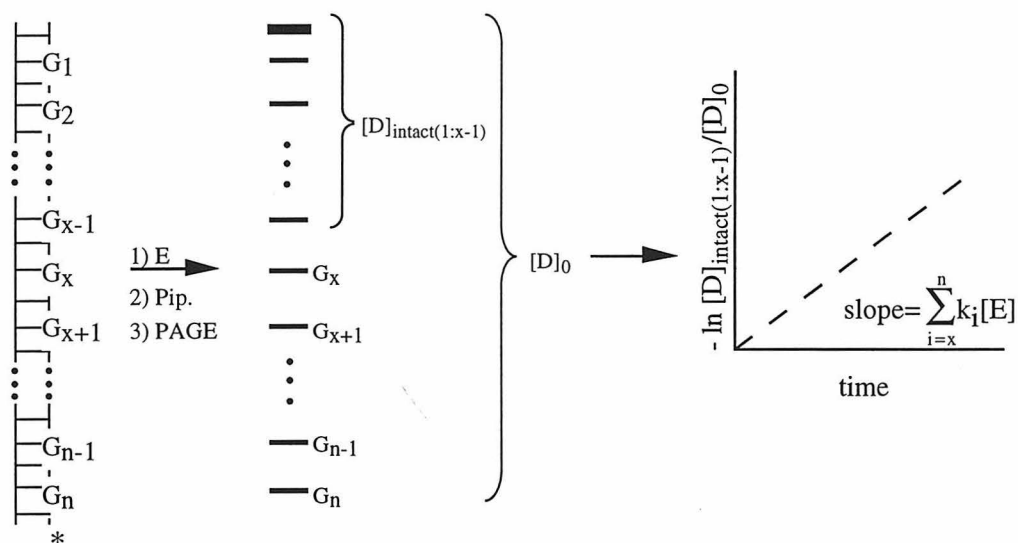


Figure 1.32. Quantitation of the data from a polyacrylamide gel gives a sum of rate constants for alkylation at individual positions in the duplex.

But what is desired in the experiment is *the rate constant k_x for alkylation at a single guanine base in the duplex*. k_x can be expressed in terms of values we can get, however - as the difference of two sums:

$$k_x = \sum_{i=x}^n k_i - \sum_{i=x+1}^n k_i$$

Multiplying both sides of the equation by $-[E]t$ gives:

$$-k_x[E]t = -\sum_{i=x}^n k_i[E]t + \sum_{i=x+1}^n k_i[E]t$$

From eq 4,

$$-k_x[E]t = \ln \frac{[D]_{\text{intact}(1:x-1)}}{[D]_0} + \sum_{i=x+1}^n k_i[E]t \quad (5)$$

From the generalization for a duplex with n guanines, definition of G_1 to G_x as intact DNA (DNA containing intact G_{x+1} to G_n guanine bases) in eq 4 then provides a sum of rate constants from G_{x+1} to G_n - *this is the second term in eq 5*.

$$-\ln \frac{[D]_{\text{intact}(1:x)}}{[D]_0} = \sum_{i=x+1}^n k_i[E]t$$

From eq 5,

$$-k_X[E]t = \ln \frac{[D]_{\text{intact}}(1:x-1)}{[D]_0} - \ln \frac{[D]_{\text{intact}}(1:x)}{[D]_0}, \quad \ln x - \ln y = \ln \frac{x}{y}$$

$$-k_X[E]t = \ln \frac{[D]_{\text{intact}}(1:x-1)}{[D]_{\text{intact}}(1:x)}$$

Therefore, quantitation of data and plotting provides the value k_X , the rate constant for alkylation at the guanine base x (Figure 1.33).

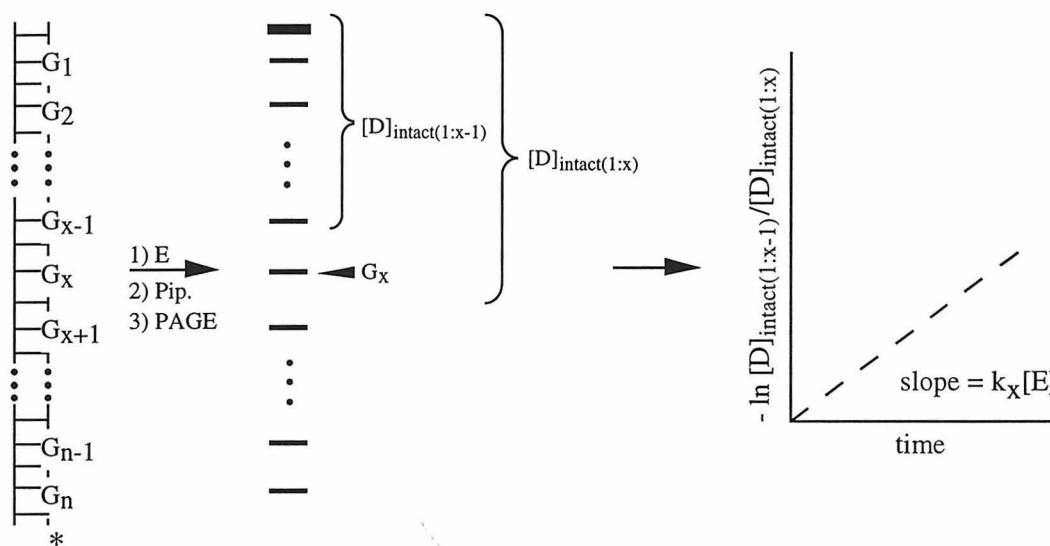


Figure 1.33. Quantitation of the data from a polyacrylamide gel gives the rate constant k_X for alkylation at the guanine base G_x .

It should be noted that total DNA concentration ($[D]_0$) cancels out in this analysis. Because the single rate constant is a difference in the values of two sums, all rate constants for

guanine bases leading to shorter fragments cancel out, obviating the need to quantitating these products in the gel analysis.

To determine a rate constant for alkylation at N7 of guanine in a restriction fragment of DNA, a 974 base pair duplex described in previous studies (7a) was 3'-³²P-end labeled and subjected to treatment with 20 and 100 mM 2-bromoacetamide. Reactions were analyzed over time by denaturing polyacrylamide gel electrophoresis (Figure 1.34). Plots of $[D]_{\text{intact}}/[D]_0$ are linear under initial rate conditions (i. e. low conversion) and provide a rate constant k_2 from an average of observed rate constants at 20 and 100 mM 2-bromoacetamide concentration of $3.8 (\pm 0.7) \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$ (Figure 1.35).

To more accurately determine this rate constant, a duplex based on the sequence from the restriction fragment encompassing the triple helix site was synthesized. As this sequence is an inverted repeat, experiments were performed to isolate the duplex from the hairpin formed in hybridization of the single stranded oligonucleotide. A non-denaturing gel was used to distinguish these two products, and as it is known that high salt conditions favor duplex formation, this was exploited to drive the equilibrium to duplex (Figure 1.36). Duplex was isolated, end labeled, and 2-bromoacetamide alkylation performed (Figure 1.37). From a series of determinations of observed rate constants, a second-order rate constant for the alkylation reaction was determined to be $2.5 (\pm 0.2) \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$ (Figure 1.38).

In order to simplify the analysis and for a more rational presentation of the results, a duplex can be designed where the products of alkylation can be resolved on a polyacrylamide gel into two experimentally distinguishable categories - those containing intact guanine base G and those resulting from alkylation at guanine base G. For this

Figure 1.34. Autoradiogram of a 6% denaturing polyacrylamide gel used to analyze the products of 2-bromoacetamide alkylation of the 3'-³²P-end labeled 974 base pair restriction fragment: lane 1, A-specific sequencing reaction; lane 2, G-specific sequencing reaction; lanes 3-8, reaction of 20 mM 2-bromoacetamide with duplex DNA at times 0, 1, 2, 5.2, 8, and 10.5 hours; lanes 9-14, reaction of 100 mM 2-bromoacetamide with triplex DNA at times 0, 1, 2, 5.2, 8, and 10.5 hours.

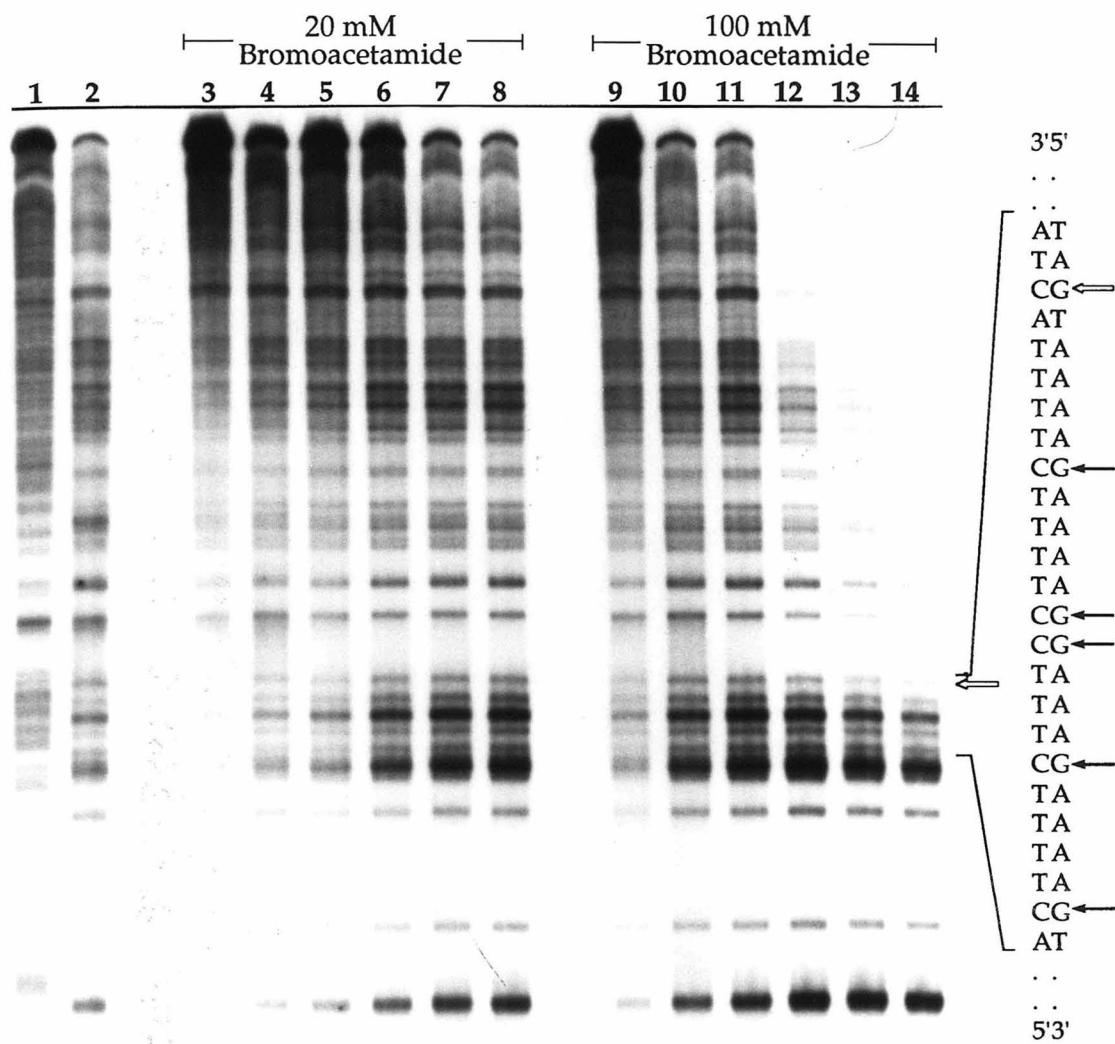
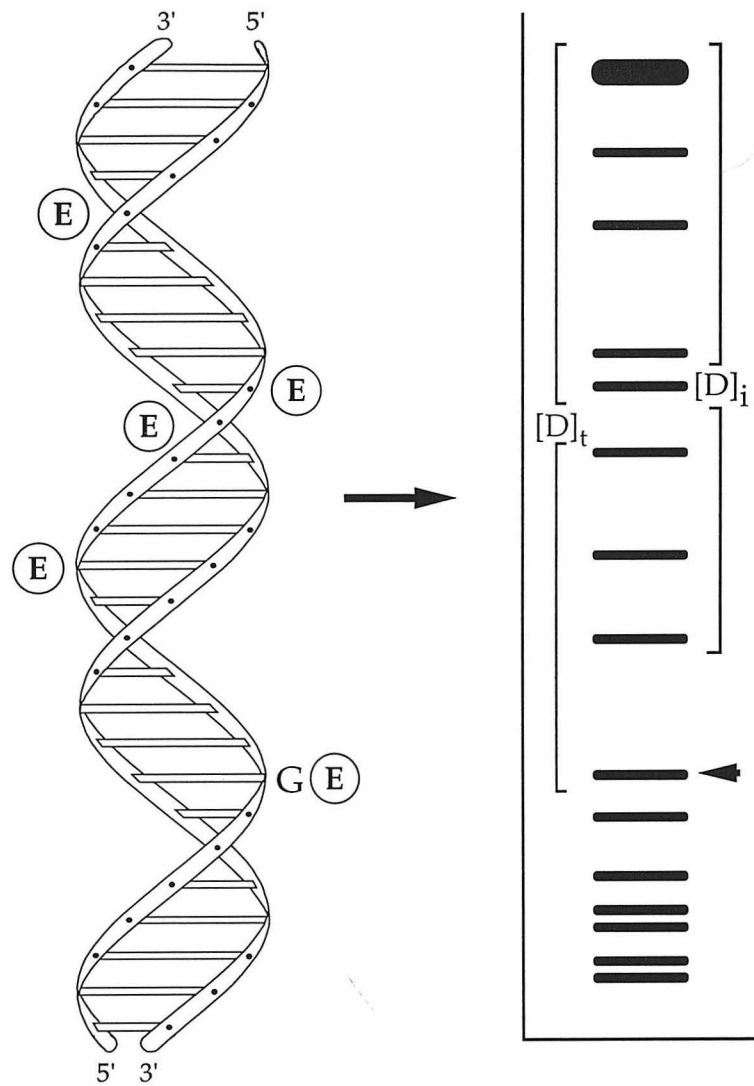


Figure 1.35. Determination of an observed rate constant k_{obs} for the intermolecular alkylation reaction of 2-bromoacetamide at the N7 position of guanine within a restriction fragment of DNA. Data is quantitated as shown in the accompanying graphic. Shown in the following graphic is a representative plot at 20 mM concentration 2-bromoacetamide of fraction intact DNA ($[D]_t/[D]_0$) versus time which provides, from the slope of the plot, an observed rate constant (k_{obs}) for the alkylation reaction of 2-bromoacetamide at the N7 position of guanine within duplex DNA of $9.1 \times 10^{-7} \text{ s}^{-1}$.



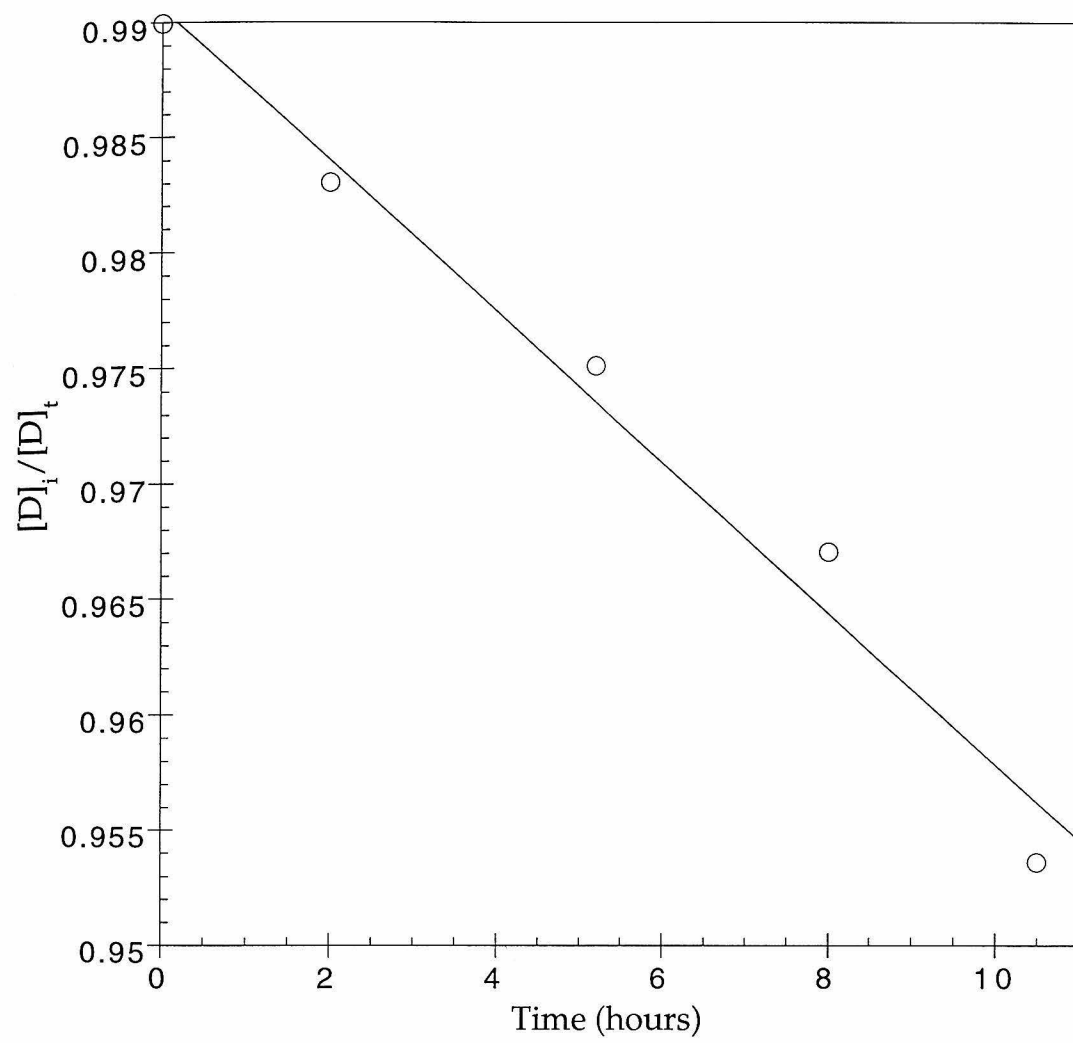


Figure 1.36. Gray-scale representation of a non-denaturing polyacrylamide gel used to analyze the effect of salt concentration on the hairpin-duplex equilibrium for the 40 base pair duplex. Samples were heated at 95°C and slow cooled in 20 mM HEPES pH 7.4 and 0-1 M sodium chloride. Gel electrophoresis then resolves the hairpin (bottom band) from the duplex (top band).

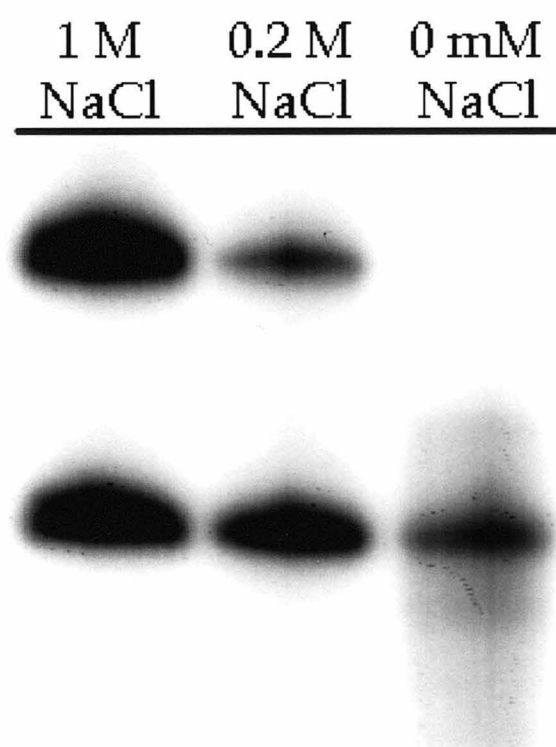


Figure 1.37. Autoradiogram of a 15% denaturing polyacrylamide gel used to analyze the products of intermolecular 2-bromoacetamide alkylation at the N7 position of guanine and base-promoted strand cleavage within the 5'-³²P-end labeled duplex: Lanes are annotated with the time points of individual time courses at 10, 50, 90 mM concentration 2-bromoacetamide.

