SEQUENCE SPECIFIC DNA RECOGNITION AND BENDING

Thesis by

David A. Liberles

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

California Institute of Technology

Pasadena, California

1998

(Submitted June 16, 1997)

To my father

ACKNOWLEDGMENTS

I would like to thank my advisor, Professor Peter B. Dervan, for his guidance of my project and support of my research goals. His insight and vision have been crucial in allowing me to accomplish what I have. He has also created a very intense educational environment through which I have learned a lot not only about science, but also about how research is properly conducted. The people in the Dervan group have also contributed heavily to this stimulating educational environment.

I would also like to thank my committee. My interactions with Judy Campbell, Barbara Imperiali, Bill Goddard, and Steve Mayo have been very positive and helpful. In addition to helpful advice and suggestions, I am grateful to Professor Campbell for the opportunity to have worked on a collaborative project with her laboratory. I thank Professor Goddard for allowing me to use his computer facilities for the modeling studies presented here. I also enjoyed TAing for Professor Imperiali and am thankful for helpful comments. I thank Professor Mayo for substituting on my thesis defense committee at the last minute.

Beyond my advisor and committee, Caltech is a tremendous place to do science. I have had many educational opportunities with people in chemistry, biology, and other departments. I have not only learned about many different research projects in different areas of science, but I have had the opportunity to discuss DNA bending with a wide array of people. Research in this thesis has been directly impacted by many people in the Dervan group. In particular, I wish to thank Ken Brameld for teaching me how to model DNA and Eldon Baird for synthesizing the polyamide used in chapter 5.

I want to thank those who took the time to read various parts of this thesis. Sue Swalley, Jeff Albert, John Trauger, Scott Carter, Paul Floreancig, Kathryn Grant, and Pernilla Wittung all helped through critical reading of this document. Eldon Baird was helpful in preparing slides for my thesis defense.

I also want to thank everyone who was friendly and supported me through the ups and downs- the people I ran with, the people I skiied with, the people I climbed with, the people I drank with, and everyone else. My parents and brother were also very supportive the whole way through. My father had a lot of input on dealing with the trials and tribulations of graduate school, although it sounds like getting a Ph.D. in the late 1950s was very different!

Finally, thank you for reading this thesis.

ABSTRACT

Intrinsic DNA curvature and protein induced DNA bending play a crucial role in the regulation of many biological processes, including transcription, replication, transposition, viral integration, site-specific recombination, and DNA packaging in nucleosomes. The design of ligands to artifically regulate DNA topology may be important in creating artificial regulators of these processes, with potential utility in biology and human medicine. Much of the work described in this thesis is based upon the ability of a triple helix forming oligonucleotide to recognize double helical DNA. This sequence specific recognition occurs in two motifs and is governed by the formation of specific hydrogen bonds. One limitation of this technique is the inability to recognize pyrimidine bases in the Hoogsteen bonded strand. A novel base, D_3, has been previously described with the potential to recognize pyrimidine bases.

Chapter 2 describes a thermodynamic analysis of the binding properties of oligonucleotides bearing single and multiple D_3 bases in the context of triple helix formation. An NMR structure has indicated the binding mode of D_3 as intercalative and the ability of D_3 containing oligonucleotides to bend DNA through this intercalative wedge was examined.

Chapter 3 describes an alternative experimental approach to triple helix mediated DNA bending. Two third strand binding domains are appended by a linker of variable size to target two cognate binding sites separated by 10 base pairs, or one turn of the helix. Using this experimental approach, large bend angles were obtained, as analyzed electrophoretically.

Chapter 4 analyzes the nature of this binding and bending event using the designed third strand ligand. The energetics of bending were examined as were other factors influencing the ability to form a bent structure, including oligonucleotide binding motif, third strand length, sequence composition, pH, ion concentration and

valence, target site spacing, and linker composition and orientation to examine the effects of electrostatics and hydrophobicity.

Chapter 5 describes the design of a complementary architectural factor to straighten DNA that has been bent by a triple helical ligand. These complementary polyamide ligands recognize DNA sequence specifically and bind in the minor groove of the linker region to straighten DNA. Such molecules may also be useful as artificial regulators of biological processes.

TABLE OF CONTENTS

LIST OF FIGURES AND TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 5

CHAPTER 1

INTRODUCTION

BIOLOGICAL IMPORTANCE OF DNA

Deoxyribonucleic acid (DNA), genetic material in every living cell, is the cornerstone of life. DNA provides a stable copy of all of the information required for an organism to survive and reproduce more copies of itself and form future generations. The central dogma of molecular biology holds that information in DNA is transferred through RNA to function in proteins, which allow for propagation of the information encoded in DNA. The discovery of reverse transcriptase added an arrow to the central dogma, allowing the transfer of information from RNA back to DNA. Maintenance of this information involves DNA polymerization from DNA and in RNA viruses, RNA polymerization from RNA. This is depicted in Figure 1. Many specific functions are carried out on DNA in the reading, copying, and maintaining of its information. In metazoa, DNA is transcribed into RNA by RNA polymerases and is replicated by DNA polymerases. Other proteins repair damage to DNA, recombine chromosomal regions, splice DNA in T cells, align chromosomes during meiosis and mitosis, and package DNA in processes that are central to the survival of metazoa (1).

While prokaryotic chromosomes are circular and lack histones, eukaryotic chromosomes are segmented linear pieces of DNA supercoiled on electropositive histones. Metazoa contain diploid chromosomes, while ploidy can be highly variable in plants and other species. The importance of DNA to living systems points towards the importance of understanding nucleic acids at a chemical and physical level. This understanding may allow for the design of artificial regulators of DNA functions and DNA binding when malfunctions occur that affect the viability of living systems.

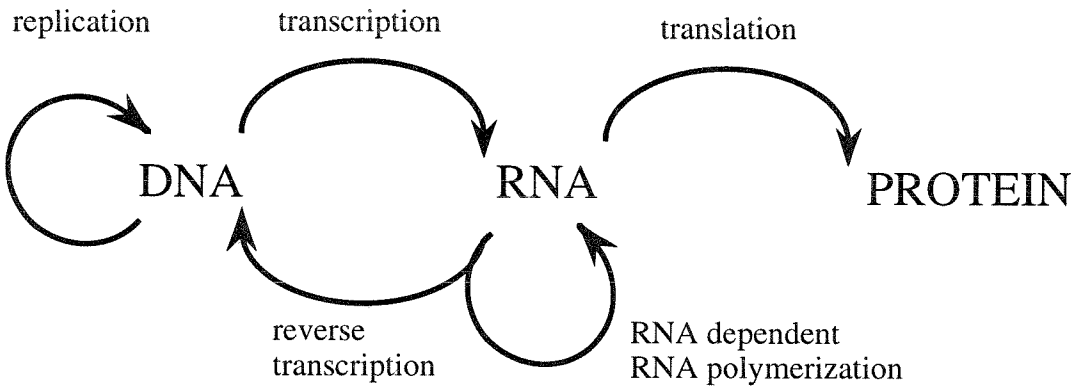


Figure 1. The central dogma of molecular biology.

DNA STRUCTURE

DNA is a polymer composed of repeating 2'-deoxyribose units linked to a purine or pyrimidine heterocycle at its 1' position (2). Deoxyribose units are linked 3' to 5' by phosphate linkages, which serve to stabilize the polymer by repulsion of anionic nucleophiles by the negative charge (3). The four bases of DNA have two or three functional groups available for hydrogen bonding, which allow formation of a regular intermolecular dimer that adopts a right handed antiparallel helical structure in its most common physiological form, B form. The hydrogen bonding patterns are depicted in Figure 2, where purine-pyrimidine pairings are always adenosine (A) with thymidine (T) and guanosine (G) with cytidine (C). Other functional groups are conceivable in this purine-pyrimidine scaffold and many of these have been synthesized and shown to lack the stability and regularity seen in the four natural bases (4). The right handed B form helix forms with the heterocyclic bases on the inside stacking upon each other. This helix shows two distinct grooves on the edge of the bases between the phosphate backbones, a major groove which is shallow and wide, and a deep and narrow minor groove between the phosphates. The helical parameters are 10.4 base pairs per helical turn, with a 34 Å rise per turn and a 36° rotation per residue. A model of this structure is depicted in Figure 3. There exists some heterogeneity in this structure due to many factors including solvation, functional group size and bonding preference, electrostatics, number of hydrogen bonds, and base stacking. Because of the different electrostatic potentials of different atoms within the bases, the optimal stacking arrangement of the bases is variable, resulting both in different stacking free energies and in different optimal stacking geometries, represented by altered twist, roll and tilt values for each dinucleotide pair (2, 5). These are defined as the three angles between the helix axis and the base pair plane, as demonstrated in Figure 4.

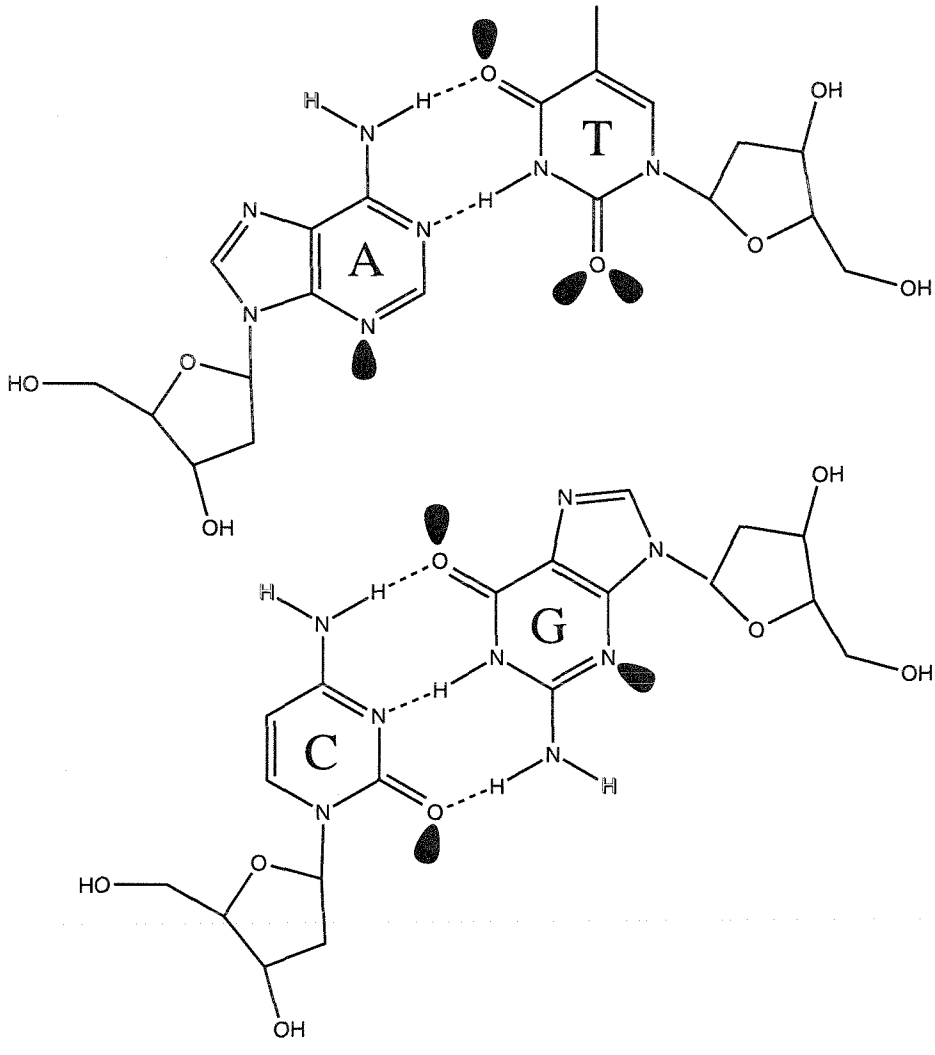


Figure 2. The hydrogen bonding pattern of AT and CG base pairs are indicated here. Lone pair orbitals and hydrogens available for bonding in the major and minor groove are also depicted.