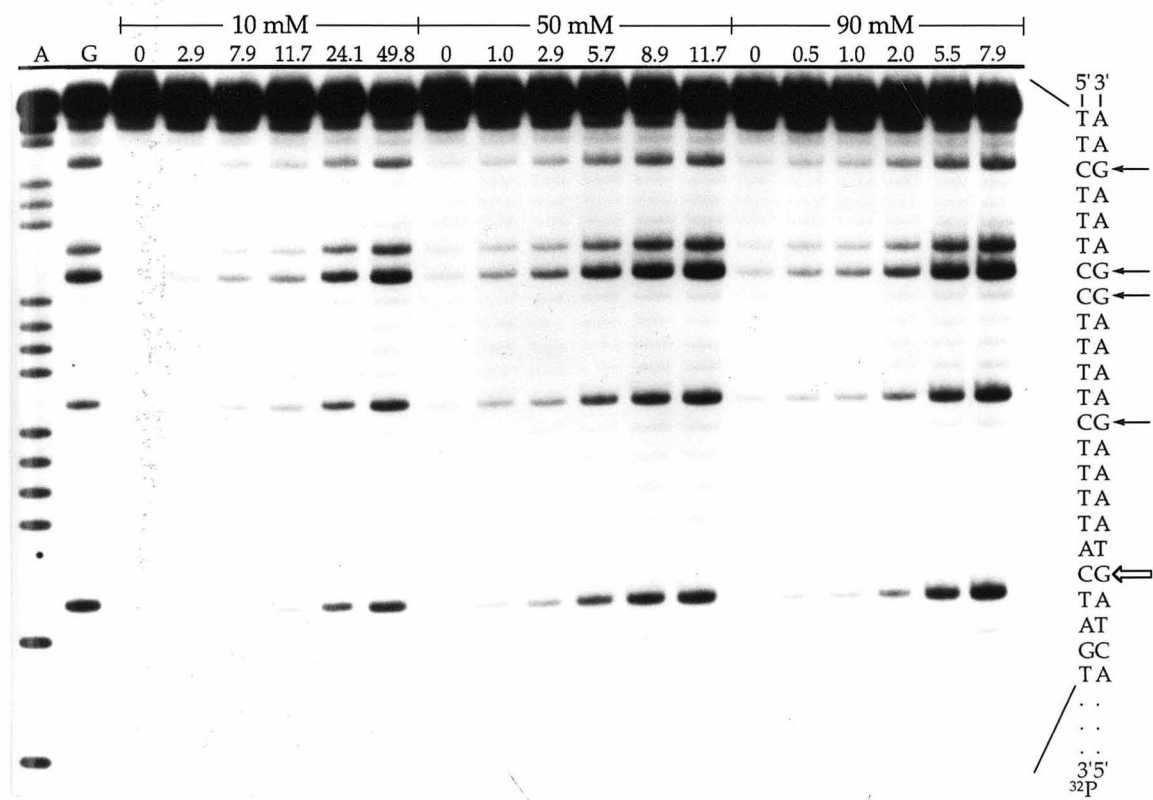
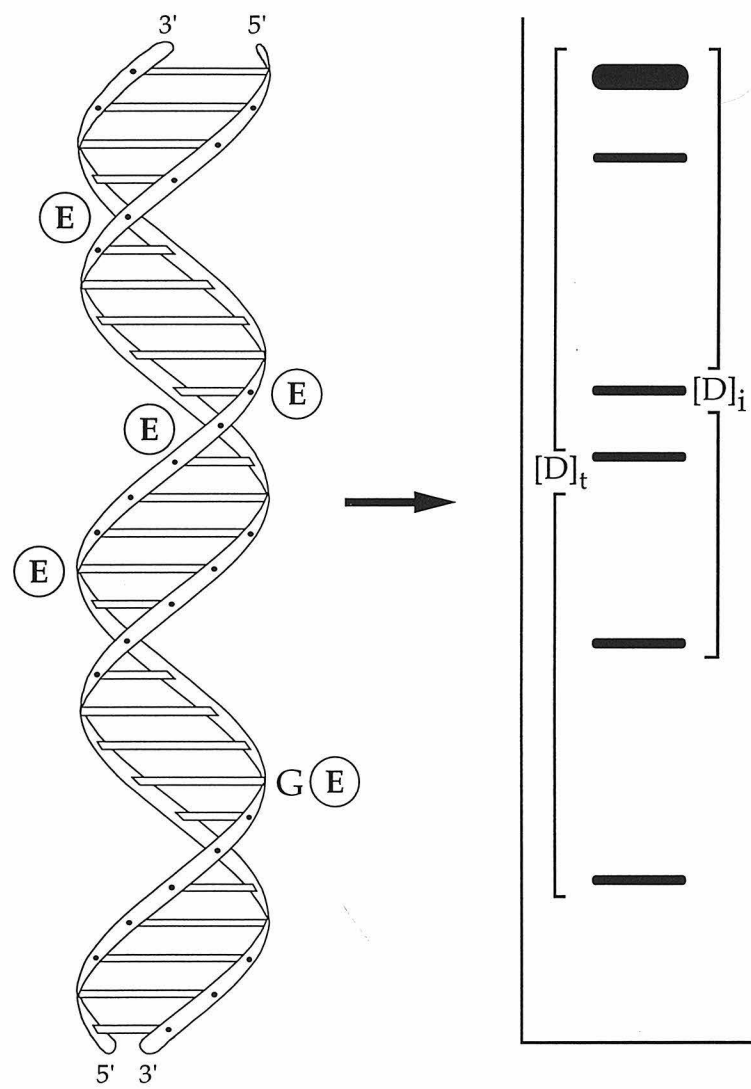
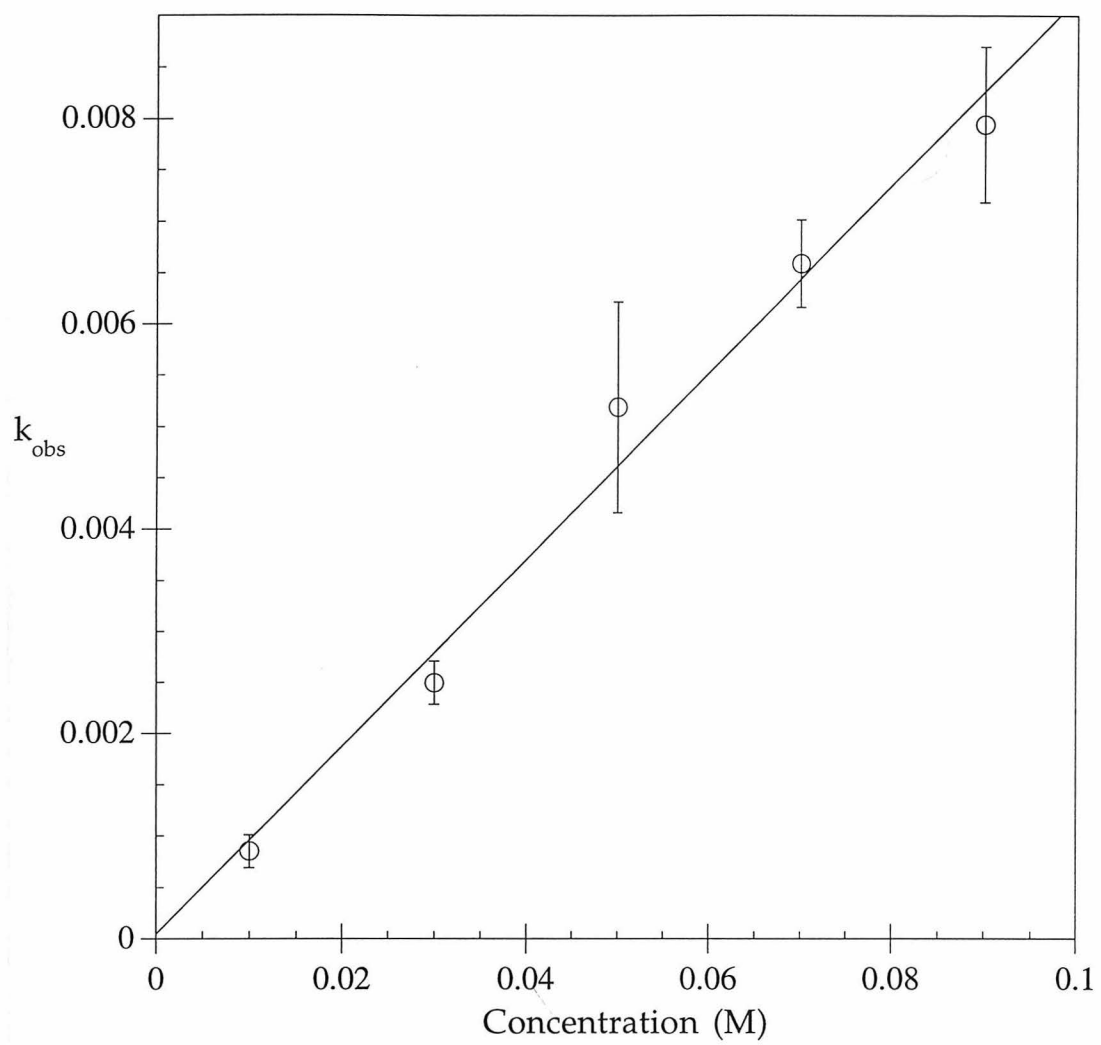


Figure 1.38. Determination of the rate constant k_2 for the alkylation of N7 of guanine in duplex DNA by 2-bromoacetamide. Data is quantitated as in the accompanying graphic. Shown in the following graphic is a plot of k_{obs} versus concentration 2-bromoacetamide, which provides from the slope a rate constant k_2 of $2.5 (\pm 0.2) \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$. Error bars represent standard deviation about the mean for two or three rate determinations at each concentration of 2-bromoacetamide.





purpose, a 39 base pair duplex of identical sequence composition to the sequence in plasmid pUCINET was synthesized.

Duplex DNA was 5'-³²P-end labeled and incubated with 2-bromoacetamide (20 mM Bis-Tris Acetate, pH 7.0, 0.8 mM cobalt hexammine trichloride, 37° C). Reactions were analyzed over time and cleavage occurs at all guanine bases in the radiolabeled strand of the duplex (Figure 1.39). Plots of $[D]_t/[D]_i$ versus time under initial rate conditions (18) provide observed rate constants (k_{obs}) for the alkylation reaction at various concentrations of 2-bromoacetamide. The slope from a plot of observed rate constant versus concentration of 2-bromoacetamide then provides a rate constant k_2 of $3.6 (\pm 0.3) \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$ for the alkylation reaction of 2-bromoacetamide at N7 of guanine within duplex DNA (Figure 1.40).

One potential oversight is that the determined rate constants were obtained by quantitation from the gels without consideration of background. That is, individual bands were quantitated for the analyses. To determine if this had an effect on the rate constant determinations, the data was quantitated as to include background. The observed rate constants obtained were found to be identical within experimental error (Figure 1.41).

All previously determined rate constants for 2-bromoacetamide alkylation are from initial rates at low conversion. To demonstrate that the alkylation reaction is first order in DNA target concentration, an alkylation reaction was performed at high concentration 2-bromoacetamide over a long period of time (Figure 1.42). Quantitation of the data results in a first-order decay of fully intact duplex, indicating that the disappearance of DNA follows the sum of a set of first-order rate constants (Figure 1.43).

Included in the observed rate enhancement is the effect of local triple helix formation on the reactivity of the targeted guanine base to nucleophilic substitution. To

Figure 1.39. Autoradiogram of a 15% denaturing polyacrylamide gel used to analyze the products of intermolecular 2-bromoacetamide alkylation at the N7 position of guanine and base-promoted strand cleavage within the 5'-³²P-end labeled duplex: lane 1, control lane - 24 hours with no 2-bromoacetamide; lane 2, A-specific sequencing reaction; lane 3, G-specific sequencing reaction; lanes 4-9, reaction of 20 mM 2-bromoacetamide with duplex DNA at times 0, 2.1, 5.8, 9.0, 12.0, and 24.0 hours; lanes 10-15, reaction of 60 mM 2-bromoacetamide with duplex DNA at times 0, 0.9, 2.1, 5.0, 8.0, and 12.0 hours.

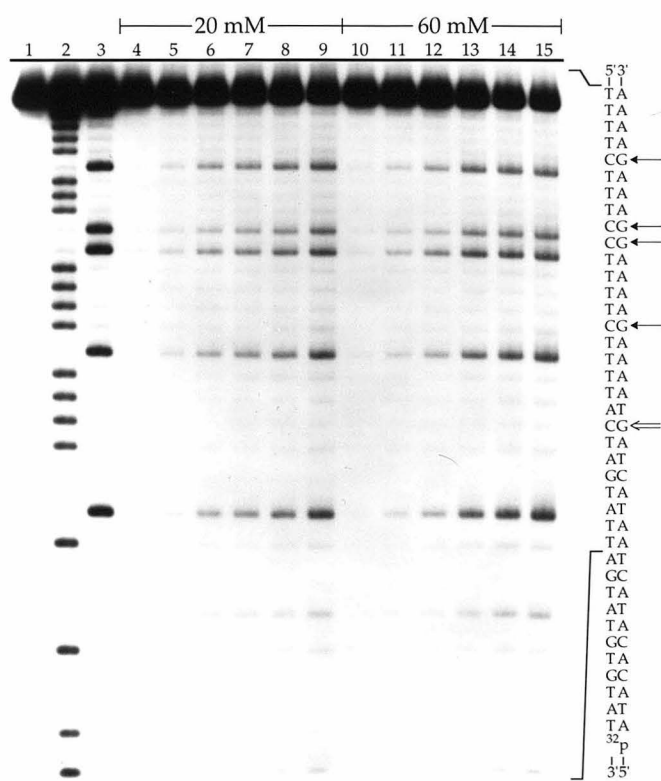


Figure 1.40. Determination of the rate constant k_2 for the intermolecular alkylation reaction of 2-bromoacetamide at the N7 position of guanine within duplex DNA. Panel a: Determination of observed rate constants for alkylation of duplex DNA at various concentrations of 2-bromoacetamide. Shown are representative plots at 10, 50, 90 mM concentrations 2-bromoacetamide of fraction intact DNA ($[D]_i/[D]_0$) versus time which provide, from the slopes of the plots, observed rate constants (k_{obs}) for the alkylation reaction of 2-bromoacetamide at the N7 position of guanine within duplex DNA of $3.4 (\pm 0.2) \times 10^{-7}$, $1.7 (\pm 0.2) \times 10^{-6}$, and $3.4 (\pm 0.1) \times 10^{-6} \text{ s}^{-1}$. Panel b: Determination of the rate constant k_2 for the alkylation of N7 of guanine in duplex DNA by 2-bromoacetamide. Shown is a plot of k_{obs} versus concentration 2-bromoacetamide, which provides from the slope a rate constant k_2 of $3.6 (\pm 0.3) \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$. Error bars represent standard deviation about the mean for two or three rate determinations at each concentration of 2-bromoacetamide.

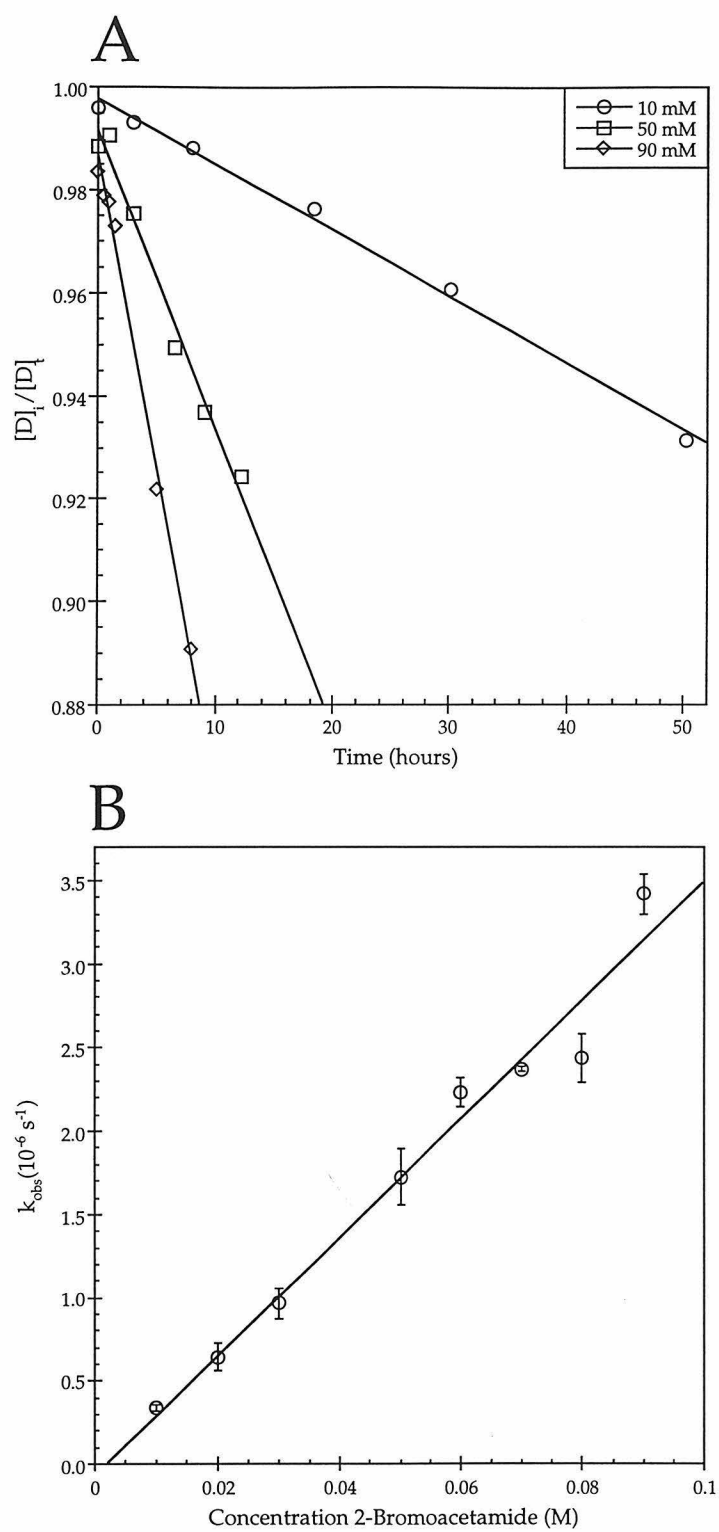
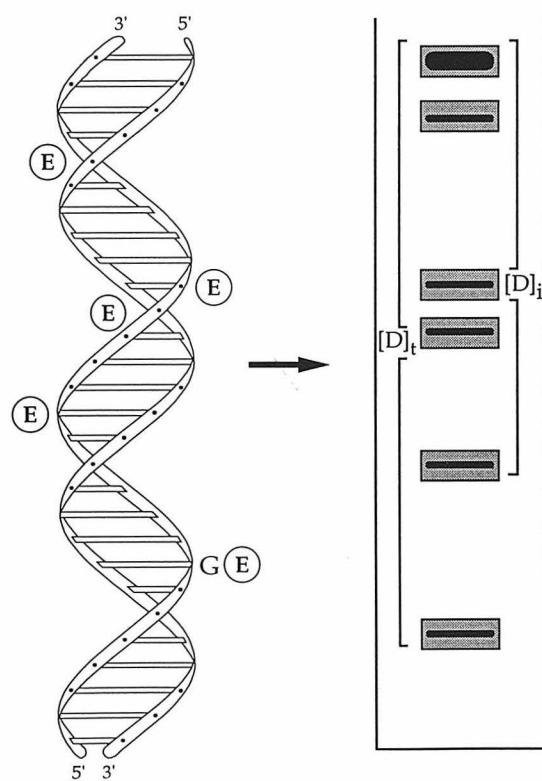
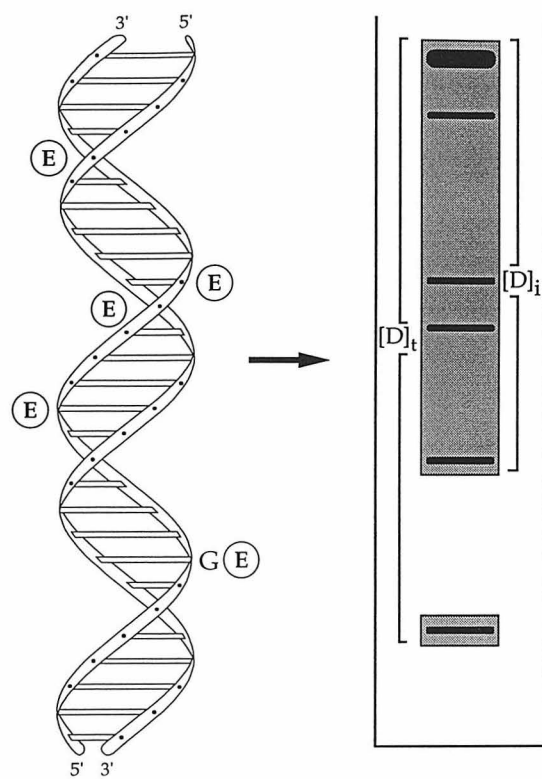


Figure 1.41. Determination of the effect of background on the rate constant k_{obs} for the alkylation of N7 of guanine in duplex DNA by 2-bromoacetamide. Data is quantitated as to include (top) or exclude (bottom) measurement of background for rate determinations. Shown in the following graphic is a plot of $[D]_i/[D]_t$ versus time at 10 mM 2-bromoacetamide, which provides identical rate constants k_{obs} from the slopes of the plots.



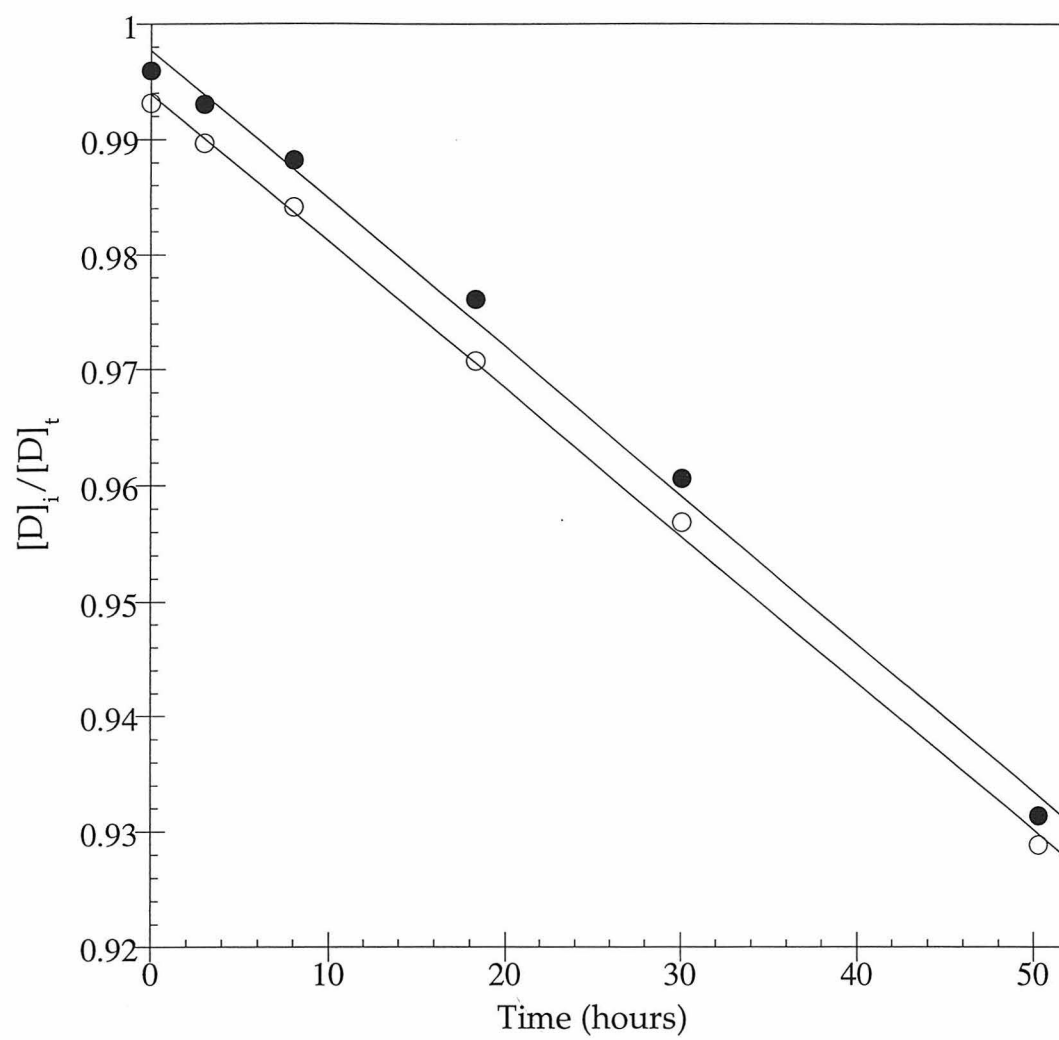


Figure 1.42. Gray-scale representation of a 15% polyacrylamide gel used to analyze the products of 2-bromoacetamide alkylation of duplex DNA at long reaction times (20 mM Bis Tris Acetate pH 7.0, 0.8 mM cobalt hexammine trichloride, 90 mM 2-bromoacetamide). Lanes are annotated with the time points of each sample.

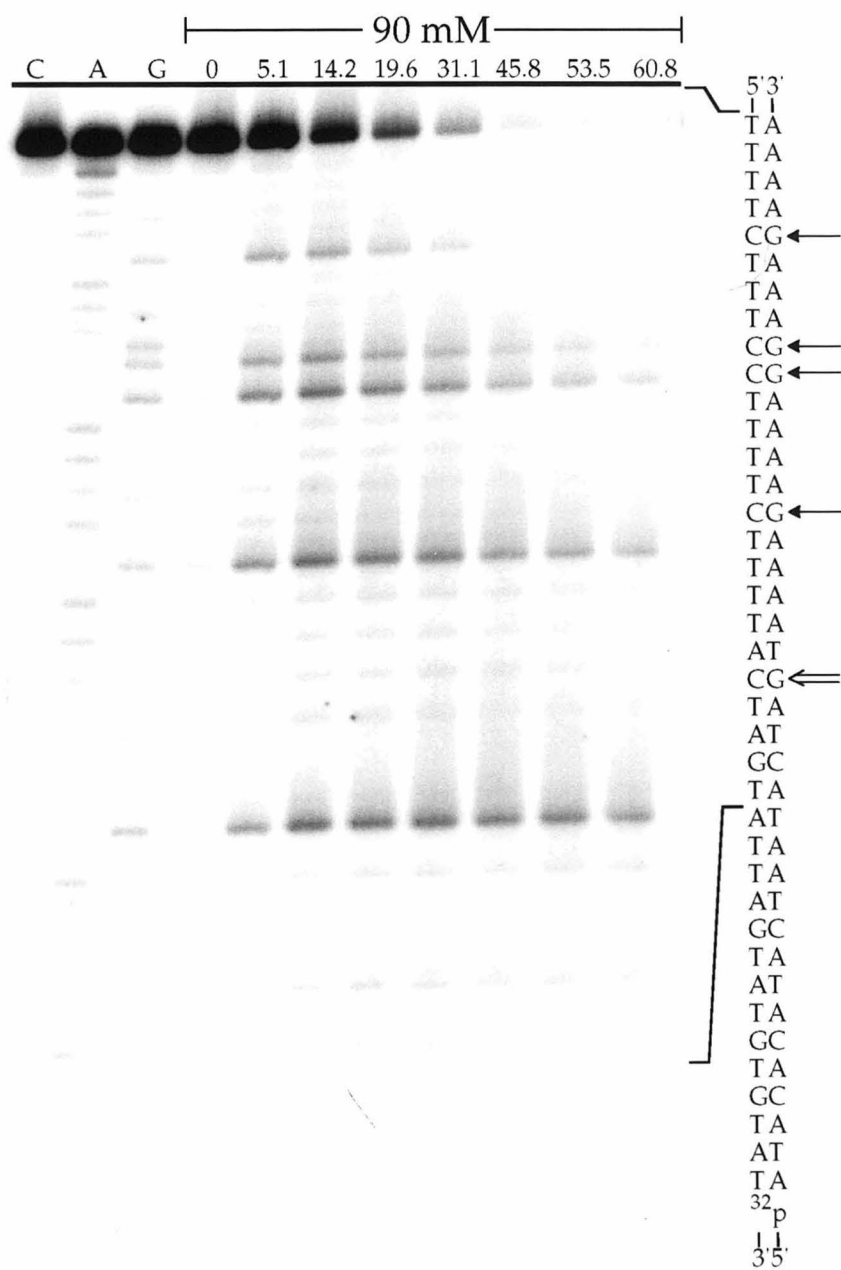
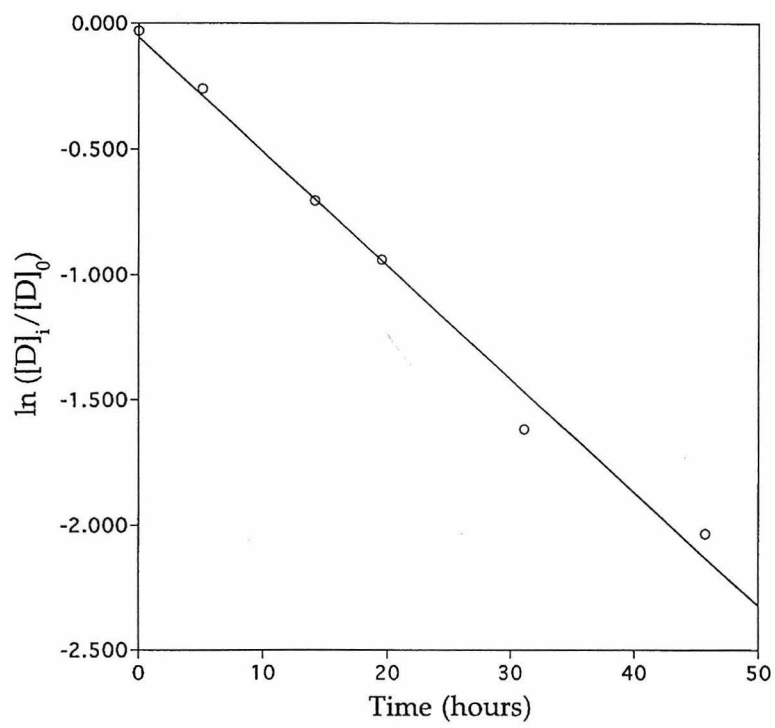
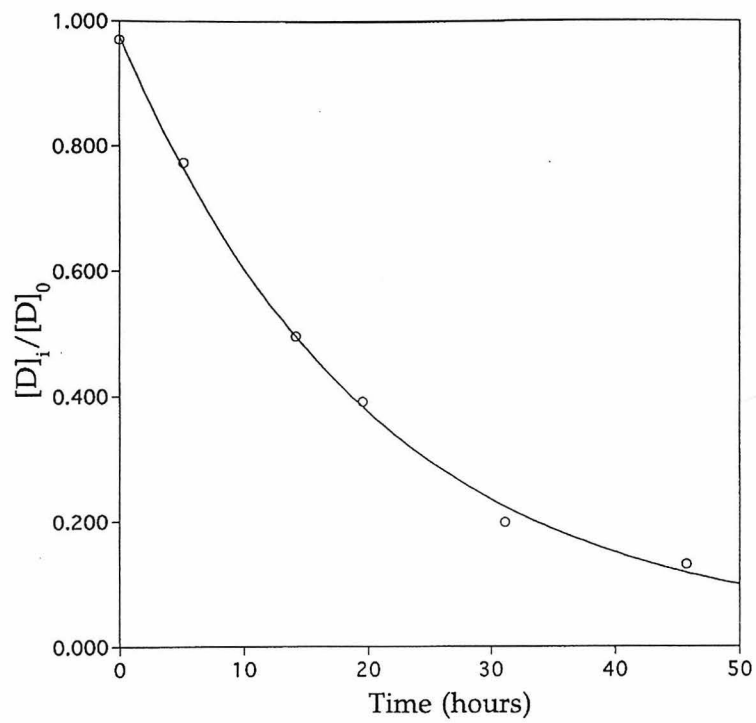


Figure 1.43. An analysis of the data at long reaction times demonstrates that loss of fully intact duplex DNA is first order in DNA concentration. Quantitation of individual bands resulting from reaction at guanine bases results in (A) exponential decay of fully intact duplex giving (B) a straight line when plotted as the natural $\ln$ versus time.



determine the magnitude of this effect, an oligonucleotide of identical sequence composition to the N-bromoacetyl oligonucleotide but lacking the linker arm and electrophilic functionality was synthesized and 2-bromoacetamide alkylation reactions were performed in the presence (20 μM) and absence of the Hoogsteen strand of the local triple-helical complex. Binding of the Hoogsteen strand in the major groove protects the N7 position of guanine from alkylation, resulting in a "footprint" at the binding site on the duplex (Figure 1.44). The data then provides observed rate constants at 80 mM 2-bromoacetamide concentration of $2.4 (\pm 0.1) \times 10^{-6} \text{ s}^{-1}$ and $1.2 (\pm 0.1) \times 10^{-6} \text{ s}^{-1}$ for alkylation of the targeted guanine base in double-helical and triple-helical DNA, respectively (Figure 1.45). This indicates an approximate two-fold decrease in the rate of nucleophilic substitution upon local triple helix formation.

The rate enhancement obtained at one position in the duplex upon localization of the alkylation reaction by oligonucleotide-directed triple helix formation can be determined from an analysis of the relative rates of alkylation of 2-bromoacetamide and the N-bromoacetyl oligonucleotide in the bound complex. These rates are conveniently expressed in terms of half-life and are summarized in Table 1.2. The shaded boxes include the most carefully measured numbers and will be used for the comparisons here. A ratio of half-lives at 200 nM concentration provides a rate enhancement of $4.0 (\pm 1.1) \times 10^6$. When the effect of local triple helix formation on the rate of alkylation by 2-bromoacetamide is considered, the rate enhancement increases to $8.0 (\pm 2.2) \times 10^6$, an almost ten-million-fold increase in the rate of nucleophilic substitution at one position in the duplex.

The maximum rate enhancement obtained from covalent attachment of bromoacetyl to any oligonucleotide is predicted to be highly dependent on the binding

Figure 1.44. Autoradiogram of a 15% denaturing polyacrylamide gel used to determine the effect of local triple helix formation on the rate of alkylation at N7 of guanine in duplex DNA by 2-bromoacetamide: lane 1, A-specific sequencing reaction; lane 2, G-specific sequencing reaction; lanes 3-8, reaction of 80 mM 2-bromoacetamide with duplex DNA at times 0, 0.5, 1.0, 2.0, 4.9, and 7.9 hours; lanes 9-14, reaction of 80 mM 2-bromoacetamide with triplex DNA at times 0, 0.5, 1.0, 2.0, 4.9, and 7.9 hours.

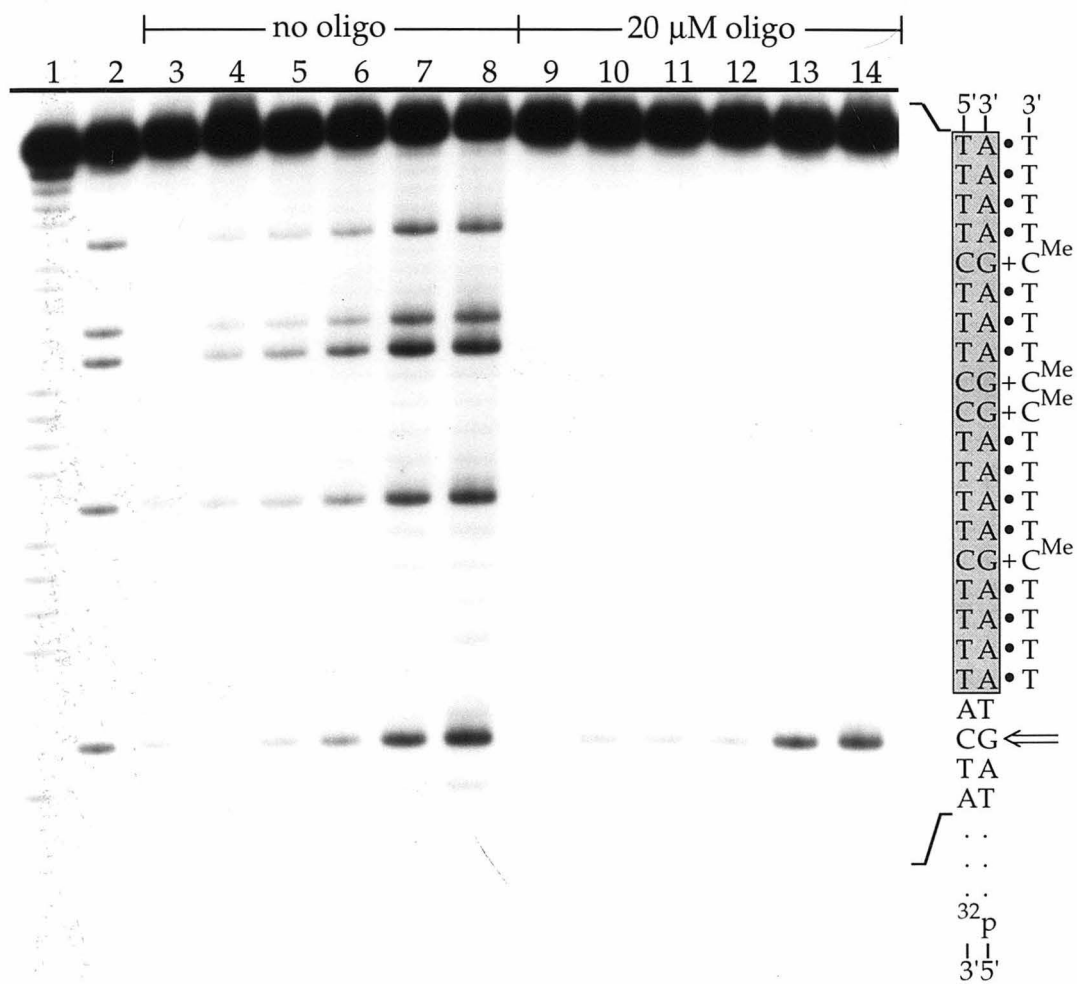


Figure 1.45. Determination of the effect of local triple helix formation on the rate of alkylation at N7 of guanine in duplex DNA by 2-bromoacetamide. Shown is a plot of $[D]_i/[D]_t$ versus time, providing observed rate constants of $2.4 (\pm 0.1) \times 10^{-6} \text{ s}^{-1}$ and $1.2 (\pm 0.1) \times 10^{-6} \text{ s}^{-1}$.