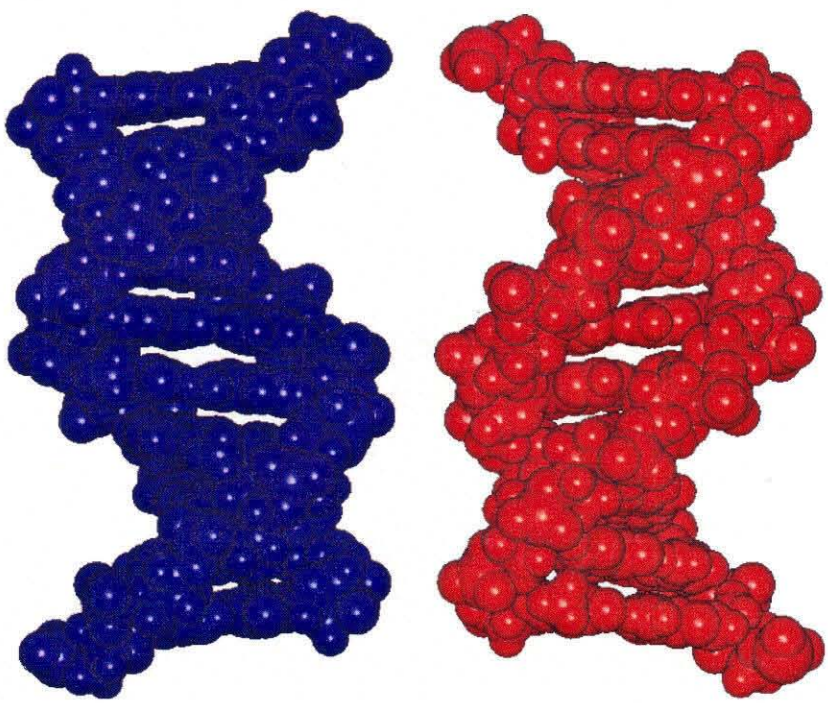


Figure 3. The blue helix on the left shows the major groove of B form DNA, while that on the right in red depicts the minor groove of B form DNA.

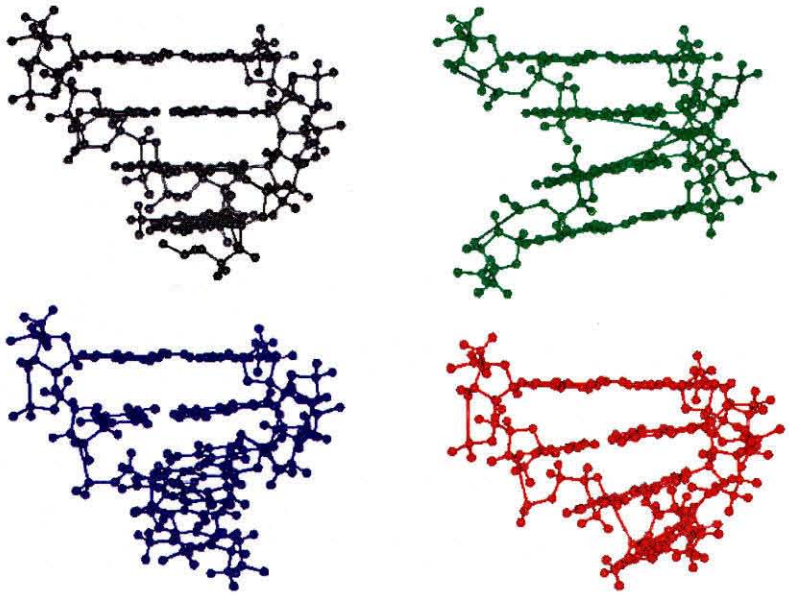


Figure 4. The upper left helix in black shows a regular B form helix. The helical angles are altered in the other three figures. The green helix has an increased helical twist angle, where the strand is overwound compared to the black helix. The blue helix has an altered roll value, where the bases are bent towards the viewer. The red helix has an altered tilt value, where the helix is bent to the side.

Short pieces of double stranded DNA are generally viewed as rigid rods, while longer DNA fragments behave as wormlike coils. The negative charge of the phosphate backbone serves to rigidify DNA through coulombic repulsion in the short range. Removal of this negative charge allows DNA to bend readily (6, 7). In fact, the large density of negative charge around DNA allows it to be viewed as a polyanion, where positively charged ions in solution such as K^+, Na^+, Mg^{++}, and organic cations including spermine are attracted to DNA to form a counterion shell. The concentration of cations in this shell around DNA has been estimated to be molar (8).

Debye-Huckel theory predicts the interionic potential energy between two hard spheres in a volume completely filled with a material containing a given dielectric constant. This approximation is only truly valid for dilute solutions, where the interionic distance is much greater than the hard sphere radius. However, extending Debye-Huckel theory to concentrated electrolyte solutions, a rough estimate of the Debye screening length can be obtained for counterion shells of this magnitude (9). Such estimates suggest a value on the order of the diameter of water. This large screening value negates long range charge effects and reduces the repulsive force between neighboring phosphates significantly, increasing the propensity for flexibility.

Other factors affect flexibility and curvature in a sequence dependent manner. Base stacking is different for each dinucleotide pair, as presented above (5). Additionally, solvation of the minor groove is different for different base pairs with different minor groove functionalities and sizes (10). Tracts of five to six A residues in a row have been demonstrated to be bent towards the minor groove. This bending is somewhat dependent upon the context of neighboring bases as well as upon the solvation state of the minor groove (10-12). Other sequence contexts show varying degrees of flexibility and directed bending (5).

DNA BENDING

Intrinsic DNA curvature appears to be important to many biological processes. Beyond inherent sequence effects, proteins that regulate such processes can work by bending or straightening DNA. Biological activities sensitive to protein-induced DNA bending include transcription, replication, site-specific recombination, viral integration, transposition, and DNA packaging (13-26). Bending is important through several biological mechanisms. First, bending can either directly or indirectly facilitate binding of an auxillary protein. This works both with proteins with altered affinities for prebent sites, or for dimeric proteins that bind at distal DNA sites, where the DNA is looped between them (27, 28). In both cases, the activation energy for protein binding is reduced. Additionally, bending can distort the bond angles in the phosphate linkage (29). This can result in a reduced energetic barrier to DNA unwinding and DNA cleavage through phosphate hydrolysis (30). It is through this mechanism that viral integrases and transposases are likely facilitated by bent DNA (19). Furthermore, some restriction endonucleases also bend DNA (31).