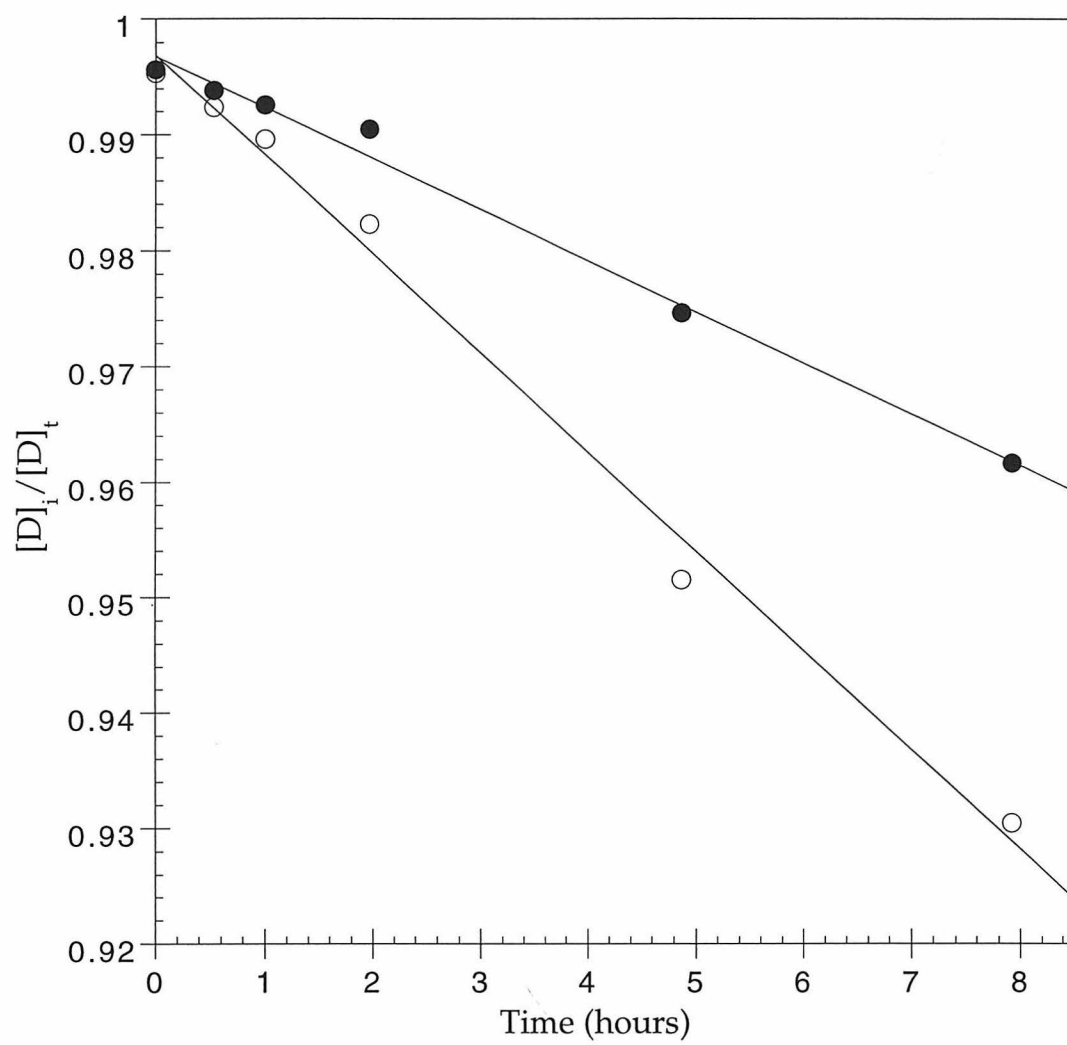


Table 1.2. Rates of Alkylation and Rate Enhancement^a

Reaction Type	Substrate	k_2 ($M^{-1}s^{-1}$)	$t_{1/2}$ at 200 nM Concentration ^b (hr)	Rate Enhancement
Oligonucleotide-Directed		-	7.2 (± 1.5) ^c	-
Bromoacetamide	Restriction Fragment	$3.8 (\pm 0.7) \times 10^{-5}$	$2.5 (\pm 0.5) \times 10^7$	$3.5 (\pm 1.2) \times 10^6$
	Old Duplex	$2.5 (\pm 0.2) \times 10^{-5}$	$3.9 (\pm 0.3) \times 10^7$	$5.4 (\pm 1.3) \times 10^6$
	New Duplex	$3.6 (\pm 0.3) \times 10^{-5}$	$2.7 (\pm 0.2) \times 10^7$	$4.0 (\pm 1.1) \times 10^6$
	Triplex	$1.5 (\pm 0.1) \times 10^{-5}$	$5.4 (\pm 0.4) \times 10^7$	$8.0 (\pm 2.2) \times 10^6$

^aError is reported as standard deviation about the mean.^bCalculated from k_2 for duplex DNA and k_{obs} at 80 mM 2-bromoacetamide for triplex DNA.^cCorresponds to an effective molarity of 0.8 M.

constant of the oligonucleotide for its duplex target site. Solution conditions which affect the binding affinity of the oligonucleotide for the duplex (i. e. polycation concentration) but, to a first approximation, leave the rate of alkylation unaffected would change the rate enhancement attainable. Similarly, oligonucleotides with intrinsically different binding constants in triple helix formation would be expected to provide different maximum rate enhancements.

Oligonucleotide-directed triple helix formation provides DNA-binding ligands with sufficient specificity for recognition of single sites within large double-helical DNA. Oligonucleotides fulfilling these criteria have dissociation constants in the micromolar to nanomolar concentration range, for which the above analysis predicts rate enhancements at a single site in duplex DNA in the 10^5 to 10^8 range. In ligand-directed alkylation reactions, the rate of alkylation by free bifunctional molecule on a fragment of double-helical DNA will formally equal the sum of the reaction rates at each nucleophilic site in the duplex. For large double-helical DNA, the sum of these rates could become significant at concentrations necessary for binding of ligands appended with electrophiles, placing a limit on attainable rate enhancements. From the above analysis, it becomes clear that oligonucleotides with high specificities (for localization at one guanine base in duplex DNA) and high affinities (for necessary rate enhancement) are crucial for this approach.

Oligonucleotide-directed alkylation of double stranded DNA has been characterized by a product analysis and by kinetic analysis relative to free electrophile. Terminal phosphate groups are produced at the cleavage site in the target duplex, making this methodology for sequence-specific cleavage of double stranded DNA compatible with molecular cloning techniques (19). By covalent attachment to the triple-helical complex, the alkylation reaction is accelerated to a reasonable rate at a specific guanine in duplex DNA under physiological conditions. This nonenzymatic approach, which couples sequence-specific recognition one million times more specific than restriction enzymes with sequence-dependent cleavage almost ten million times faster than its bimolecular counterpart, represents a powerful methodology for directed chemistry on double stranded DNA.

Experimental

^1H NMR and ^{13}C NMR spectra were recorded at 300 MHz on a General Electric QE 300 Spectrometer. Chemical shifts were recorded in parts-per-million using the proteo NMR solvent as a reference. Biochemical manipulations were carried out according to standard procedures (20) unless otherwise noted. Adenine-specific (21) and guanine-specific (20) sequencing reactions were performed as previously described. Radioactive nucleotides were purchased from Amersham and ICN. T4 Polynucleotide Kinase, Calf Alkaline Phosphatase, Klenow Fragment, and glycogen were obtained from Boehringer Mannheim. Hind III was obtained from New England Biolabs. Ssp I was obtained from Gibco/BRL. Terminal transferase was obtained from USB. Dimethyl sulfate, piperidine, and 2-bromoacetamide were obtained from Aldrich. Purity of 2-bromoacetamide was

verified by ^1H NMR. Acrylamide was purchased from BioRad. Cobalt hexamine trichloride was obtained from Fluka. Snake venom phosphodiesterase (type VIII) was purchased from Sigma. NAP-5, NAP-25, and NICK columns were purchased from Pharmacia. All reagents were used without further purification unless otherwise noted. High-resolution FAB mass spectra were obtained at the California Institute of Technology Chemistry / Biology Mass Spectrometry Facility. Matrix-assisted laser desorption / ionization time of flight mass spectra were obtained at the Protein and Peptide Microanalytical Facility at the California Institute of Technology.

High pressure liquid chromatography was performed on a Hewlett-Packard 1090 liquid chromatograph with a HPLC 3D Chemstation. Analytical reversed phase HPLC was performed on a Vydac 201HS Reversed Phase HPLC column. Oligonucleotide purification was accomplished on a Pharmacia LKB FPLC using a ProRPC HR 16/10 reversed phase column (160 mm $\times$ 10 mm). Ion retardation resin (AG 11A8) and cation exchange cellulose (Cellex-CM) were purchased from BioRad. Reversed phase resin (LiChroprep RP-18) was purchased from EM Separations.

Synthesis and Purification of Oligonucleotides. Oligonucleotides were synthesized on an Applied Biosystems Model 380B DNA Synthesizer using standard phosphoramidite chemistry (22) and deprotected with concentrated NH_4OH at 55°C for 24 hours. Following lyophilization, purification of 5'-OH-oligonucleotide was achieved using denaturing polyacrylamide gel electrophoresis. Purification of the 5'-O-dimethoxytrityl-oligonucleotide was achieved using reversed phase FPLC (solvent A: 100 mM ammonium acetate, pH 7.0; solvent B: 40% acetonitrile in solvent A). Following lyophilization of appropriate fractions and removal of the dimethoxytrityl protecting group (80% acetic acid for 20 min followed by lyophilization), purification of the deprotected

oligonucleotide was achieved by reversed phase FPLC. Following purification, oligonucleotides were desalted on a Sep-Pak cartridge (Waters) and quantitated by UV-visible spectroscopy (extinction coefficients in $M^{-1}cm^{-1}$: dT=8,700, dG=11,500, dA=15,400, dC=7,400).

Synthesis and Analysis of N-Bromoacetyloligonucleotides. Fmoc-protected aminonucleoside (6c) was incorporated at the 5'-end of the oligonucleotide 5'-TTT^{Me}CTTTT^{Me}C^{Me}CTTT^{Me}CTTTT-3' as the 5'-O-DMT-3'-phosphoramidite with a modification of standard coupling protocols. The last coupling step in the oligonucleotide synthesis was performed with a 0.15 M solution of the β -cyanoethyl phosphoramidite in tetrahydrofuran (distilled from calcium hydride) on a 1 μ mol scale. Following deprotection at 55° C for 15 hours, the 5'-O-DMT-oligonucleotide was purified by reversed phase FPLC, deprotected, and further purified by reversed phase FPLC. Acylation of the resulting oligonucleotide-amine was then performed. Ten nmol of oligonucleotide were dissolved in 10 μ L of 200 mM borate buffer, pH 8.9. 10 μ L of a 30 mM solution of freshly prepared N-hydroxysuccinimidyl bromoacetate in N,N-dimethylformamide (stored over molecular sieves) was added, the resulting solution mixed well, and after 30 seconds diluted to 0.5 mL with water and passed through a NAP-5 column (Sephadex G-25, equilibrated with water). Following elution with 1.0 mL water, the N-bromoacetyloligonucleotide was quantitated by UV-visible spectroscopy (extinction coefficient of 153,000 $M^{-1}cm^{-1}$ determined from the sum of dT=8,700, dT^E=8,700, and 5-methyl-dC=5,700) and stored at -20° C. The reaction was judged quantitative by HPLC analysis of enzyme degradation products. To the oligonucleotide was added 18 μ L water, 2.5 μ L 10 $\times$ dephosphorylation buffer (500 mM Tris-HCl, 1 mM EDTA, pH 8.5), 2.5 μ L 2

M MgCl₂, and 2 µL snake venom phosphodiesterase (2 units / 0.7 mL enzyme). Following incubation at 37° C for 2 hours, samples were filtered (0.2 µm Nylon-66 membrane, Rainin) and 23.5 µL injected. HPLC conditions- Vydac C18 201HS reverse phase column preequilibrated to 50° C, solvent A: 100 mM ammonium acetate, pH 6.0, solvent B: acetonitrile, gradient: 0-10 min at 0% B; 10-40 min to 5% B; 40-60 min at 5% B; 60-70 min to 10% B; 70-85 min at 10% B; 85-90 min to 0% B; Retention times and relative peak areas: 5-methyl-2'-deoxycytidine 5'-monophosphate 10.3 min, 2.6; thymidine 5'-monophosphate, 15.0 min, 14.6; hydroxy-(linker)-thymidine 35.2 min, 0.2; bromo-(linker)-thymidine 52.6 min, 0.8. Mass spectrum of N-bromoacetyloligonucleotide (MALDI-TOF) calcd 5934.9, found 5935.7.

5'-³²P-End Labeling of the 29 Base Pair Duplex. 50 pmol single stranded oligonucleotide 5'-GCTAGTAAAAGAAAAGGAAAGAAAAGTCG-3' was labeled with ³²P at the 5'-end by treatment with 45 units T4-polynucleotide kinase and [γ-³²P]dATP for 45 minutes at 37° C. Following phenol / chloroform extraction, ethanol precipitation, and hybridization with 100 pmol complementary single stranded oligonucleotide, the 5'-³²P-end labeled duplex was subjected to 14% non-denaturing polyacrylamide gel electrophoresis, visualized by autoradiography, excised from the gel, and eluted with 200 mM NaCl, 1 mM EDTA at 37° C overnight. The DNA was filtered (0.45 µm), ethanol precipitated, passed through a NICK column (Sephadex G-50, equilibrated with water), and stored at -20° C.

3'-³²P-End Labeling of 29 Base Pair Duplex. 50 pmol single stranded oligonucleotide 5'-GCTAGTAAAAGAAAAGGAAAGAAAAGTCG-3' was labeled with [α-³²P]ddA at the 3'-end by treatment with 85 units terminal deoxynucleotidyl

transferase and [α - 32 P]ddATP for 45 minutes at 37° C. Unincorporated nucleotide was removed by passage through a NICK column (Sephadex G-50, equilibrated with water), then butanol extracted and lyophilized to dryness. Radiolabeled single stranded oligonucleotide was then subjected to 20% denaturing polyacrylamide gel electrophoresis, visualized by autoradiography, excised from the gel, and eluted with 200 mM NaCl, 1 mM EDTA at 37° C overnight. The DNA was filtered (0.45 μ m) and ethanol precipitated. Following hybridization with 100 pmol complementary single-strand, the 3'- 32 P-end labeled duplex was purified by 14% non-denaturing polyacrylamide gel electrophoresis, visualized by autoradiography, excised from the gel, and eluted with 200 mM NaCl, 1 mM EDTA at 37° C overnight. The DNA was filtered (0.45 μ m), ethanol precipitated, passed through a NICK column (Sephadex G-50, equilibrated with water), and stored at -20° C.

G-Specific Sequencing Reactions for 29 Base Pair Duplex. 5'- or 3'- 32 P-end labeled duplex (100,000 cpm) in 2 μ l water was added to 50 μ l DMS buffer (50 mM sodium cacodylate, 1 mM EDTA, pH 8.0). To the resulting solution was added 1 μ l 10% aqueous dimethylsulfate (freshly prepared). After 20 minutes at room temperature, 1 μ l glycogen in water (20 mg/ml), 12.5 μ l DMS stop solution (1.5 M NaOAc (pH 5), 1 M β -mercaptoethanol, 250 μ g/ml yeast tRNA), and 190 μ l absolute ethanol were added, mixing well after each addition. Following chilling at -70° C for 20 minutes and centrifugation at 14,000 rpm for 40 minutes, the supernatant was discarded. 70% ethanol/water was added and following centrifugation at 14,000 rpm for 5 min. the supernatant discarded. 50 μ l water was added and, after vigorous mixing, the solution was dried *in vacuo*. To the resulting pellet was added 100 μ l 10% aqueous piperidine. The solution was heated to 90°

C for 30 min., then dried *in vacuo*. Samples were then lyophilized twice from 50 μ L water and stored at -20° C.

Enzymatic Characterization of DNA Termini. End labeled 29 base pair duplex (100,000 cpm) with and without N-bromoacetyl oligonucleotide (2.5 μ M) was incubated in 20 mM HEPES (pH 6.8) with 1 mM $\text{Co}(\text{NH}_3)_6\text{Cl}_3$ in a volume of 20 μ L at 4° C for 2 hours, then at 37° C for 24 hours. 2 μ L glycogen in water (20 mg/mL), 2 μ L 3 M sodium acetate (pH 5), and 60 μ L absolute ethanol were then added. Following chilling at -70° C for 20 minutes and centrifugation at 14,000 rpm for 40 min, the supernatant was discarded. 70% ethanol/water was added and following centrifugation at 14,000 rpm for 5 minutes the supernatant was discarded. 50 μ L water was added and, after vigorous mixing, the reaction was dried *in vacuo*. To each tube was added 100 μ L of 10% piperidine. Following incubation at 90° C for 30 min, reactions were dried *in vacuo*. Samples were then lyophilized twice from 50 μ L water.

For the 3'-end analysis, a portion (10,000 cpm) of 5'-³²P-end labeled product as well as products of dimethyl sulfate cleavage were dissolved in separate reactions in dephosphorylation buffer (20 mM Tris-HCl, pH 6.6, 20 mM MgCl_2 , 10 mM β -mercaptoethanol) and 10 units T4-polynucleotide kinase was added. After 1 hour at 37° C, reactions were ethanol precipitated. For the 5'-end analysis, a portion (20,000 cpm) of the 3'-[α -³²P]ddA-end labeled product as well as products of dimethyl sulfate cleavage were dissolved in separate reactions in calf alkaline phosphatase (CAP) buffer and 2 units CAP added. After 45 minutes at 37° C, reactions were ethanol precipitated. Reaction products, as well as control and G-reactions, were dissolved in formamide loading buffer and subjected

to gel electrophoresis on 1:20 crosslinked 20% denaturing polyacrylamide gels. Following electrophoresis, gels were visualized by autoradiography.

N-Hydroxysuccinimidyl 2-bromoacetate. (23) To a solution of N-hydroxysuccinimide (0.6 g, 5.4 mmol) and 2-bromoacetic acid (0.75 g, 5.4 mmol) in dioxane (30 mL) was added N,N-dicyclohexylcarbodiimide (1.3 g, 6.5 mmol). The resulting white suspension was stirred for 2 hours, then filtered into 100 mL petroleum ether. The resulting white precipitate was collected by filtration, washed with petroleum ether, and dried *in vacuo* to provide 0.5 g (38%) N-hydroxysuccinimidyl 2-bromoacetate as a white, crystalline solid. ^1H NMR (d_6 -DMSO) δ 2.82 (s, 4H), 4.62 (s, 2H). ^{13}C NMR (d_6 -DMSO) δ 23.1, 26.0, 164.7, 170.3. IR (KBr) 3060 (vw), 2990 (vw), 2943 (vw), 1812 (m), 1781 (m), 1736 (s), 1356 (m), 1211 (m), 1074 (m). HRMS calcd for $\text{C}_6\text{H}_6\text{NO}_4\text{Br}$ 234.9479, found 234.9469.

Alkene Precursor to Thymidine Methyl Ester 1. To a solution of 2'-deoxy-5-iodouridine (5 g, 14.2 mmol), o-toluyyl-phosphine (360 mg, 0.6 mmol), palladium(II) chloride (100 mg, 1.2 mmol), and triethylamine (30 mL) in 500 mL acetonitrile was added freshly distilled methyl acrylate (4 g, 46.4 mmol). The flask was protected from light and the solution heated to reflux for 14 hours. Following evaporation *in vacuo* the brown residue was dissolved in methanol, filtered, and adsorbed onto silica gel. Flash chromatography provided 3 g (67%) of alkene precursor as a yellow foam. The product was further purified by recrystallization from water to give 2.5 g (56%) of alkene precursor as a white solid. ^1H NMR (DMSO- d_6) δ 2.15-2.18 (m, 2H), 3.58-3.69 (m, 2H), 3.66 (s, 3H), 3.78 (m, 1H), 4.24 (m, 1H), 5.19 (b, 1H), 5.27 (b, 1H), 6.12 (t, $J=6.5$ Hz, 1H), 6.84 (d, $J=15.9$ Hz, 1H), 7.36 (d, $J=15.9$ Hz, 1H), 8.41 (s, 1H).

Thymidine methyl ester (1). 2.5 g (8 mmol) of **1** was dissolved in 300 mL methanol and added to 10% palladium on carbon (300 mg). The suspension was shaken under 55 psi H₂ until judged complete by TLC (8 hours). Following filtration, the clear solution was evaporated *in vacuo* providing 2.4 g (95%) **1** as a white foam. ¹H NMR (DMSO-d₆) δ 2.07-2.11 (m, 2H), 2.50-2.54 (m, 4H), 3.60 (s, 3H), 3.57-3.71 (m, 2H), 3.78 (m, 1H), 4.26 (m, 1H), 5.10 (b, 1H), 6.18 (t, J=6.8 Hz, 1H), 7.73 (s, 1H).

Amino-(linker)-Thymidine (T-NH₂). To a flask containing 237 mg (0.75 mmol) thymidine methyl ester **1** (13) was added 5 mL ethylenediamine (freshly distilled from sodium hydroxide). The clear solution was stirred at 50° C for 7 hours, then at room temperature for 16 hours. The ethylenediamine was removed *in vacuo* to give a clear oil. Evaporation with methanol/toluene to remove trace ethylenediamine yielded a clear oil, and upon repeated evaporation with acetonitrile, 247 mg (96%) of **T-NH₂** was obtained as a white, hygroscopic, fluffy solid. ¹H NMR (CD₃OH) δ 2.08-2.15 (m, 2H), 2.29-2.33 (m, 2H), 2.47-2.52 (m, 2H), 2.56-2.61 (m, 2H), 3.11 (dd, J = 6.4 Hz, J = 2.3 Hz, 2H), 3.66 (dd, J = 21 Hz, J = 4 Hz, 2H), 3.79 (q, J = 3.4 Hz, 1H), 4.28 (m, 1H), 6.18 (t, J = 6.7 Hz, 1H), 7.70 (s, 1H). ¹³C NMR (CD₃OH) δ 24.3, 35.8, 41.2, 41.9, 42.8, 62.9, 72.1, 86.3, 88.8, 114.1, 138.8, 152.5, 166.1, 175.3. IR (KBr) 3362 (s), 3261 (m), 2937 (w), 2855 (w), 1702 (s), 1637 (s), 1560 (m), 1282 (m), 1104 (m), 926 (w). HRMS calcd for C₁₄H₂₂N₄O₆ 342.1534, found 342.1561.