At a mechanistic level, DNA bending proteins bend DNA in three physically distinct ways. They can alter the stacking of the bases by intercalation of hydrophobic groups, alter the effective Debye length of the DNA surface through charge screening, and bend DNA as dimers through energetic compensation with tight binding events for conformational restriction (25, 27, 32-37). Figure 5 depicts a schematic of these three mechanisms. An example of a protein that appears to work through all three mechanisms in bending DNA by greater than 160° is integration host factor (IHF) (37). Figure 6 shows a depiction of IHF binding to and bending DNA into a U shaped structure. Given the importance of DNA bending in biological processes, these mechanisms can be envisioned as a starting point for the design of artificial regulators of DNA topology.

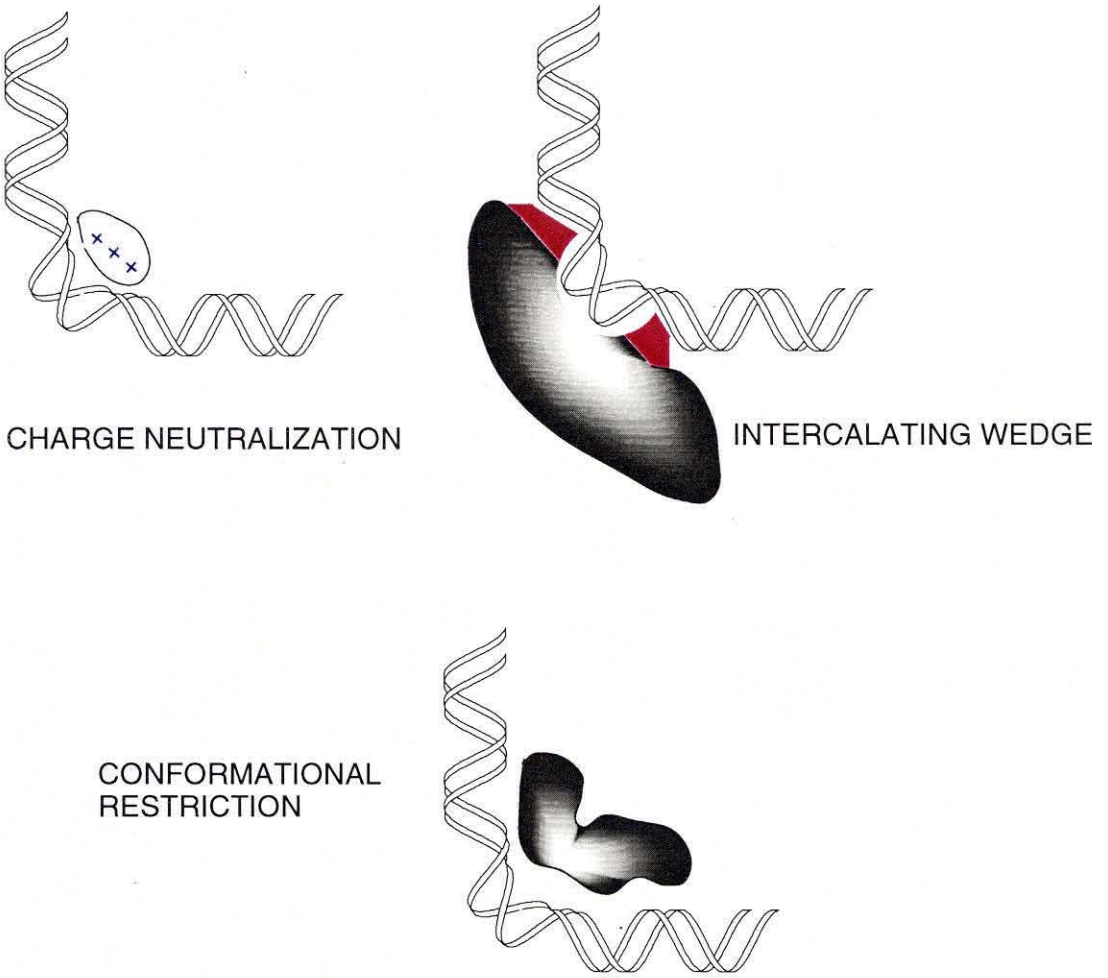


Figure 5. Three motifs of ligand induced bending are indicated. These involve charge neutralization, placement of an intercalating group into the helix, and conformational restriction.

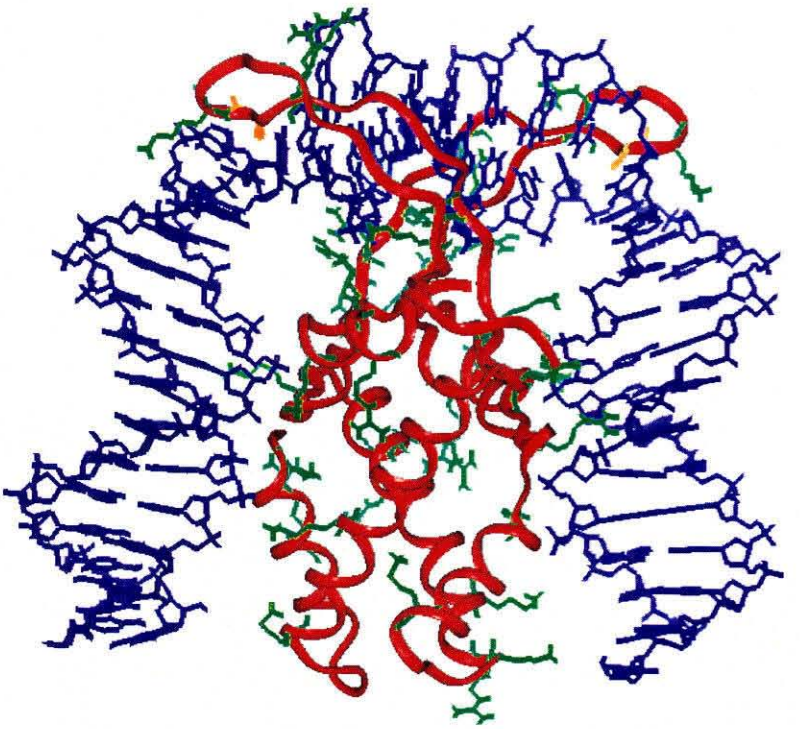


Figure 6. The bent structure of dimeric IHF-bound DNA is depicted here (37). The net bend angle is 160°. Here, orange proline residues are intercalated at the major kink sites. The rest of the bent structure is stabilized by specific van der Waals contacts and hydrogen bonds, as well as an electropositive surface. Positively charged residues are displayed in green and many of these are seen to complex specific phosphate groups.

ENGINEERED DNA BINDING MOLECULES

In designing a molecule to bind to DNA sequence specifically within a genome, towards the end of therapeutic gene regulation, a large binding site is required. For a ligand to be gene-specific, it should recognize only a single site within the entire genome. Assuming a random distribution of DNA sequences, a ligand needs to specifically recognize 17 base pairs of DNA to be specific in the 3 x 10^9 diploid base pairs of the human genome (38). In Figure 7, the number of unique sites for a given ligand size using this random model and some sample genome sizes are juxtaposed. Of course, promoters of genes contain combinatorial recognition elements for the same small number of transcription factors and can not really be considered to be random. Thus, the actual binding site size required to uniquely recognize a site within a genome is probably larger than estimated from the random model. However, the model is a good rough estimate of what a DNA binding ligand must be capable of binding.

Molecules can be designed to bind to DNA in either the major groove or the minor groove, or to intercalate between the bases. Modes of molecular recognition involve specific hydrogen bonds with functional groups in the grooves, van der Waals contacts with the bases, electrostatic interactions both with the bases and with the phosphate backbone, and steric clashes to preclude nonspecific binding. Many different DNA binding molecules are known. Only a fraction of those molecules are presented below.

INTERCALATORS

Intercalators are designed to be sequence specific through shape selective recognition in the absence of hydrogen bonding. The hydrophobic planar aromatic

SITE SIZE	# UNIQUE SITES
1	2
2	10
3	32
4	136
5	512
6	2,080
7	8,192
8	32,896
9	131,072
10	524,800
11	2,097,152
12	8,390,656
13	33,554,432
14	134,225,920
15	536,870,912
16	2,147,516,416
17	8,589,934,592
18	34,359,869,440
19	137,438,953,472
20	549,756,338,176

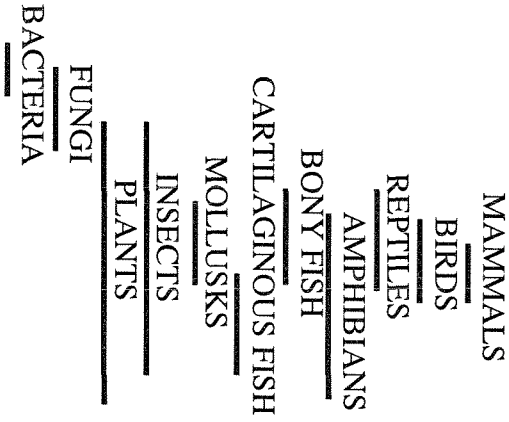


Figure 7. The number of unique sites is presented assuming a random genomic distribution of sequences (38). Given the size of organismal genomes, the number of base pairs that must be recognized to be site-specific using this model is indicated.

rings stack well between the DNA bases. An example of a compound that selectively recognizes an eight base pair target sequence is $Rh(DPB)_2phi^{+3}$ depicted in Figure 8 (39, 40). Recognition of its palindromic target site is through both specific van der Waals and steric interactions. To date, no specific intercalating compound has been designed with a larger binding site in the absence of added hydrogen bonding in one of the grooves.

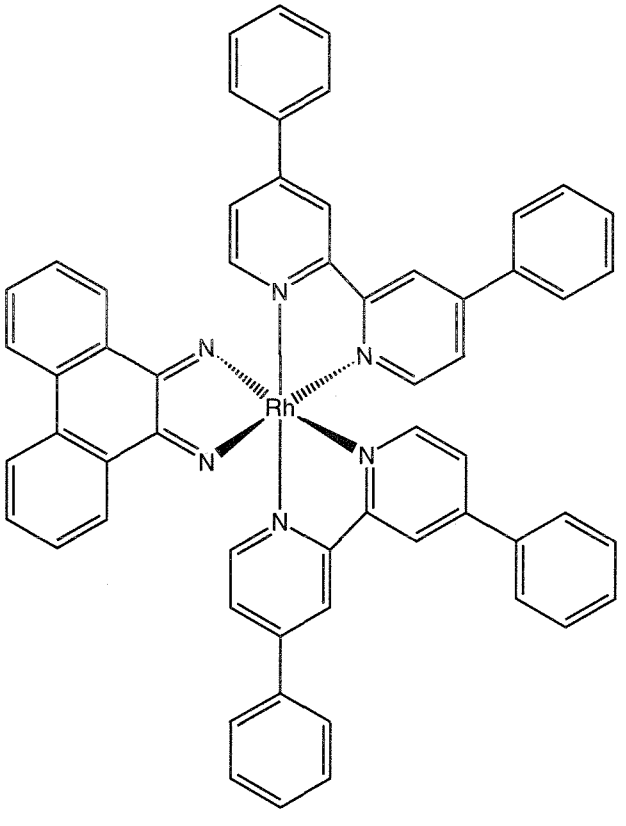
MAJOR GROOVE BINDERS

Groove binding is the preferred mode for specifically recognizing a large binding site with specific hydrogen bonds. Proteins naturally recognize specific DNA sites by groove binding.

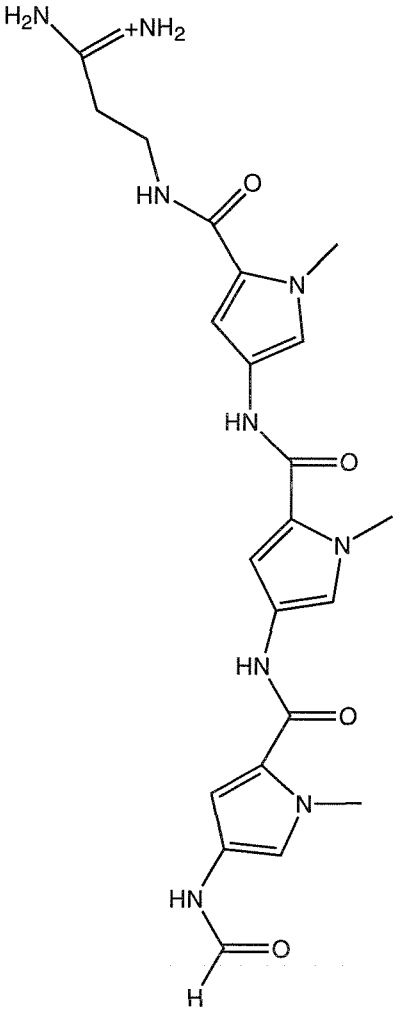
Previously, DNA binding domains have been redesigned to alter sequence specificity. Two examples have now been demonstrated of domains that can be stitched together to specifically recognize large sites.