Bromo-(linker)-Thymidine (T-Br). To a solution of **T-NH₂** (19 mg, 55 μmol) in water (0.2 mL) was added N-hydroxysuccinimidyl 2-bromoacetate (25 mg, 106 μmol). The solution was vigorously mixed for 5 min, then loaded onto a 2 gram Sep-Pak cartridge equilibrated with water. The column was washed first with water, then with 50%

acetonitrile/water. Lyophilization afforded 6 mg (24%) **T-Br** as a fluffy, white solid. ^1H NMR (D_2O) δ 2.29-2.38 (m, 2H), 2.40-2.45 (t, $J = 7.1$ Hz, 2H), 2.56-2.61 (t, $J = 5.9$ Hz, 2H), 3.30 (s, 4H), 3.71-3.81 (m, 2H), 3.86 (s, 2H), 3.87-3.89 (m, 1H), 3.98-4.02 (m, 1H), 4.41-4.46 (m, 1H), 6.23-6.28 (t, $J = 6.7$ Hz, 1H), 7.65 (s, 1H). ^{13}C NMR (D_2O) δ 23.0, 28.1, 34.4, 38.4, 38.8, 39.3, 61.2, 70.5, 85.2, 86.6, 113.3, 138.3, 151.6, 165.6, 170.2, 175.5. IR (KBr) 3507 (m), 3365 (s), 3049 (m), 2931 (m), 2802 (w), 1686 (s), 1655 (s), 1625 (m), 1560 (m), 1424 (m), 1048 (m). HRMS calcd for $\text{C}_{16}\text{H}_{24}\text{N}_4\text{O}_7\text{Br}$ 463.0823, found 463.0821.

Hydroxy-(linker)-Thymidine (T-OH). To a solution of N-hydroxysuccinimide (1.5 g, 13.2 mmol) and glycolic acid (1 g, 13.2 mmol) in dioxane (16 mL) was added N,N-dicyclohexylcarbodiimide (3.3 g, 15.8 mmol). The resulting white suspension was stirred for 2 hours, then filtered into 100 mL petroleum ether. The resulting white precipitate was collected by filtration, washed with petroleum ether, and dried *in vacuo* to provide 1.2 g of a white solid, used without further purification. To a solution of **T-NH₂** (44 mg, 129 μmol) in water (0.44 mL) was added the NHS-ester of glycolic acid (34 mg). The solution was vigorously mixed for 5 min, then stirred for 4 hours. The reaction mixture was dried *in vacuo* to a yellow oil. The oil was dissolved in 50% aqueous acetonitrile (300 μL), then adsorbed onto silica gel. Column chromatography (0-7% water/acetonitrile), followed by adsorption and elution from reversed phase resin (LiChroprep RP-18) afforded 7 mg (14%) **T-OH** as a white solid. ^1H NMR (d_6 -DMSO): δ 2.03-2.07 (m, 2H), 2.20-2.26 (t, $J = 7.4$ Hz, 2H), 2.38-2.43 (t, $J = 7.4$ Hz, 2H), 3.08-3.14 (m, 4H), 3.53-3.58 (m, 2H), 3.73-3.75 (m, 1H), 3.77 (s, 2H), 4.21-4.23 (m, 1H), 5.01-5.05 (m, 1H), 5.22-5.24 (d, $J = 4$ Hz, 1H), 6.12-6.17 (t, $J = 6.9$ Hz, 1H), 7.63 (s,

1H), 7.81-7.84 (m, 1H), 7.85-7.90 (m, 1H), 11.28 (s, 1H). ^{13}C NMR (D_2O) δ 23.0, 34.4, 38.4, 38.6, 38.8, 60.9, 61.2, 70.5, 85.2, 86.6, 113.3, 138.3, 151.5, 165.5, 175.2, 175.5. IR (KBr) 3448 (s), 3413 (s), 3072 (m), 2931 (m), 1690 (s), 1655 (s), 1561 (m), 1543 (m), 1273 (m), 1096 (m), 1032 (w). HRMS calcd for $\text{C}_{16}\text{H}_{24}\text{N}_4\text{O}_8$ 400.1588, found 400.1607.

N-7-carboxymethyl Guanine. (24) To a flask containing guanosine (1.8 g, 6.6 mmol), iodoacetic acid (4.8 g, 26 mmol), and lithium hydroxide (0.31 g, 12.8 mmol) was added 12 mL water. The resulting suspension was stirred at reflux for 70 minutes, during which time a clear, dark red solution was formed. The solution was cooled and solid sodium thiosulfate was added until the red solution turned to a light pink color. The solution was then adjusted to pH 5.5 with LiOH(s). After storage overnight at 4° C, the pink precipitate was collected by filtration. Following recrystallization from 5% aqueous acetic acid 350 mg (26%) of N-7-carboxymethyl guanine was obtained as a pale pink solid. ^1H NMR ($\text{NaOD}/\text{D}_2\text{O}$): δ 4.86 (s, 1H), 7.71 (s, 1H). ^{13}C NMR ($\text{NaOD}/\text{D}_2\text{O}/d_6\text{-DMSO}$) δ 40.0 ($d_6\text{-DMSO}$), 50.8, 111.6, 144.0, 160.0, 162.2, 166.5, 176.5. IR (KBr) 3537 (s), 3316 (s), 3124 (s), 2694 (s), 1659 (s), 1502 (w), 1474 (m), 1232 (m), 1103 (w), 897 (w). HRMS calcd for $\text{C}_7\text{H}_7\text{N}_5\text{O}_3$ 209.0548, found 209.0545.

Guanine-(linker)-Thymidine (T-G). To 1.5 mL of 50% aqueous pyridine (pH 8.5) was added N-7-carboxymethyl guanine (20 mg, 96 μmol), 4-nitrophenol (133 mg, 960 μmol), and thymidine-amine **T-NH₂** (36 mg, 105 μmol). 1-(3-Dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride (EDCI) (total 576 μmol) was then added to the resulting yellow solution in portions of 1/5 equivalent. The solution was lyophilized, dissolved in 2 mL 10 mM NH_4OAc , pH 4.5 and extracted with an equal amount of chloroform. Removal of residual pyridine was achieved by passage of the

aqueous solution through cation exchange cellulose (CM Sephadex). The reaction was then subjected to purification by reversed phase chromatography (LiChroprep C-18 resin), affording a further purified mixture of adduct **T-G** and p-nitrophenol, which was then lyophilized, dissolved in water, and passed through ion retardation resin (Biorad AG 11A8) to remove p-nitrophenol. Lyophilization afforded the authentic standard **T-G** (2 mg) as a fluffy, white solid. ^1H NMR (d_6 -DMSO): δ 2.04-2.08 (m, 2H), 2.21-2.26 (t, J = 7.4 Hz, 2H), 2.38-2.43 (t, J = 7.5 Hz, 2H), 3.31 (s, 4H), 3.53-3.57 (m, 2H), 3.73-3.77 (m, 1H), 4.21-4.23 (m, 1H), 4.85 (s, 2H), 5.01-5.04 (t, J = 5.1 Hz, 1H), 5.23-5.24 (d, J = 4.2 Hz, 1H), 6.10 (s, 2H), 6.13-6.17 (t, J = 6.9 Hz, 1H), 7.65 (s, 1H), 7.79-7.81 (s, 1H), 7.83 (s, 1H), 8.17 (s, 1H), 10.72 (s, 1H), 11.31 (s, 1H). ^{13}C NMR (d_6 -DMSO): δ 23.0, 34.5, 38.5, 38.9, 39.8, 48.6, 61.7, 70.7, 84.3, 87.6, 109.0, 113.1, 137.6, 144.5, 150.8, 153.2, 155.2, 160.1, 163.8, 167.4, 172.0. IR (KBr) 3358 (s), 2924 (m), 2848 (w), 1686 (s), 1560 (m), 1384 (m), 1273 (w), 1108 (w). HRMS calcd for $\text{C}_{21}\text{H}_{27}\text{N}_9\text{O}_8$ 533.1977, found 533.1989.

HPLC Analysis of Products of Oligonucleotide-Directed Alkylation. 20 nmol

single stranded oligonucleotide 5'-TCGGTACCCGGGGATCTAAAGTAAAAGAAAAGGAAAGAAAAAGCTTTCTTC TTCCCTATC-3' ($\epsilon=6.0 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$) and complementary oligonucleotide ($\epsilon=5.7 \times 10^5 \text{ M}^{-1} \text{ cm}^{-1}$) were hybridized as follows: oligonucleotides were combined in 400 μL water and passed through a NAP-5 column (G-25 Sephadex, equilibrated with water). To the 1.0 mL of eluent in a 1.7 mL microfuge tube was added 100 μL of 200 mM HEPES, pH 7.0. The solution was heated to 90° C for 9 minutes, then cooled to room temperature. After 15 hours, the resulting solution was then combined in a 15 mL Falcon tube with 4.5 mL of

water, 570 μ L of 200 mM HEPES, pH 7.0, and a stirbar. After cooling to 4° C with stirring, 20 nmol N-bromoacetyloligonucleotide in 1 mL water was added, followed by 170 μ L of 8 mM cobalt hexammine trichloride (to give a final concentration of 2.7 μ M oligonucleotides). This solution was stirred at 4° C for 1.5 hours, then placed in an oil bath equilibrated to 37° C. The cloudy solution was stirred and after 10 min the first aliquot (1.1 mL, 3 nmol) was removed for workup. The sample was pipetted into a 1.7 mL microfuge tube with 100 μ L of 2 M sodium chloride (170 mM NaCl final concentration). The resulting clear solution was heated to 90° C for 15 minutes, then cooled to room temperature. The solution was diluted to 2.5 mL and passed through a NAP-25 column (Sephadex G-25, equilibrated with water), the 3.5 mL eluent collected in two 2 mL microfuge tubes, and the solutions dried *in vacuo*. Product was then transferred to a 0.7 mL microfuge tube with 100 μ L of water, followed by washes of three 100 μ L aliquots of water to ensure complete transfer of material. Samples were then dried *in vacuo* and stored at -20° C. The above procedure was repeated at time points of 3, 7, 10.5, 19.2, 29, and 69.5 hours (relative to the time of the first sample withdrawn from the reaction mixture).

Just prior to HPLC analysis, to each tube was added 18 μ L of water, 2.5 μ L of 10X dephosphorylation buffer (500 mM Tris-HCl, 1 mM EDTA, pH 8.5), 2.5 μ L of 2 M $MgCl_2$, and 2 μ L of snake venom phosphodiesterase (2 units/0.7 mL enzyme). Following incubation at 37° C for 2 hours, samples were filtered (0.2 μ m Nylon-66 membrane, Rainin) and 23.5 μ L were injected. A blank run with 23.5 μ L of water injected between each run was performed to prevent contamination of HPLC chromatograms from previous runs. HPLC conditions- Vydac C18 201HS reverse phase column preequilibrated to 50° C. Solvent A: 100 mM ammonium acetate, pH 6.0, solvent B: acetonitrile. Gradient: 0-10 min at 0% B; 10-40 min to 5% B; 40-60 min at 5% B; 60-70 min to 10% B; 70-85 min at

10% B; 85-90 min to 0% B. Retention times: 2'-deoxycytidine 5'-monophosphate - 7.7 min; 5-methyl-2'-deoxycytidine 5'-monophosphate - 10.7 min; thymidine 5'-monophosphate - 14.5 min; 2'-deoxyadenosine 5'-monophosphate - 26.6 min; hydroxy-(linker)-thymidine nucleoside - 38.4 min; guanine-(linker)-thymidine nucleoside - 47.8 min; bromo-(linker)-thymidine nucleoside - 55.2 min; undigested oligodeoxyribonucleotides - 70-80 min; authentic standard **T-G** injected under identical solution conditions - 48.5 min; hydroxy-(linker)-thymidine - 39.0 min; bromo-(linker)-thymidine - 55.5 min.

HPLC Analysis of 2-Bromoacetamide. Samples were prepared, aliquoted into sealed HPLC vials, and maintained at 37°C. Over time, vials were removed and 10 µL injected onto a Rainin Microsorb reversed phase HPLC column. solvent A: 0.1% trifluoroacetic acid

3'-³²P-End Labeling of the 974 Base Pair Restriction Fragment. pUCLEU2C (7a) was digested with the restriction enzyme Nar I to produce 5.6 and 1.1 kbp fragments. 10 µg of DNA was labeled at the 3'-end using Sequenase (v. 2.0) to incorporate [α -³²P]dCTP and [α -³²P]dGTP. Unincorporated nucleotides were removed by passage through a NICK column (Sephadex G-50, equilibrated with water), the eluent extracted with phenol/chloroform, and the 3'-³²P-end labeled DNA ethanol precipitated. The DNA was digested with EcoRI, producing fragments 4.5, 1.1, 0.9, and 0.2 kbp in size. Following separation by 5% polyacrylamide gel electrophoresis and visualization of the gel by autoradiography, the 0.9 kbp fragment containing the target site was excised from the gel and eluted with 200 mM NaCl, 1 mM EDTA at 37°C overnight. The radiolabeled

restriction fragment was filtered (0.45 μ m), ethanol precipitated, and passed through a NICK column.

Kinetic Measurements: 2-Bromoacetamide Alkylation of the 974 Base Pair Restriction Fragment. The NarI / EcoRI 3'-³²P-end labeled pUCLEU2C restriction fragment (70,000 cpm) was dissolved in 40 mM HEPES buffer, pH 7.4 with 1.6 mM cobalt hexammine trichloride in a volume of 65 μ L. Following incubation at 37°C for 10 min, an equal volume of either 40 mM or 200 mM 2-bromoacetamide in water was added. The resulting solution was aliquoted into six 700 μ L microfuge tubes (20 μ L per tube, equilibrated at 37°C). and to one tube was added 2 μ L glycogen in water (20 mg/mL), 2 μ L 3M sodium acetate pH 5, and 48 μ L ethanol. Reactions were incubated at 37°C and quenched as above over time. Following quenching of the last time point the reactions were centrifuged at 14,000 rpm for 30 min, the supernatant discarded and the pellet washed with 70% ethanol/water. The supernatant was discarded and the pellets dried *in vacuo*. To each tube was added 100 μ L 5% piperidine. Following incubation at 90°C for 30 min, the solutions were dried. Samples were lyophilized twice from 50 μ L water and dissolved in formamide buffer. Separation of products was performed by 6% denaturing polyacrylamide gel electrophoresis. Dried polyacrylamide gels were analyzed by phosphorimagery and cleavage efficiencies determined by quantitative analysis of storage phosphor autoradiograms using a PhosphorImager (Model 400S) and Image Quant (Molecular Dynamics) software.

Cloning of Plasmid pUCINET. The plasmid pUCINET was constructed by ligation of the duplex formed between oligonucleotides of sequence 5'-GATCATACACATACTAATACTAGTAAAAGAAAAGGAAAGAAAA-3' and 5'-TCGATTTTCTTTCCTTTTCTTTTACTAGTATTAGTATGTGTAT-3' with pUC19

previously digested with Bam HI and Sal I. Ligation products were used to transform Epicurian Coli XL 1 Blue competent cells. Colonies were selected for α -complementation on Luria-Bertani medium agar plates containing 50 μ g/mL ampicillin and XGAL and IPTG solution. Large scale plasmid purification was performed using Qiagen purification kits according to the manufacturer's protocol. Plasmid DNA concentration was determined from absorbance at 260 nm using the relation 1 OD unit = 50 μ g/mL duplex DNA. The sequence of the inserted region was verified by dideoxy sequencing of the resulting recombinant plasmid.

3'-³²P-End Labeling of the 666 Base Pair Restriction Fragment. Plasmid pUCINET was linearized with the restriction enzyme Hind III. 8 μ g of DNA was labeled at the 3'-end using Klenow to incorporate [α -³²P]dATP and [α -³²P]TTP. Following a chase fill-in with dATP, TTP, dCTP, and dGTP the reaction was extracted with phenol/chloroform, and the 3'-³²P-end labeled DNA ethanol precipitated. The DNA was digested with Ssp I, producing fragments 2 and 0.7 kbp in size. Following ethanol precipitation, separation by 5% polyacrylamide gel electrophoresis, and visuallization of the gel by autoradiography, the 666 base pair fragment containing the target site was excised from the gel and eluted with 200 mM NaCl, 1 mM EDTA at 37° C overnight. The radiolabeled restriction fragment was filtered (0.45 μ m), ethanol precipitated, passed through a NICK column (Sephadex G-50, equilibrated with water), 1/10 th volume 200 mM Bis-Tris Acetate pH 7.0 was added, and stored at -20° C.

A-Specific Sequencing Reaction for Restriction Fragment. The Hind III / Ssp I 3'-³²P-end labeled restriction fragment (100,000 cpm) from plasmid pUCINET in 10 μ l was added to 153 μ l water. To the resulting solution was added 4 μ l calf thymus DNA (1

mM in base pair concentration) followed by 40 μ l of a solution of 1.6 mg/100 μ l potassium tetrachloropalladate (II) in 100 mM HCl/NaOH, pH 2.0. After 5 minutes at room temperature, 5 μ l glycogen in water (20 mg/ml), 50 μ l DMS stop solution (1.5 M NaOAc, 1 M β -mercaptoethanol, 250 μ g/ml yeast tRNA), and 750 μ l absolute ethanol were added. Following chilling at -80° C for 5 minutes and centrifugation at 14,000 rpm for 30 minutes, the supernatant was discarded. The resulting yellow pellet was dried *in vacuo*. 100 μ l 10% aqueous piperidine was added, the solution heated to 90° C for 30 min., then dried *in vacuo*. Samples were then lyophilized twice from 50 μ l water and stored at -20° C until analyzed by high resolution polyacrylamide gel electrophoresis.

G-Specific Sequencing Reaction for Restriction Fragment. The Hind III / Ssp I 3'-³²P-end labeled restriction fragment (100,000 cpm) from plasmid pUCINET in 10 μ l was added to 190 μ l DMS buffer (50 mM sodium cacodylate, 1 mM EDTA, pH 8.0). To the resulting solution was added 4 μ l calf thymus DNA followed by 5 μ l 10% dimethylsulfate in absolute ethanol (freshly prepared). After 5 minutes at room temperature, 5 μ l glycogen in water (20 mg/ml), 50 μ l DMS stop solution, and 750 μ l absolute ethanol were added. Following chilling at -80° C for 5 minutes and centrifugation at 14,000 rpm for 30 minutes, the supernatant was discarded. The resulting pellet was dried *in vacuo*. 100 ml 10% aqueous piperidine was added, the solution heated to 90° C for 30 min., then dried *in vacuo*. Samples were then lyophilized twice from 50 μ l water and stored at -20° C until analyzed by high resolution polyacrylamide gel electrophoresis.

Kinetic Measurements: N-Bromoacetyloligonucleotide Alkylation of the 666 Base Pair Restriction Fragment. A 2X stock solution was made containing 3'-³²P-end labeled Hind III/Ssp I pUCINET restriction fragment (4,000 cpm/ μ L), 40 mM Bis-Tris acetate (pH 7.0), and 1.6 mM cobalt hexammine trichloride. The stock solutions were then

aliquoted into microfuge tubes and to each was added an equivalent volume of 2X aqueous N-bromoacetyloligonucleotide at 2, 0.4, and 0.2 μM concentrations. The solutions were incubated overnight at 8° C, then each was split into 7 tubes and put at 37° C. After 30 minutes, the first tube (arbitrarily designated time equals zero) was quenched by the addition of 2 μL glycogen (20 mg/mL), 2 μL 3 M sodium acetate, and 60 μL ethanol. Samples were stored at -80° C until completion of the experiment. Subsequent time points were treated as described. The samples were then centrifuged at 14,000 rpm for 30 minutes, the supernatants discarded, and reactions were dried *in vacuo*. 100 μL 10% aqueous piperidine was added and reactions heated to 90° C for 30 minutes. Reactions were lyophilized, then lyophilized twice from 50 μL water and dissolved in formamide loading buffer. Separation of products was performed by 1:20 cross-linked 8% denaturing polyacrylamide gel electrophoresis (5,000 cpm/lane). Dried polyacrylamide gels were analyzed by storage phosphor autoradiogram.

Band densities were determined by quantitative analysis of storage phosphor autoradiograms using a PhosphorImager (Model 400S) and Image Quant (Molecular Dynamics) software. Integrated volumes were determined for individual bands and numerical values were transferred to Microsoft Excel. Bands were assigned as depicted in Figure 7 to determine values for "[D]_i" and "[D]_t" and " $\ln ([D]_i/[D]_t)$ " was plotted versus time using the program KaleidaGraph to determine from the slope the value for k_1 . Kinetics were first-order in target DNA concentration over two half-lives and provided rate constants of $2.4 (\pm 0.3) \times 10^{-5}$, $2.7 (\pm 0.5) \times 10^{-5}$, and $1.6 (\pm 0.4) \times 10^{-5} \text{ s}^{-1}$ for 1, 0.2, and 0.1 μM concentration N-bromoacetyloligonucleotide, respectively.

5'-³²P-End Labeling of the 39 Base Pair Duplex. 50 pmol single stranded oligonucleotide 5'-ATACACATACTAATACTAGTAAAAGAAAAGGAAAGAAAA-

3' was labeled with ^{32}P at the 5'-end by treatment with 45 units T4-polynucleotide kinase and $[\gamma\text{-}^{32}\text{P}]\text{dATP}$ for 45 minutes at 37°C . Following ethanol precipitation the DNA was resuspended in formamide loading buffer and purified by 20% denaturing polyacrylamide gel electrophoresis. Following elution, ethanol precipitation, and hybridization with excess complementary single stranded oligonucleotide, the 5'- ^{32}P -end labeled duplex was then subjected to 12% non-denaturing polyacrylamide gel electrophoresis, visualized by autoradiography, excised from the gel, and eluted with 200 mM NaCl, 1mM EDTA at 37°C overnight. The DNA was filtered (0.45 μm), ethanol precipitated, passed through a NICK column (Sephadex G-50, equilbrated with water), 1/10 th volume 200 mM Bis-Tris Acetate pH 7.0 was added, and stored at -20°C .

A-Specific Sequencing Reaction for 39 Base Pair Duplex. To the 5'- ^{32}P -end labeled 39-mer duplex (200,000 cpm) in 158 μl water was added 4 μl calf thymus DNA (1mM base pair) followed by 40 μl 1.6 mg/100 μl potassium tetrachloropalladate(II) in 100 mM HCl/NaOH, pH 2 (freshly prepared). After 5 minutes at room temperature, 5 μl glycogen in water (20 mg/ml), 50 μl DMS stop solution, and 750 μl absolute ethanol were added. Following chilling at -20°C for 15 minutes and centrifugation at 14,000 rpm for 30 minutes, the supernatant was discarded. The resulting pellet was dried *in vacuo*. 100 μl 10% aqueous piperidine was added, the solution heated to 90°C for 30 min., then dried *in vacuo*. Samples were then lyophilized twice from 50 μl water and stored at -20°C until analyzed by high resolution polyacrylamide gel electrophoresis.

G-Specific Sequencing Reaction for 39 Base Pair Duplex. The 5'- ^{32}P -end labeled 39-mer duplex (200,000 cpm) in 10 μl water was added to 190 μl DMS buffer (50 mM sodium cacodylate, 1 mM EDTA, pH 8.0). To the resulting solution was added 4 μl calf thymus DNA (1 mM base pair) followed by 5 μl 10% dimethylsulfate in absolute

ethanol (freshly prepared). After 5 minutes at room temperature, 5 μ l glycogen in water (20 mg/ml), 50 μ l DMS stop solution, and 750 μ l absolute ethanol were added. Following chilling at -20° C for 15 minutes and centrifugation at 14,000 rpm for 30 minutes, the supernatant was discarded. The resulting pellet was dried *in vacuo*. 100 μ l 10% aqueous piperidine was added, the solution heated to 90° C for 30 min., then dried *in vacuo*. Samples were then lyophilized twice from 50 μ l water and stored at -20° C until analyzed by high resolution polyacrylamide gel electrophoresis.

Kinetic Measurements: 2-Bromoacetamide Alkylation of the 39 Base Pair Duplex. A 2X stock solution was made containing 5'-³²P-end labeled 39 base pair duplex (8,000 cpm/ μ L), 40 mM Bis-Tris acetate (pH 7.0), and 1.6 mM cobalt hexammine trichloride. The stock solutions were then aliquoted into microfuge tubes and to each was added an equivalent volume of 2X aqueous 2-bromoacetamide (made fresh, volumetrically) at 20, 40, 60, 100, 120, 140, 160, 180 mM concentrations. Reactions were each then aliquoted to six microfuge tubes and incubated at 37° C for 25 minutes. The first tube (arbitrarily designated time equals zero) was quenched by brief exposure to liquid nitrogen, followed by the addition of 2 μ L glycogen (20 mg/mL), 2 μ L 3 M sodium acetate, and 60 μ L ethanol. Samples were stored at -80° C until completion of the experiment and subsequent time points were treated as described. The samples were then centrifuged at 14,000 rpm for 30 minutes, the supernatants discarded, and reactions dried *in vacuo*. 100 μ L of 10% aqueous piperidine was added and reactions were heated to 90° C for 30 minutes. Reactions were lyophilized, then lyophilized twice from 50 μ L water and dissolved in formamide loading buffer. Separation of products was performed by 1:20 cross-linked 15% denaturing polyacrylamide gel electrophoresis (10,000 cpm/lane). Dried polyacrylamide gels were analyzed by storage phosphor autoradiography.

Band densities were determined by quantitative analysis of storage phosphor autoradiography as described above. Bands were assigned as depicted in Figure 9 to determine values for "[D]_i" and "[D]_t" and "[D]_i/[D]_t" plotted versus time using the program KaleidaGraph to determine from the slope values for k_{obs} for two or three rate determinations at each concentration 2-bromoacetamide. An analysis of the data indicated that disappearance of fully intact DNA at the highest concentration of 2-bromoacetamide (90 mM) over three half-lives followed pseudo-first order kinetics. In subsequent determinations of observed rate constants, initial rates were used providing values of $3.4 (\pm 0.2) \times 10^{-7}$, $6.4 (\pm 0.8) \times 10^{-7}$, $9.6 (\pm 0.9) \times 10^{-7}$, $1.7 (\pm 0.2) \times 10^{-6}$, $2.2 (\pm 0.1) \times 10^{-6}$, $2.4 (\pm 0.01) \times 10^{-6}$, $2.4 (\pm 0.1) \times 10^{-6}$, and $3.4 (\pm 0.1) \times 10^{-6} \text{ s}^{-1}$ for 10, 20, 30, 50, 60, 70, 80, and 90 mM 2-bromoacetamide concentration, respectively. Average k_{obs} values were then plotted against concentration 2-bromoacetamide to determine, from the slope of the line a rate constant k_2 of $3.6 (\pm 0.3) \times 10^{-5} \text{ M}^{-1} \text{ s}^{-1}$ from which rates at very low concentration 2-bromoacetamide can be calculated.

References

- (1) (a) Hertzberg, R. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 313. (b) Hertzberg, R. P.; Dervan, P. B. *Biochemistry* **1984**, *23*, 3934. (c) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861. (d) Mack, D. P.; Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7572. (e) Mack, D. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 4604.
- (2) (a) Baker, B. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 8266. (b) Dervan, P. B.; Baker, B. F. *Ann. N. Y. Acad. Sci.* **1986**, *471*, 51. (c) Baker, B. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 2700.
- (3) (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci., U.S.A.* **1989**, *86*, 5723. (b) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783. (c) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci., U.S.A.* **1992**, *89*, 7586.
- (4) (a) Moser, H. E.; Dervan, P. B. *Science* **1987**, *238*, 645. (b) Strobel, S. A.; Moser, H. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7927. (c) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1989**, *111*, 3059. (d) Strobel, S. A.; Dervan, P. B. *Science* **1990**, *249*, 73. (e) Maher, L. J.; Dervan, P. B.; Wold, B. J. *Biochemistry* **1990**, *29*, 8820. (f) Plum, G. E.; Park, Y.-W.; Singleton, S. F.; Dervan, P. B.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 9436. (g) Strobel, S. A.; Dervan, P. B. *Nature* **1991**, *350*, 172. (h) Strobel, S. A.; Doucette-Stamm, L. A.; Riba, L.; Housman, D. E.; Dervan, P. B.

Science **1991**, 254, 1639. (i) Singleton, S. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, 114, 6957. (j) Best, G.C.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, 117, 1187.

(5) (a) LeDoan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.-L.; Thuong, N. T.; Lhomme, J.; Helene, C. *Nucleic Acids Res.* **1987**, 15, 7749. (b) Perrouault, L.; Asseline, U.; Rivalle, C.; Thoun, N. T.; Bisagni, E.; Giovannangeli, C.; Le Doan, T.; Helene, C. *Nature* **1990**, 344, 358. (c) Mergny, J.-L.; Sun, J.-S.; Rougée, M.; Montenay-Garestier, T.; Barcelo, F.; Chomilier, J.; Helene, C. *Biochemistry* **1991**, 30, 9791. (d) Rougée, M.; Faucon, B.; Mergny, J. L.; Barcelo, F.; Giovannangeli, C.; Garestier, T.; Helene, C. *Biochemistry* **1992**, 31, 9269

(6) (a) Fedorova, O. S.; Knorre, D. G.; Podust, L. M.; Zarytova, V. F. *FEBS Letters* **1988**, 228, 273. (b) Vlassov, V. V.; Gaidamakov, S. A.; Zarytova, V. F.; Knorre, D. G.; Levina, A. S.; Nikonova, A. A.; Podust, L. M.; Fedorova, O. S. *Gene* **1988**, 72, 313. (c) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, 112, 9428. (d) Shaw, J-P.; Milligan, J. F.; Krawczyk, S. H.; Matteucci, M. *J. Am. Chem. Soc.* **1991**, 113, 7765. (e) Young, S. L.; Krawczyk, S. H.; Matteucci, M. D.; Toole, J. J. *Proc. Natl. Acad. Sci., U.S.A.* **1991**, 88, 10023.

(7) (a) Povsic, T. J.; Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, 114, 5934. (b) Kutyavin, I. V.; Gamper, H. B.; Gall, A. A.; Meyer, R. B. *J. Am. Chem. Soc.* **1993**, 115, 9303.

(8) (a) E is bromoacetyl attached to the 5-position of thymine (6c) and ¹³C is 5-methyl-2'-deoxycytidine, previously shown to facilitate binding in oligonucleotide-directed triple helix formation (4c).

(9) Synthesis of the N-bromoacetyloligonucleotide was performed as described previously (6c).

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

(11) Cameron, V.; Uhlenbeck, O. C. *Biochemistry* **1977**, 16, 5120.

(12) Razzell, W. E.; Khorana, H. G. *J. Biol. Chem.* **1959**, 234, 2114.

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

(14) Hemminki, K.; Peltonen, K.; Vodicka, P. *Chem.-Biol. Int.* **1989**, 70, 289.

(15) Barrio, J. B.; Secrist III, J. A.; Leonard, N. J. *Biochem. Biophys. Res. Comm.*, **1972**, 46, 597.

(16) March, J. *Advanced Organic Chemistry, 3rd Ed.*; John Wiley & Sons, Inc.: New York, 1985.

(17) Errors are reported as standard deviation about the mean. Bevington, P. R.; Robinson, D. K. *Data Reduction and Error Analysis for the Physical Sciences, 2nd Ed.*; McGraw-Hill, Inc.: New York, 1990.

(18) Frost, A. A.; Pearson, R. G. *Kinetics and Mechanism, 2nd Ed.*; John Wiley & Sons, Inc.: New York, 1961; pp 40-46.

(19) It has been previously shown that products of cleavage can be ligated into plasmid DNA which are competent for replication in *E. coli* (7a).

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

(21) Iverson, B. L.; Dervan, P. B. *Methods in Enzymology* **1993**, 218, 222.

(22) Gait, M. J. *Oligonucleotide Synthesis: A Practical Approach*; IRL Press: Oxford, 1984; and references cited therein.

(23) (a) de Groot, N.; Lapidot, Y.; Panet, A.; Wolman, Y. *Biochem. Biophys. Res. Comm.* **1966**, 25, 17. (b) Pelligrini, M; Oen, H.; Cantor, C. R. *Proc. Natl. Acad. Sci., U.S.A.* **1972**, 69, 837.

(24) Zurlo, J.; Curphey, T. J.; Hiley, R.; Longnecker, D. S. *Cancer Research* **1982**, 42, 1286.

Chapter 2

Sequence Specific Alkylation of Double Stranded RNA Mediated by Oligonucleotide-Directed Triple Helix Formation

Introduction

This project was performed in collaboration with Jean-Francois Mouscadet, a postdoctoral fellow in the Dervan Group.

Chemical reagents for the sequence specific cleavage of RNA involve two mechanistic classes, oxidation of the ribose backbone (1a) and hydrolysis of the phosphodiester bond (2). These methods generally lead to a mixture of products, due to the inherent non-sequence specificity of diffusible oxidants and the flexibility of linkages to metal complexes from RNA-binding ligands for phosphodiester hydrolysis. Additionally, for phosphodiester hydrolysis cleavage of double stranded RNA is difficult relative to single stranded RNA, presumably by an increased barrier to rearrangement of the coordination sphere about phosphorus in the pentacoordinate transition state. Development of a new mechanistic class of reagents for the sequence-specific cleavage of double stranded RNA would improve on existing methodology for cleavage of double stranded RNA (Figure 2.1).

Oligonucleotide-directed triple helix formation is a powerful approach for the recognition of double-helical DNA (3). It has been shown that sequence-specific recognition of double stranded RNA and RNA-DNA hybrids can be achieved by oligoribonucleotide-directed triple helix formation, where a backbone composition of ribonucleotides is necessary for binding when the purine strand of the duplex is RNA (1).

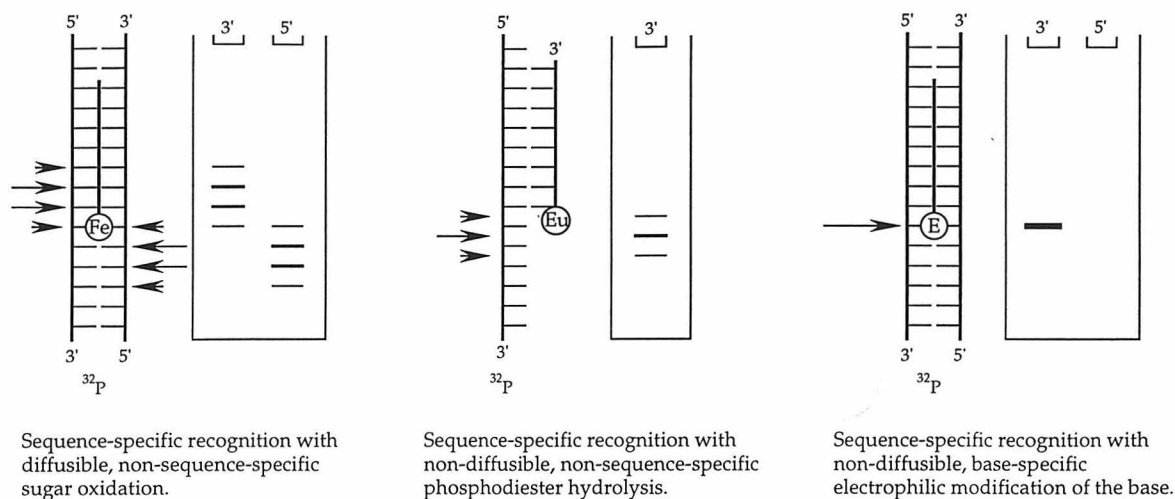


Figure 2.1. Different mechanistic classes of sequence-specific reagents for RNA cleavage afford different specificities.

Previous work has shown that an oligodeoxyribonucleotide can direct an electrophile to react at the N7 position of guanine in the major groove of duplex DNA by triple helix formation (4). Use of this system for directing electrophiles to react at N7 of guanine in the major groove of *double stranded RNA* would represent a new mechanistic class of sequence specific RNA cleaving agents (Figure 2.2).

Discussion

For development of a system for oligonucleotide-directed alkylation of double stranded RNA, synthesis of an oligoribonucleotide-electrophile conjugate is necessary because it has been shown that only RNA binds with reasonable affinity to duplex RNA (1). This required the synthesis of the phosphoramidite **8**, possessing the same base, linker

Figure 2.2. Binding of N-bromoacetyloligoribonucleotides to adjacent inverted triple helix sites would afford double strand alkylation and cleavage of double stranded RNA.

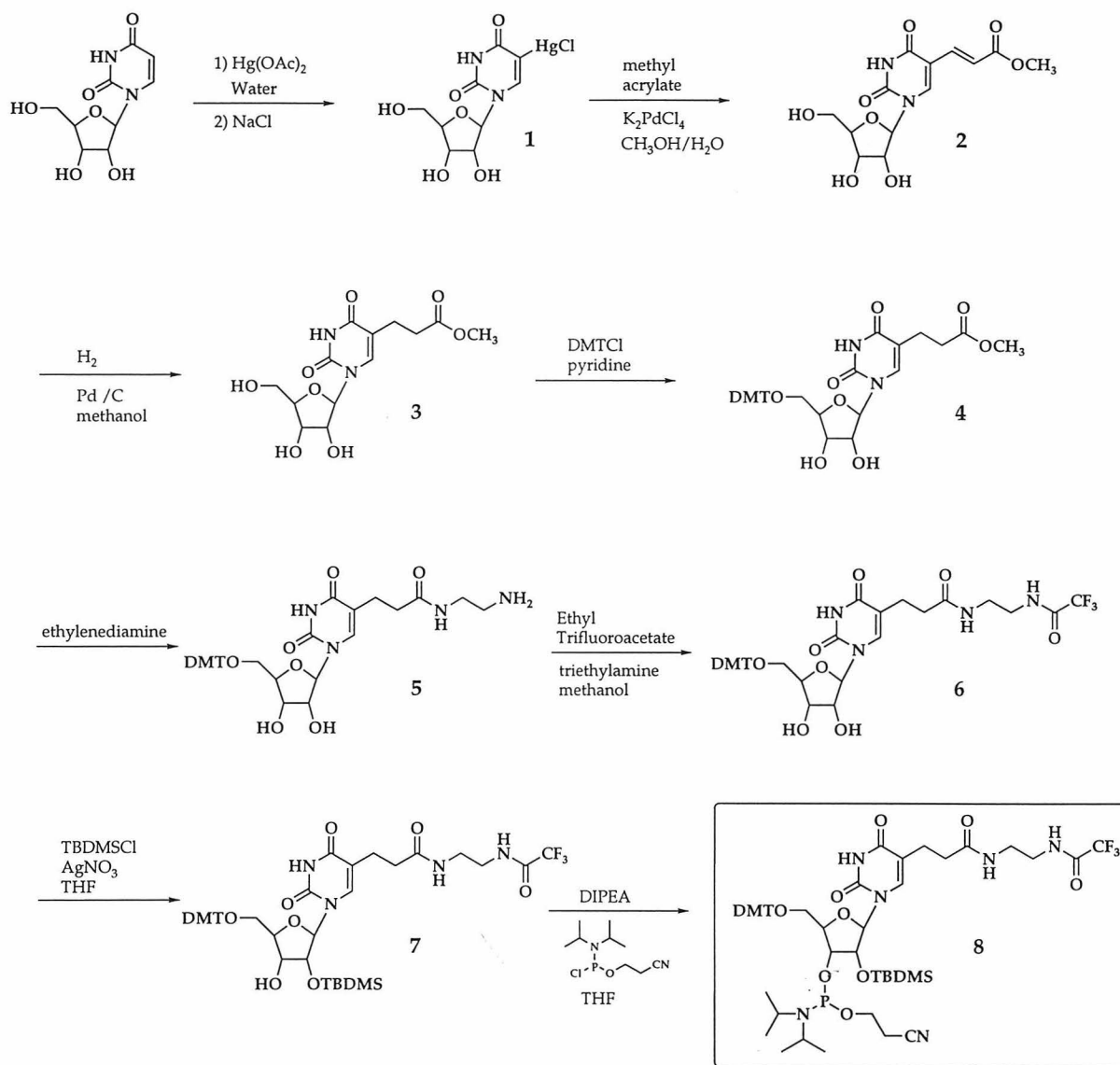
arm, and electrophile as that previously successful in triplex alkylation of duplex DNA. The phosphoramidite was synthesized in eight steps from uridine (Figure 2.3) (5).

Reaction of uridine with mercuric acetate, followed by precipitation as the mercurichloro-salt afforded **1** in 89% yield. Metal-mediated coupling of **1** with methyl acrylate in a Heck-type reaction afforded **2** in 18% yield following flash chromatography and recrystallization. Hydrogenation of **2** over palladium on carbon afforded **3** in 95% yield. Dimethoxytrityl-protection of the 5'-hydroxyl group was achieved in 54% yield. Amidation of the methyl ester with ethylenediamine, followed by removal of excess ethylenediamine by co-evaporation with a number of solvents, afforded **5** in 95% yield. Protection of the amino group of **5** with ethyl trifluoroacetate afforded an inseparable mixture of **6** and triethylammonium trifluoroacetate. Silylation of the 2'-hydroxyl group of **6** was achieved in 31% yield to provide **7** as the major isomer and the 3'-O-silyl protected material as the minor isomer (6). Phosphitylation of **7** with chloro-N,N-diisopropyl- β -cyanoethylphosphoramidite provided the 5'-O-dimethoxytrityl-2'-O-t-butyldimethylsilyl phosphoramidite **8** in 56% yield.