The DNA binding domain from the v-jun leucine zipper protein has been linked in a modular sense by disulfide residues to form large recognition units (41, 42). These synthetic peptide multimers have been shown to recognize a 16 base pair DNA target site specifically. The nature of the linkage may allow for the joining of different recognition motifs in a single molecule.

Zinc fingers are modular DNA recognition units that bind in the major groove specifically (43, 44). These proteins can also be selected for binding to a specific site, and have the advantage of the possibility of biological expression (45-50). Affinities and specificities are both high, with affinities in the subnanomolar range and specificities against single base pair mismatches as high as 100 fold. The major disadvantage of these compounds, like leucine zippers, is the lack of direct readout of



$Rh(DPB)_2phi^{+3}$



Distamycin A

Figure 8. $Rh(DPB)_2phi^{+3}$ is shown as an example of a flat aromatic intercalator that shows site selectivity. Distamycin A is drawn as an example of a crescent shaped minor groove binding ligand.

the protein amino acid sequence corresponding to a desired DNA binding site. Thus, peptides with altered specificities need to be selected for rather than designed.

Recognition in both leucine zipper and zinc finger proteins derives both from specific hydrogen bonds and from van der Waals contacts with the target DNA bases and backbone. Additionally, positively charged residues contact the phosphate backbone. Metal ions in zinc fingers are crucial to the retention of secondary and tertiary structure of the motif, although Zn (II) can be replaced with other divalent metals (51).

Another class of DNA recognition molecules that bind in the major groove are third strand oligonucleotides. The early observation by Rich and coworkers in 1957 was that nucleic acids composed of three strands underwent 2:1 binding of poly (rU) to poly (rA) in the presence of Mg^{++} (52). Subsequent research elucidated the rules of recognition for two classes of oligonucleotides binding as third strands. In 1987, Moser and Dervan demonstrated that an oligonucleotide could direct intermolecular triple helix formation in the pyrimidine motif (53). Direct readout of double helical base pairs is accomplished by major groove binding of the oligonucleotide through specific Hoogsteen base pairing. Here, the third strand oligonucleotide binds parallel to the purine rich strand in the double helix, where the specific triplet interactions are T•AT and protonated C•GC base triplets. The need to protonate C in the third strand leads to a pH dependence for binding and C is commonly replaced with 5-methyl C, which is more readily protonated at N3. These triplets are depicted in Figure 9. A third weaker match is G recognizing TA, although this triplet is noncanonical and shows an altered glycosidic bond angle (χ) in the NMR structure (54, 55). This NMR structure of a pyrimidine motif triple helix indicates little distortion form canonical B form DNA and is depicted in Figure 10. Characterization of third strand oligonucleotide binding has shown triple helix formation to be sensitive to the length of the third strand, single base

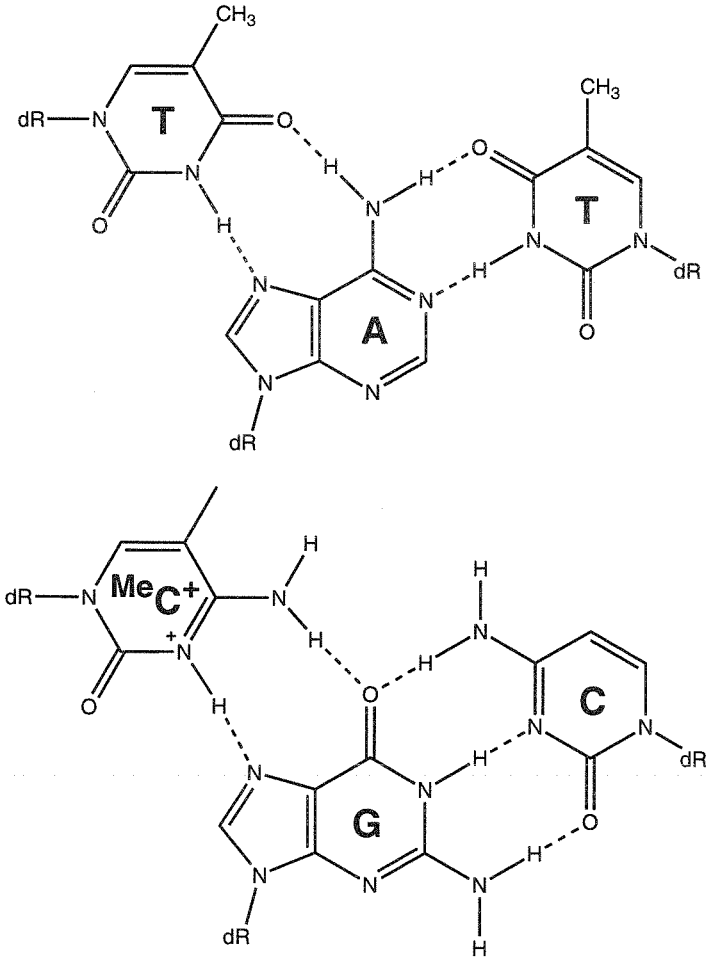


Figure 9. The canonical pyrimidine motif triplets are presented, where the third strand binds parallel to the purine rich strand.

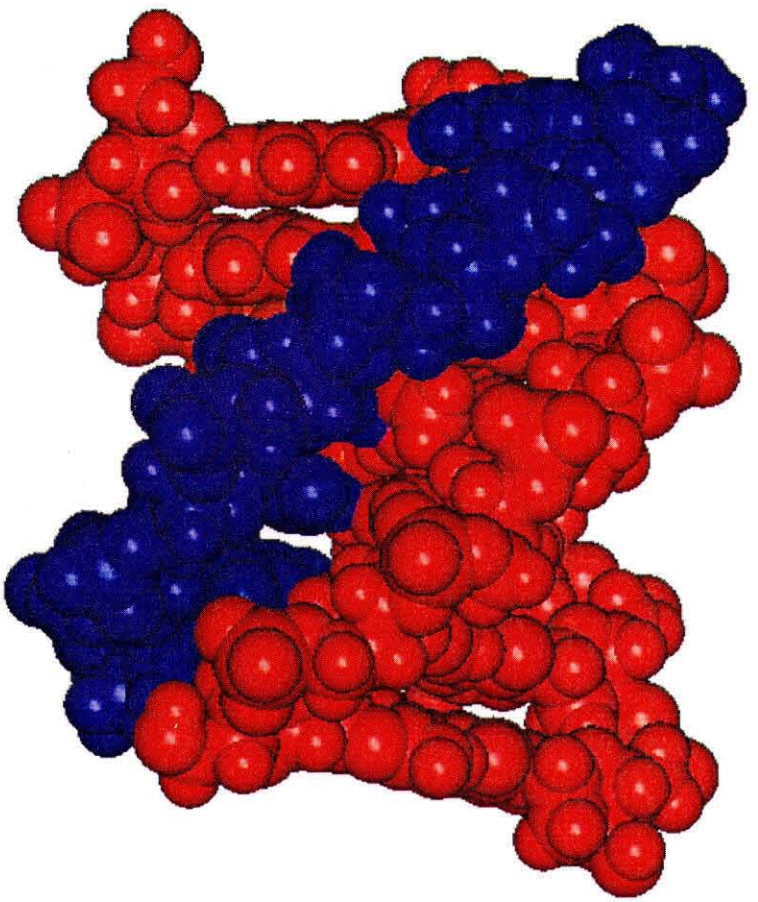


Figure 10. An NMR structure of a pyrimidine motif triple helix is shown (55). This particular complex contains a G•TA base triplet. In this structure, the blue third strand is shown bound in the B-like major groove of the red double helix.

mismatches, pH, cation concentration and valence, temperature, and backbone composition of the three strands (53, 56-67).

A second class of third strand oligonucleotide was characterized through the work of Beal and Dervan and of Greenberg and Dervan (68-70). This motif involves the antiparallel binding of a purine rich oligonucleotide through specific reverse Hoogsteen base pairs in a pH independent association. Here, specific G•GC and A•AT triplets are formed, where T•AT has been shown to be isoenergetic with A•AT. These triplets are depicted in Figure 11.

Third strand oligonucleotides bind with high affinity and specificity and can readily accommodate large target sites. In-vitro, these molecules have been demonstrated to displace DNA binding proteins, including transcription factors, indicating their potential utility as in-vivo DNA recognition agents (71, 72). However, while a third strand oligonucleotide can readily form under physiological conditions, evidence suggests that third strand oligonucleotides may be substrates for DNA repair enzymes (73). This remains a problem, but routes around this may be possible. It has been demonstrated that cis-platin adducts are protected from DNA repair by simultaneous binding of an HMG protein in the minor groove of DNA (74). For third strand oligonucleotides, simultaneous binding of minor groove binders has been demonstrated with little difference in affinity (75). The binding of these polyamides, unlike HMG proteins which bend DNA, is thought to rigidify DNA providing one difference between the studies. On the other hand, the sequence targeted by Moser and Dervan contains an A_5 tract which is bent towards the minor groove and bound with high affinity by a third strand oligonucleotide (53).

A second limitation of third strand oligonucleotides exists. Using the natural bases, there is no solution for binding of pyrimidine sequences in the Hoogsteen bonded strand. The one example of a designed base that successfully recognizes pyrimidine bases at a single site is D_3 (76-79). The chemical structure of D_3 is shown

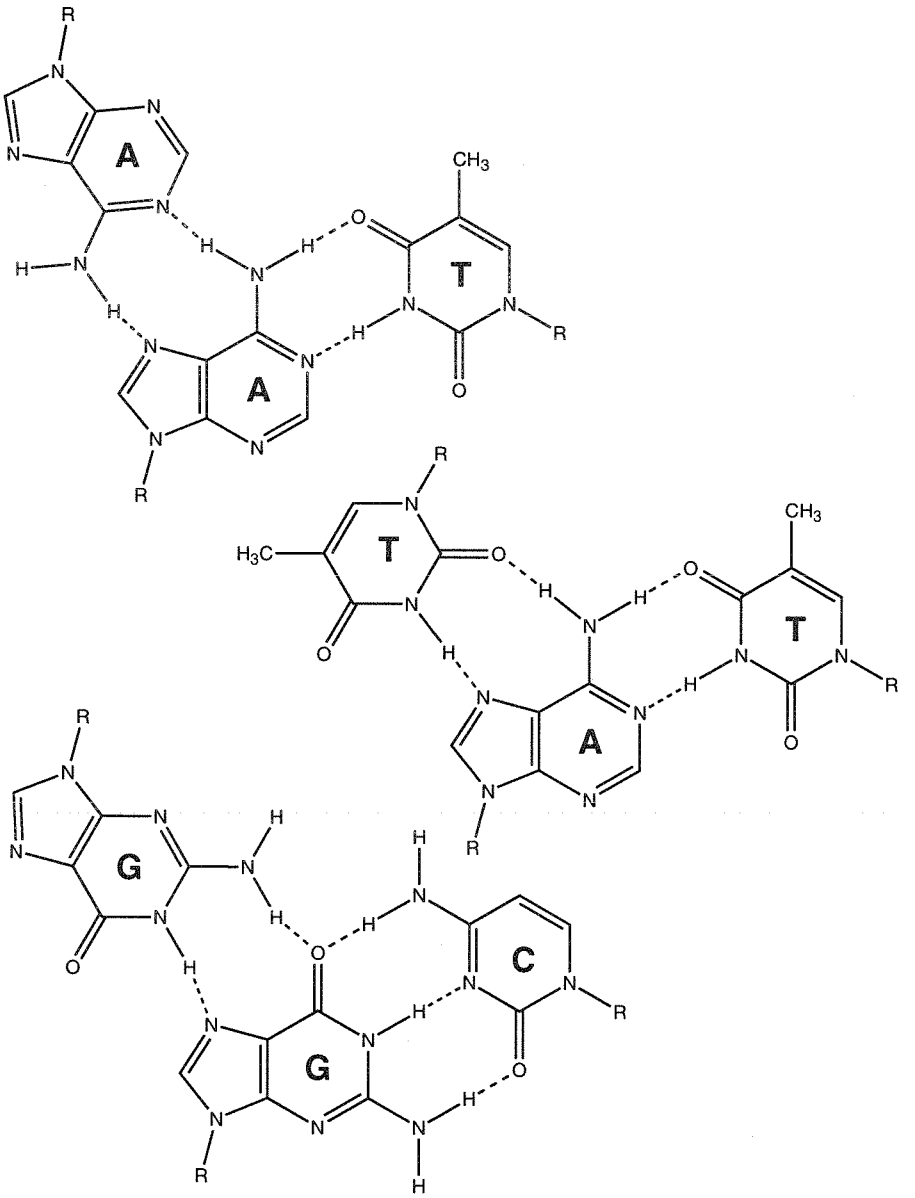


Figure 11. The canonical purine motif triplets are presented, where the third strand binds antiparallel to the purine rich strand.

in Figure 12. Further characterization of D_3 in triple helix formation is presented in chapter 2.