For synthesis of the RNA containing the amino group at the 5-position of uracil the phosphoramidite **8**, as a 0.15 M solution in acetonitrile, was coupled chemically to the RNA sequence to the 5'-end of 5'-UUUCUUUUUCUUUUCU-3' (Figure 2.4) (7). The resulting aminooligoribonucleotide was deprotected, then purified by denaturing polyacrylamide gel electrophoresis. Acylation of the aminooligoribonucleotide with N-hydroxysuccinimidyl 2-bromoacetate (8) provided the N-bromoacetyloligoribonucleotide which was used for oligonucleotide-directed alkylation of double stranded RNA by Jean-Francois Mouscadet, a postdoctoral fellow in the Dervan group.

A summary of the alkylation experiments on double stranded DNA is as follows (9). Site-specific alkylation of one strand of duplex RNA was achieved in reasonable yield

Figure 2.3. Synthesis of phosphoramidite **8** for oligoribonucleotide-directed alkylation of double stranded RNA was achieved in eight steps from uridine.



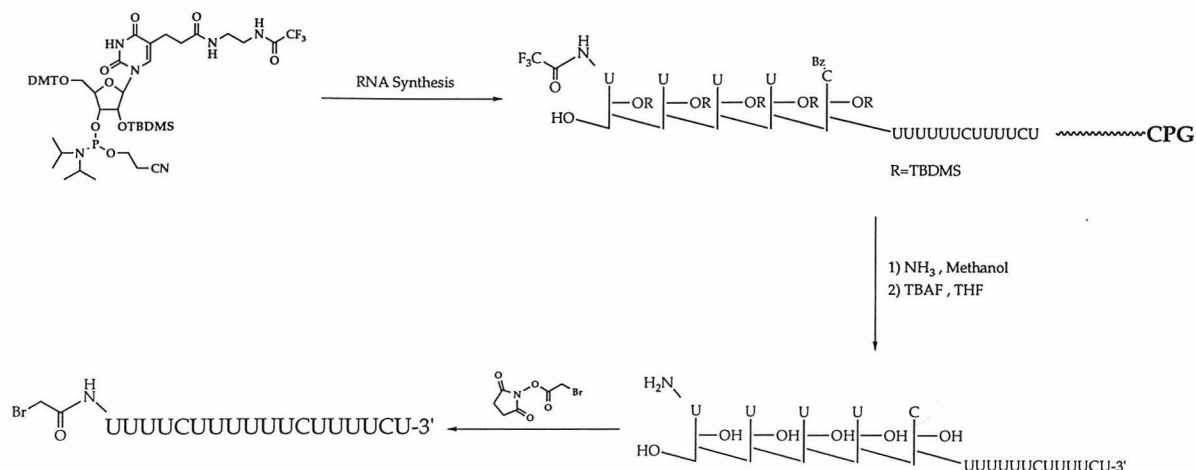


Figure 2.4. Incorporation of the amine-containing phosphoramidite into chemically synthesized RNA by standard methods followed by deprotection and acylation affords the N-bromoacetyl oligoribonucleotide.

(ca 60%). A stable crosslink was observed on denaturing polyacrylamide gels, likely due to the increased stability of the glycosidic bond of N7-alkylated guanine bases in ribonucleotides (10). Cleavage of this crosslink by sodium borohydride reduction and basic workup was inefficient, ultimately resulting in low cleavage yields.

Conclusion.

Oligonucleotide-directed alkylation and cleavage of double stranded RNA was achieved in low yield. The unanticipated complexity of this problem highlights the general observation that where DNA is structurally "well behaved" RNA is structurally diverse even in the context of a repeating double-helical structure. 2-Bromoacetamide alkylation experiments indicated that one problem may be lack of accessibility to the major groove by the electrophile (9). Solutions to the problem of major groove accessibility may involve

targeting of RNA substrates with a widened major groove. Crothers et al. have demonstrated that the effect of bulges in RNA helices is to widen the major groove (11). Use of stronger electrophiles for RNA alkylation should also be considered as an alternate and parallel solution. Underscored is the observation that, though problems of DNA recognition and reactivity are well understood, problems in RNA recognition and cleavage, though solvable, are non-trivial and will likely represent long-term investments.

Experimental

Materials and Methods. ^1H NMR and ^{13}C NMR spectra were recorded at 300 MHz on a General Electric QE 300 Spectrometer. Chemical shifts were recorded in parts-per-million using the proteo NMR solvent as a reference. All reagents were used without further purification unless otherwise noted. RNA synthesis was performed on an Applied Biosystems 394 DNA synthesizer. RNA phosphoramidites were purchased from Biogenex.

5-Mercurichlorouridine (1). Uridine (5.0 grams, 20.5 mmole) was dissolved in 30 ml water. To the resulting solution was added mercuric acetate (6.9 grams, 21.6 mmole). The white suspension was stirred for 20 hours at 55° C. A solution of sodium chloride (3 grams) in water (20 ml) was added, and the suspension dried *in vacuo*., then dried repeatedly with methanol and acetonitrile. The solid was collected by filtration from methanol (20 ml) and washed with four 20 ml portions methanol then diethyl ether (20 ml) to provide 8.7 grams (89%) product as a white solid.

Alkene (2). In 125 mL methanol was combined 5-mercurichlorouridine (5.0 g, 10.4 mmol) and freshly distilled methyl acrylate (4.5 g, 52.2 mmol). To the white suspension was added potassium tetrachloropalladate(II) (3.4 g, 10.4 mmol) in 30 mL water. The dark green solution was stirred for 2 hours under argon, then hydrogen sulfide was bubbled in for 3 min. Celite was added and the suspension filtered to give a clear solution, which was dried *in vacuo* to give a pale yellow solid. The product mixture was dried *in vacuo* from methanol and acetonitrile to remove water. 50 mL methanol was added and the solution decanted from the insoluble material. Silica gel was added to the solution and the suspension dried *in vacuo*. Flash chromatography in 0-12% methanol/dichloromethane provided 2.0 g (60%) of a yellow solid. As this material was

judged unsuitably pure by NMR, the product was recrystallized from methanol to provide 0.6 g (18%) of a white solid. ^1H NMR (d_6 -DMSO): δ 3.55-3.75 (m, 2H, 5'-CH₂), 3.67 (s, 3H, OCH₃), 3.84-3.86 (m, 1H, 4'-CH), 3.97-4.02 (dd, J = 5.1 Hz, J = 10.2 Hz, 1H, 2'-CH or 3'-CH), 4.04-4.08 (dd, J = 4.8 Hz, J = 9.2 Hz, 1H, 3'-CH or 2'-CH), 5.07-5.09 (d, J = 5.6 Hz, 1H, 2'-OH or 3'-OH), 5.29-5.32 (t, J = 5.0 Hz, 1H, 5'-OH), 5.45-5.47 (d, J = 5.1 Hz, 1H, 3'-OH or 2'-OH), 5.74-5.75 (d, J = 4.1 Hz, 1H, 1'-CH), 6.81-6.86 (d, J = 15.8 Hz, 1H, =CH), 7.31-7.36 (d, J = 15.8 Hz, 1H, =CH), 8.50 (s, 1H, ArH), 11.7 (s, 1H, NH). ^{13}C NMR (d_6 -DMSO): δ 51.9, 60.7, 69.6, 74.6, 85.2, 89.3, 108.8, 116.9, 138.5, 144.6, 150.1, 162.3, 167.8. IR (KBr) 3341 (m), 3059 (w), 1721 (s), 1690 (s), 1624 (m), 1474 (w), 1437 (w), 1295 (m), 1086 (m), 1049 (m), 986 (w).

Alkane (3). Alkene **2** (0.52 g, 1.6 mmol) was dissolved in 250 mL methanol and placed in a bomb with 110 mg 10% palladium on carbon. The reaction was stirred under 75 psi hydrogen for 5 hours. Celite was added, the reaction mixture filtered, and the solution dried *in vacuo* to provide 0.50 g (95 %) alkane **3** as a white foam. ^1H NMR (d_6 -DMSO): δ 2.4-2.6 (m), 3.6 (s, 3H), 3.6-3.7 (m, 1H), 3.9-4.2 (m, 2H), 5.1-5.3 (m, 2H), 5.4 (m, 1H), 5.8 (m, 1H), 7.8 (s, 1H), 11.4 (s, 1H). ^{13}C NMR (d_6 -DMSO): δ 23.3, 33.2, 52.3, 61.8, 70.8, 74.5, 85.8, 88.7, 112.8, 115.4, 138.0, 164.2, 173.5.

DMT alkane (4). Alkane **3** (0.47 g, 1.4 mmol) was dried *in vacuo* from pyridine (Sure-Seal). 3 mL freshly distilled pyridine was added under argon and the solution cooled to 0°C. Dimethoxytrityl chloride (0.53 g, 1.6 mmol) was added. The solution was brought to room temperature and stirred for 2 hours. Methanol (4 mL) was added and after 30 min the solution dried *in vacuo*. 5% sodium bicarbonate (15 mL) and dichloromethane (45 mL) were added, and the resulting organic layer extracted with water (15 mL), then dried over magnesium sulfate. The solution was dried *in vacuo* to a yellow foam. Flash

chromatography in 0-4% methanol/dichloromethane provided 0.49 g (54%) of **4** as a white foam. ^{13}C NMR (d_6 -DMSO): δ 15.4, 25.3, 44.4, 48.1, 55.7, 63.4, 68.1, 76.8, 79.7, 82.7, 106.1, 106.3, 119.9, 120.8, 120.9, 121.0, 123.0, 123.1, 128.2, 128.3, 129.8, 137.3, 144.1, 151.4, 151.5, 156.7, 165.6.

DMT-TFA-amine (6). DMT-alkane (0.45 g, 0.71 mmol) was dissolved in ethylenediamine (10 mL, freshly distilled from NaOH). The solution was stirred at room temperature for 40 hours, then dried *in vacuo*. The product was dried *in vacuo* from 20 mL dioxane, 30 mL toluene, 20 mL acetonitrile and 20 mL toluene, 20 mL acetonitrile, and 5 mL anhydrous methanol to afford **5** (445 mg, 95%) as a white foam. The amine (0.21 g, 0.32 mmol) and triethylamine (226 μL , 1.6 mmol) were dissolved in 5 mL anhydrous methanol. Ethyl trifluoroacetate (193 μL , 1.6 mmol) was added dropwise with stirring. The solution was stirred for 24 hours, then dried *in vacuo* to a yellow oil. The product was dissolved in 10 mL methanol and adsorbed onto silica gel. Flash chromatography two times in 0-10% methanol/dichloromethane provided 0.3 g of a white solid as a mixture of product and triethylammonium trifluoroacetate salt. ^1H NMR (d_6 -DMSO): δ 1.3 (m, many H's, triethylammonium), 2.1-2.2 (m, 4H), 3.1-3.3 (m, many H's, triethylammonium), 3.8 (s, 6H), 3.8-4.0 (m, 2H), 4.1 (q, 1H), 5.2 (d, 1H), 5.4 (d, 1H), 5.8 (d, 1H), 6.8-7.5 (m, 14H), 7.9 (t, 1H), 9.4 (t, 1H), 11.4 (s, 1H).

Silyl-protected nucleoside (7). DMT-TFA-amine **6** (*ca.* 200 mg, 0.27 mmol) and pyridine (80 mg, 1 mmol) were dissolved in 10 mL THF (freshly distilled from Na/benzophenone). Silver nitrate (54.5 mg, 0.32 mmol) and t-butyldimethylsilyl chloride (53.0 mg, 0.35 mmol) were added. The suspension was stirred for 24 hours, then 1 equivalent of silver nitrate and TBDMS-Cl were added. After 24 hours, the solution was decanted from the purple solid and 40 mL methylene chloride and 10 mL 5% sodium

bicarbonate were added. The organic layer was dried over magnesium sulfate, filtered, and dried *in vacuo* to a yellow foam. Flash chromatography from 0-10% methanol/dichloromethane provided a major 2'-O-silyl protected product (73 mg, 31%) and a minor 3'-O-silyl protected product (*ca.* 10 mg) as well as remaining starting material. ^1H NMR (d_6 -DMSO): δ 0.01 (s, 3H), 0.03 (s, 3H), 0.82 (s, 9H), 2.05-2.18 (m, 4H), 3.08-3.29 (m, 4H), 3.72 (s, 6H), 3.77-3.98 (m, 2H), 4.22 (t, $J = 5.2$ Hz, 1H), 5.07 (d, $J = 6.1$ Hz, 1H), 5.81 (d, $J = 5.3$ Hz, 1H), 6.87-6.90 (m, 4H), 7.18-7.40 (m, 9H), 7.46 (s, 1H), 7.86 (t, $J = 5.1$ Hz, 1H), 9.39 (t, $J = 5.0$ Hz, 1H), 11.4 (s, 1H). ^{13}C NMR (d_6 -DMSO): δ -4.7, -4.3, 18.5, 23.3, 26.1, 34.6, 37.9, 39.6, 55.5, 64.1, 70.6, 75.6, 83.8, 86.4, 88.3, 113.7, 113.8, 114.6, 118.3, 127.3, 128.1, 128.4, 130.3, 135.7, 135.8, 136.7, 145.3, 151.0, 156.3, 156.7, 157.2, 158.7, 163.7, 171.8.

Phosphoramidite (8). Silyl-protected nucleoside **7** (87 mg, 99.4 μmol) was transferred to a flask with acetonitrile, then dried *in vacuo*. THF (800 μL , distilled from Na/benzophenone) and diisopropylethylamine (40 μL , 398 μmol) were added via syringe. Chloro-N,N-diisopropyl- β -cyanoethylphosphoramidite (21 μL , 129 μmol) was added dropwise via syringe. After one hour a second aliquot was added. After one more hour, the solution was dried *in vacuo*. Flash chromatography with 0-1-0.5% methanol/dichloromethane/triethylamine afforded 60 mg (56%) of **8** as mixture of diastereomers. ^1H NMR (CDCl_3): δ 0.1-0.3 (m), 0.7-0.9 (m), 1.0-1.1 (m), 1.2-1.3 (m), 1.3-1.4 (m), 1.9-2.2 (m), 2.4 (m), 2.7 (m), 3.2 (m), 3.3-3.5 (m), 3.5-3.7 (m), 3.8 (m), 4.2-4.3 (m), 4.3-4.6 (m), 6.0-6.2 (m), 6.3-6.5 (m), 6.7-6.9 (m), 7.3-7.4 (m), 7.4-7.5 (m), 7.7 (m), 8.3 (m).

Synthesis and Purification of the N-Bromoacetyloligoribonucleotide. The phosphoramidite was dissolved in acetonitrile (Sure Seal) to 0.15 M concentration and

coupled to the 5'-end of the RNA sequence 5'-UUUCUUUUUUCUUUUCU-3' by standard phosphoramidite chemistry (7). Coupling efficiency was good (60-70%). The RNA was deprotected with ammonia saturated methanol at room temperature for 24 hours. Following lyophilization, the t-butyldimethylsilyl groups were removed by treatment with 1M tetrabutylammonium fluoride in tetrahydrofuran for 24 hours at room temperature. 100 mM HEPES pH 7.4 was added and samples were lyophilized. Ten mL water was added and samples filtered (Centrex) to remove traces of CPG, and the eluent was dialyzed against 0.5 mM TE. RNA solutions were then lyophilized. Preparative denaturing polyacrylamide gel electrophoresis, excision of the band corresponding to RNA-amine, electroelution, and elution through NAP-25 columns provided 57 nmol RNA-amine ($\epsilon = 170.1 \text{ mL}/\mu\text{mol}\cdot\text{cm}$). Five nmol aliquots were lyophilized and stored at -80°C .

To a 5 nmol aliquot of RNA-amine was added 5 μL 200 mM sodium borate buffer pH 8.9. To the resulting solution was added 5 μL 30 mM N-hydroxysuccinimidyl 2-bromoacetate in DMF. After mixing for 1 min, 490 μL water was added and the solution passed through a NAP-5 column. One mL of eluent was collected and stored at -20°C . Analytical denaturing polyacrylamide gel electrophoresis, performed by J.-F. Mouscadet, verified that the acylation reaction was quantitative.

References.

- (1) (a) Han, H.; Dervan, P. B. *Proc. Natl. Acad. Sci. USA*, **90**, 3806 (1993). (b) Roberts, R. W.; Crothers, D. M. *Science*, **258**, 1463 (1992).
- (2) (a) Hall, J. Husken, D., Pieleles, U., Moser, H.E., Haner, R. *Current Biology* **1**, 185-190 (1994). (b) Magda, D., Miller, R.A., Sessler, J.L. and Iverson, B.L. *J. Am. Chem. Soc.* **116**, 7439-7440 (1994).
- (3) Moser, H. E.; Dervan, P. B. *Science* **238**, 645 (1987).
- (4) (a) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.* **112**, 9428 (1990) (b) Povsic, T. J.; Strobel, S. A.; Dervan, P. B. *J. Am. Chem. Soc.* **114**, 5934 (1992).
- (5) Dreyer, G. B.; Dervan, P. B. *Proc. Natl. Acad. Sci., U.S.A.* **82**, 968 (1985).
- (6) (a) Corey, E. J.; Venkatateswarlu, A. *J. Am. Chem. Soc.*, **94**, 6190. (1972). (b) Hakimelahi, G. H.; Proba, Z. A.; Obilvie, K. K. *Tetrahedron Lett.*, **22**, 5243 (1981).
- (7) Gait, M. J. *Oligonucleotide Synthesis: A Practical Approach*; IRL Press: Oxford, 1984; and references cited therein.
- (8) Taylor, M. J.; Dervan, P. B. *J. Am. Chem. Soc.*, submitted.
- (9) Jean-Francois Mouscadet, unpublished observations.
- (10) Hemminki, K.; Peltonen, K.; Vodicka, P. *Chem.-Biol. Int.* **70**, 289 (1989).
- (11) Weeks, K. M.; Crothers, D. M.; *Science*, **261**, 1574 (1993).

Chapter 3

X-Ray Crystallography of Triple-Helical Complexes of the Pyrimidine•Purine•Pyrimidine Type

Introduction

A collaborative effort between the laboratories of Professor Peter B. Dervan and Professor Douglas Rees at Caltech involved the attempted structural characterization of triple-helical complexes of the pyrimidine•purine•pyrimidine type using x-ray crystallography. Though structural information is available for these complexes from both chemical methods, such as footprinting (1) and affinity cleavage (2), and spectroscopic methods, such as NMR (3), fiber diffraction (4), and IR (5), as of yet no high resolution crystal structure has been reported.

Discussion

For a description of the x-ray crystallography experiments mentioned in this Chapter, the reader is referred to: Kim, K.-H. Ph.D. Dissertation, California Institute of Technology, pp 101-116, 1995.

Section I. Among the first of our earliest efforts in the x-ray crystallography of triple-helical complexes were intermolecular complexes of the pyr•pur•pyr type (YM1-YM5).

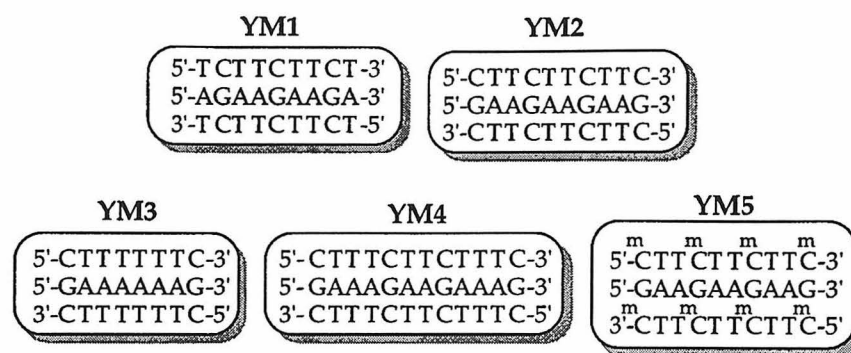


Figure 3.1. Complexes YM1-YM5 were among the first attempts in crystallography experiments.

Arbitrary design elements included i) varying length and sequence composition, ii) an element of symmetry to reduce sample preparation to two oligonucleotides per complex, and iii) blunt-ended complexes. Crystallization trials performed by Leemor Joshua-Tor in the Rees Group yielded only microcrystals unsuitable for x-ray diffraction experiments.

A second early approach involved the crystallization of intramolecular complexes A-G including pyrimidine•purine•pyrimidine (prepared by Scott Singleton), purine•purine•pyrimidine (prepared by Peter Beal), and mixed

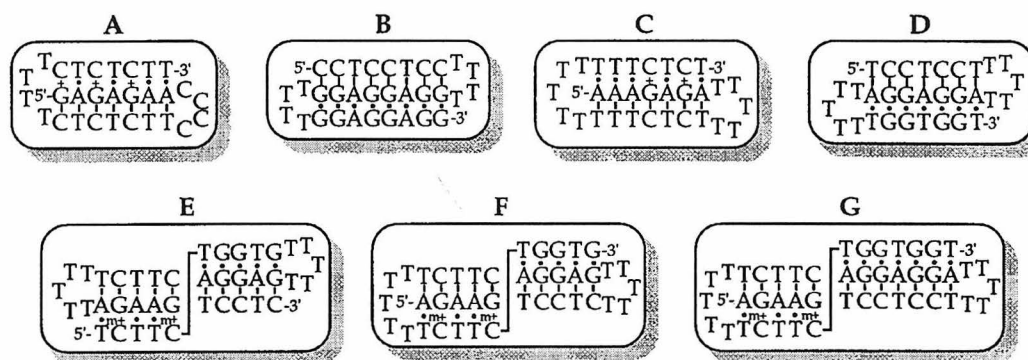


Figure 3.2. Intramolecular triple-helical complexes were prepared in early approaches to crystallography experiments.

pyrimidine•purine•pyrimidine-purine•purine•pyrimidine (prepared by Peter Beal) type triple helices. Incorporating elements known by NMR to favor formation of these structures (i. e. loop size) (3), these molecules were synthesized and crystallization trials attempted. Crystallization trials performed in the Rees Group yielded only microcrystals unsuitable for x-ray diffraction experiments.

Because of the failure of these early experiments, it was decided to explore a more "rational" approach where specific design elements were incorporated to facilitate crystallization. Complexes YM6-YM9 were prepared.

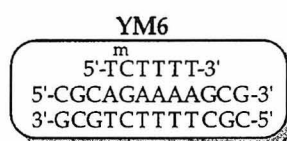


Figure 3.3. YM6 provides duplex DNA termini identical to that in drug-DNA co-crystal structures (6).

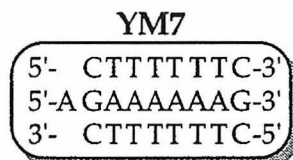


Figure 3.4. YM7 provides an overhanging residue to potentially increase intermolecular contacts in the crystal.

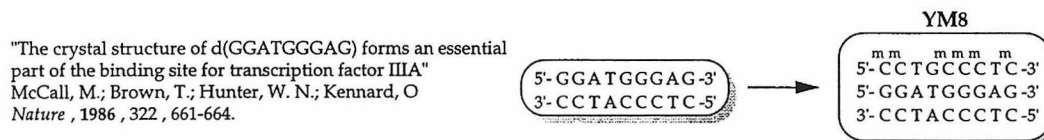


Figure 3.5. In YM8, the duplex from the known crystal structure of the TFIIIA transcription factor binding site provides a recognition site for pyrimidine motif triple helix formation.

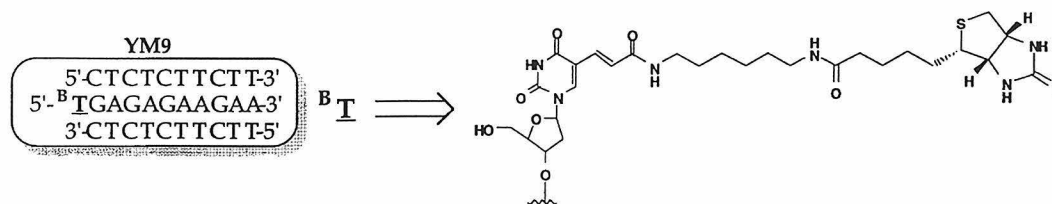


Figure 3.6. In YM9, incorporated is an overhanging thymidine nucleotide with a biotin moiety at the 5-position for co-crystallization trials with avidin and streptavidin proteins.

Attempts YM6-YM9 all failed to give crystals of the quality necessary for x-ray diffraction studies.

Section 2. A more systematic approach was then attempted based conceptually on the work of Schultz et al. (7).

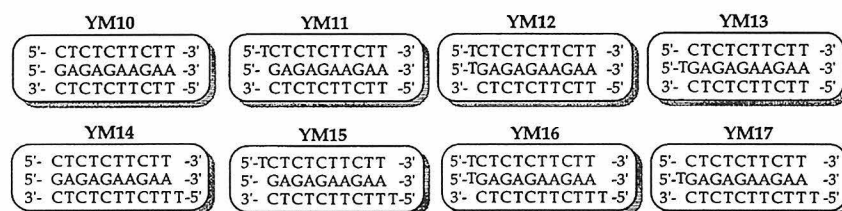


Figure 3.7. A "mix-and-match" approach provides more triple-helical structures for crystallography trials from a small set of oligonucleotides.

A "mix and match" approach to more rapidly screen structure space where, building off of a core sequence of ten base triplets, one base overhangs were incorporated to provide eight structures from six oligonucleotides.

Complexes YM10-YM17 were prepared and crystallization trials performed by Kyoung-Hee Kim in the Rees Group. For the first time, x-ray diffraction quality crystals

were obtained. Some diffracted to high resolution, but the data sets were incomplete (8). At this point, experiments were performed to determine the nature of the complex present in the crystal (i. e. duplex or triplex). A sample of YM12 crystals was provided by Kyoung-Hee Kim in the Rees Group and subjected to analysis by HPLC and denaturing polyacrylamide gel electrophoresis. Results from both techniques were consistent with the presence of triple-helical complexes in the crystal (Figure 3.8).

For complex YM11, preliminary data sets were consistent with a unit cell containing approximately six molecule stacked end-to-end (8). In an attempt to make a rational, incremental change to reduce the unit cell size, complex YM18 was prepared where two TAT triplets were inserted into the YM11 complex.

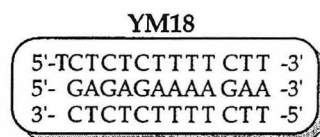


Figure 3.9. Complex YM18 results from the addition of two TAT triplet to complex YM11, in an attempt to reduce the unit cell size.

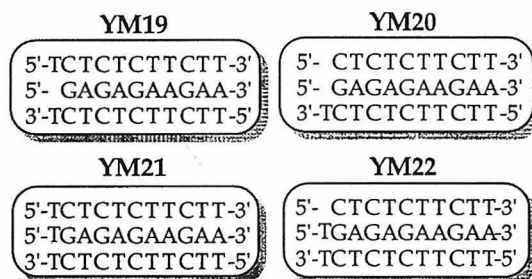
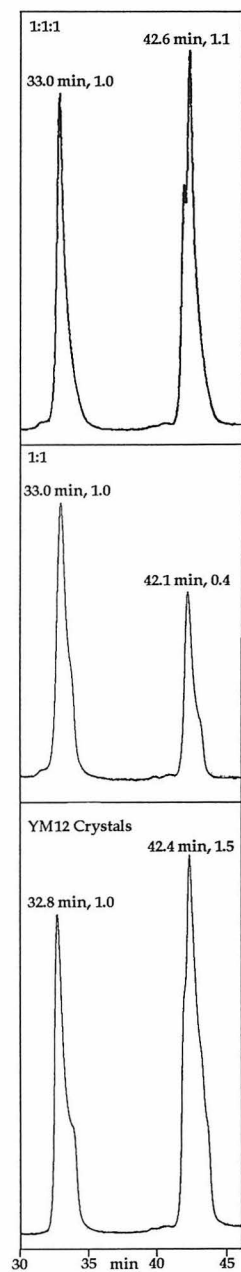


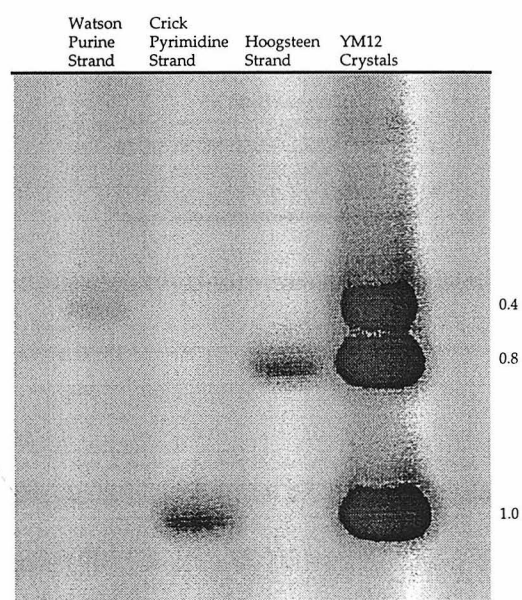
Figure 3.10. Complexes YM19-YM22 were prepared to complete the series of 5'-overhangs.

Figure 3.8. Characterization of crystals obtained from sample YM12. (A) Reversed-phase HPLC analysis of standards of triplex (top) and duplex (middle) DNA and YM12 crystals (bottom) indicates the presence of complexes consistent with triplex DNA. Peaks are annotated with retention time and relative peak area. (B) PAGE analysis of a sample of crystals treated with T4-polynucleotide kinase and γ - ^{32}P -ATP. Shown are standards of all three strands, indicating the presence of all three strands in a 1:1:1 ratio in the crystal. Bands are annotated with relative intensity.

A



B



This change would be predicted to reduce the unit cell size to one molecule, assuming a helical repeat size of B-form DNA in an end-to-end stacked continuous helix. Results from diffraction experiments with YM18 were consistent with fiber-type diffraction, previously observed by others (9), overlaid with a more ordered pattern giving diffraction to high resolution. No solution of this structure was possible.

From the above study it seemed that good quality crystals were only obtained from triple-helical complexes having blunt-ends on the 3'-end of the duplex homopyrimidine strand of the complex (YM10-YM13) (8). To complete the series of one-base thymine overhangs on the 5'-side, complexes YM19-YM22 were prepared. None, however, yielded satisfactory crystals for diffraction experiments.

To assist in solution of the data sets collected for complexes previously crystallized, bromine and iodine-containing analogs of complexes YM11 and YM18 were prepared.

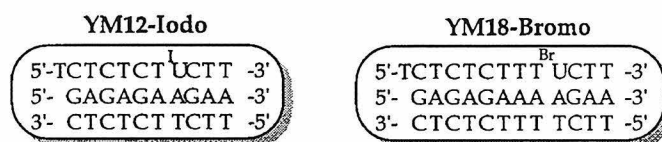


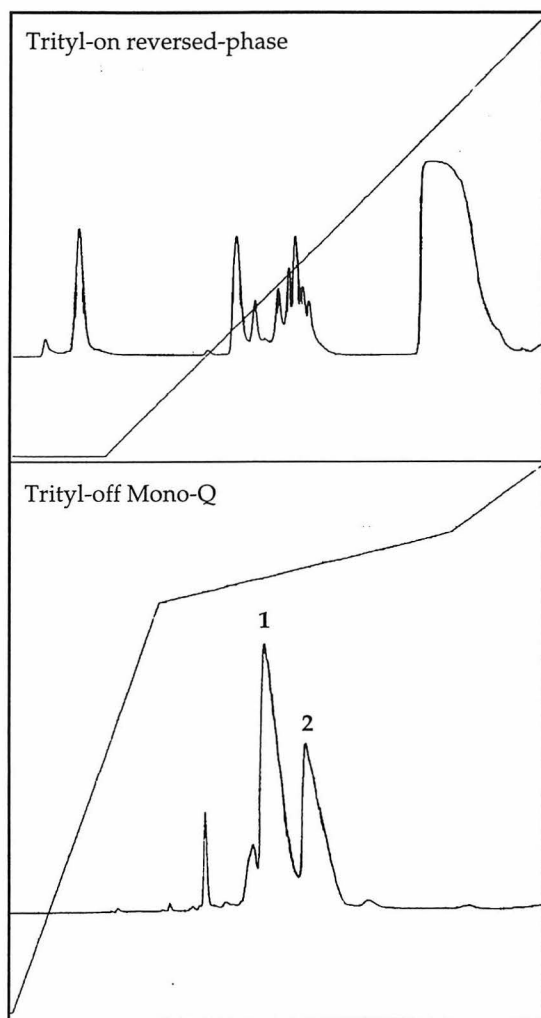
Figure 3.11. 5-Bromo- and 5-Iodo-2'-deoxyuridine-containing triple-helical complexes for x-ray crystallography trials.

Because it is known that halogens at the 5-position of uracil facilitate triple helix formation when in the Hoogsteen strand (10), 5-bromo-2'-deoxyuridine and 5-iodo-2'-deoxyuridine-containing triplexes were prepared. Brominated oligonucleotides were synthesized and purified by standard methods, but the iodinated oligonucleotides required special consideration.

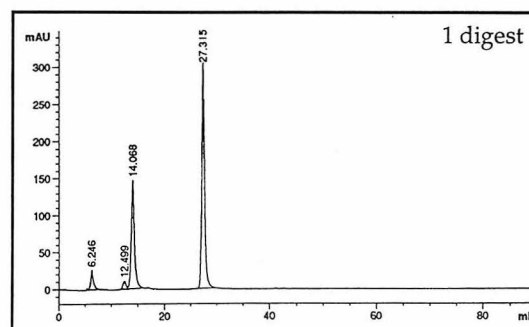
Trityl-on reversed-phase FPLC purification of 5-iodo-dU-containing oligonucleotides provide results similar to those for unmodified oligonucleotides. Trityl-off

Figure 3.12. Purification and characterization of the 5-iodo-2'-deoxyuridine-containing oligonucleotide 5'-TCTCTCT^IUCTT-3'. Chromatograms were monitored at 260 nm. Trityl-on reversed-phase FPLC purification, followed by trityl-off Mono-Q ion-exchange FPLC purification yields two major products (A). Enzyme digestion with SVP/CAP then demonstrates the absence (B) and presence by coinjection (C) of the 5-iodo-dU nucleoside in products 1 and 2, respectively.

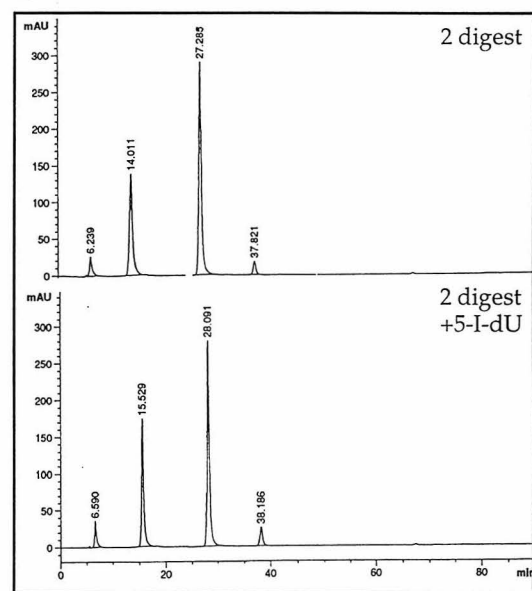
A



B



C



reversed-phase FPLC purification, while sufficient for unmodified oligonucleotides, does not resolve oxidation products from intact material when 5-iodo-dU oligonucleotides are used (11). An ion-exchange Mono-Q FPLC purification was therefore required at the trityl-off stage, providing two major products. Subsequent HPLC analysis of oligonucleotides digested with snake venom phosphodiesterase and calf alkaline phosphatase then allowed for identification of the correct product from the presence of the 5-iodo-dU nucleoside (Figure 3.11). Crystallization trials with the halogenated complexes yielded no suitable crystals, and attempts to purify oligonucleotides containing two 5-iodo-dU nucleotides failed to produce quantities sufficient for crystallization trials.

Conclusions

Though no high resolution x-ray crystal structure was obtained from these studies, recent results demonstrated that good progress could be made with a relatively small library of triple-helical complexes (~20) in that some of the later complexes gave single crystals diffracting to high resolution. "Semi-rational" approaches seemed to indicate that as much or more progress could be made by a more random approach. Based on this work, it is likely that a non-standard approach involving novel triple-helical complexes (12, 13) will be necessary for the solution to this elusive problem.

Experimental

Radioactive nucleotides were purchased from ICN. T4 Polynucleotide Kinase was obtained from Boehringer Mannheim. Biochemical manipulations were performed by standard methods (14). Snake venom phosphodiesterase (type VIII) was purchased from Sigma. High pressure liquid chromatography was performed on a Hewlett-Packard 1090 liquid chromatograph with a HPLC 3D Chemstation. Analytical reversed phase HPLC was performed on a Vydac 201HS Reversed Phase HPLC column. Oligonucleotide purification was accomplished on a Pharmacia LKB FPLC using a ProRPC HR 16/10 reversed phase column (160 mm $\times$ 10 mm) or a Mono-Q 10/10 ion-exchange column. Cation exchange cellulose (Cellex-CM) was purchased from BioRad. SepPak cartridges were purchased from Waters.

Oligonucleotides were synthesized on a ten μ mol scale on an Applied Biosystems Model 380B DNA Synthesizer using standard phosphoramidite chemistry and deprotected with 2 mL concentrated NH_4OH at 55° C for 24 hours. Following lyophilization, purification of the 5'-O-dimethoxytrityl-oligonucleotide was achieved using reversed phase FPLC (solvent A: 100 mM ammonium acetate, pH 7.0; solvent B: 40% acetonitrile in solvent A). Following lyophilization of appropriate fractions and removal of the dimethoxytrityl protecting group (80% acetic acid for 20 min followed by lyophilization), purification of the deprotected oligonucleotide was achieved by reversed phase FPLC. Following purification, oligonucleotides were desalted by two different methods. For YM1-YM9, oligomers were desalted and exchanged to sodium form on a Sep-Pak cartridge. For the remaining complexes, oligomers were desalted by Sep-Pak and passed through cation exchange cellulose equilibrated with sodium chloride. Oligomers were then quantitated by UV-visible spectroscopy (extinction coefficients in $\text{M}^{-1}\text{cm}^{-1}$: dT=8,700,

dG=11,500, dA=15,400, dC=7,400), aliquoted, lyophilized, combined at 24 mg/mL and 50 mM sodium cacodylate buffer, heated to 90°C for 8 min, cooled slowly to room temp, then sent to the Rees group for crystallography trials.

For 5-iodo-2'-deoxyuridine-containing oligonucleotides, trityl-on purification was performed as described. Trityl-off purification was performed on an FPLC Mono-Q 10/10 ion-exchange column (solvent A: 10 mM sodium hydroxide, pH 11; solvent B: 0.75 M sodium chloride/10 mM sodium hydroxide, pH 11). The eluent containing the product was adjusted to pH 7 with acetic acid and passed through a C-18 reversed-phase SepPak cartridge. Triple-helical complexes containing this oligomer were treated as previously described.

HPLC analysis of enzymatic digests was performed as follows: Oligonucleotide digests (2.5 μ L 10x dephosphorylation buffer, 2.5 μ L 2 M $MgCl_2$, 18 μ L oligomer, 1 μ L calf alkaline phosphatase, and 2 μ L snake venom phosphodiesterase) were injected in 23.5 μ L aliquots - solvent A, 100 mM ammonium acetate, pH 6.0; solvent B, acetonitrile; 0-60 min, 0-10% B; 60-85 min, 10-20%; 80-90 min, 20-0% B. Chromatograms were monitored at 260 nm.

References

- (1) Best, G.C.; Dervan, P. B. *J. Am. Chem. Soc.* **117**, 1187 (1995).
- (2) Moser, H. E.; Dervan, P. B. *Science* **238**, 645 (1987).
- (3) (a) Rajagapol, R.; Feigon, J. *Nature*, **339**, 637 (1989). (b) Live, D. H.; Radhakrishnan, I.; Misra, V.; Patel, D. J. *J. Am. Chem. Soc.*, **113**, 4687 (1991).
- (4) Arnott, S.; Selsing, E. *J. Mol. Biol.*, **88**, 509 (1974).
- (5) Liquier, J.; Coffinier, P.; Firon, M.; Taillandier, E. *J. Biomol. Struct. Dyn.*, **9**, 457 (1991).
- (6) Coll, M.; Frederick, C. A.; Wang, A., H.-J.; Rich, A. *Proc. Natl. Acad. Sci. USA*, **84**, 8385 (1987).
- (7) Schultz, S. C.; Shields, G. C.; Steitz, T. A. *J. Mol. Biol.*, **213**, 159 (1990).
- (8) Kyoung-Hee Kim, personal communication
- (9) Liu, K.; Miles, H. T.; Parris, K. D.; Sasisekharan, V. *Nature Structural Biol.*, **1**, 11 (1994).
- (10) Povsic, T. J.; Dervan, P. B. *J. Am. Chem. Soc.*, **111**, 3059 (1989).
- (11) Leemor Joshua-Tor, personnel communication.
- (12) Vanmeervelt, L.; Vlieghe, D.; Dautant, A.; Gallois, B.; Precigoux, G.; Kennard, O. *Nature*, **374**, 742 (1995).
- (13) Betts, L.; Josey, J. A.; Veal, J. M.; Jordan, S. R. *Science*, **270**, 5243 (1995).
- (14) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*: Cold Spring Harbor Laboratory: Cold Spring Harbor, N. Y., 1989.