MINOR GROOVE BINDERS

Many natural products exist that recognize DNA sequence specifically utilizing an array of binding modes. An example of a molecule that binds in the minor groove of DNA that has been used for sequence specific recognition in the inhibition of transcription complex formation is calicheamycin (80). However, the binding site size is only four bases and it is not obvious how a generalizable class of DNA binding ligands can be generated from it.

Generalizable DNA recognition can also be accomplished through the minor groove. Natural products such as DAPI, netropsin, and distamycin A bind through the formation of specific hydrogen bonds in the minor groove of DNA, extensive van der Waals contacts, and an electrostatic interaction with the phosphates of DNA. The structure of distamycin A is depicted in Figure 8, as an example of a crescent-shaped minor groove binding compound.

These molecules have been the basis for the design of sequence specific polyamides that bind in the minor groove of DNA. Distamycin and netropsin bind primarily as monomers to AT rich DNA. Subsequent structural studies led to the development of a model for binding and allowed rational design of polyamides with specific binding capabilities for GC recognition by replacing an N-methyl-pyrrole (Py) β amino acid with an N-methyl-imidzaole (Im) (81).

Shortly thereafter, NMR structural studies showed that high concentrations of distamycin bound as an antiparallel side by side dimer in the minor groove of DNA (82, 83). Figure 13 depicts this NMR structure. Subsequent NMR structural studies of polyamides containing Im and Py in conjunction with affinity cleavage data were

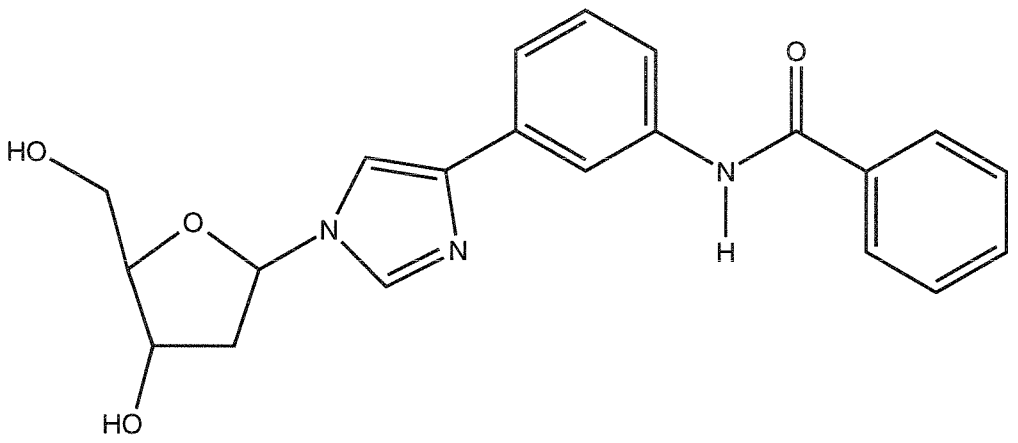


Figure 12. The chemical structure of novel nucleoside D_3 is shown. This nucleoside can recognize target pyrimidines in a purine rich stretch at a single position when incorporated into a third strand oligonucleotide.

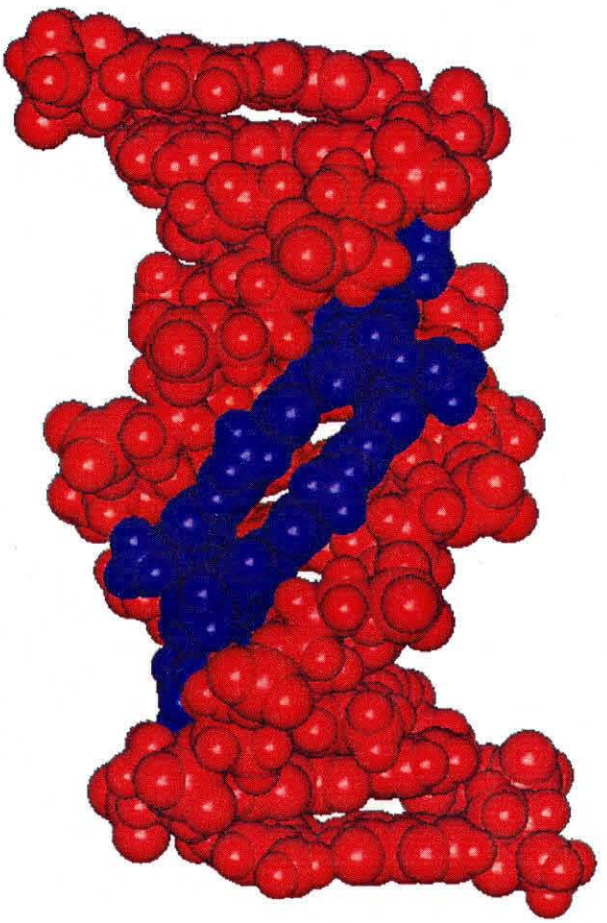


Figure 13. An NMR structure of blue distamycin bound as a side-by-side dimer in the minor groove of a red double helix is depicted (82, 83).

consistent with GC recognition by the designed compound occurred through dimeric binding (84-87). One model for this binding is that the compressed minor groove of AT rich DNA is 3-4 Å wide compared with that of mixed sequence B DNA, which is 5-6 Å wide (88). This wider minor groove contains GC base pairs, which are more sterically demanding with functional groups forming a third hydrogen bond proximal to the minor groove. Together, these allow for close van der Waals contacts of a dimeric complex for polyamides.

General rules for predictably modifying the sequence specificity of designed polyamides have been well established, with a Py/Py pair degenerate for AT and TA, Im/Py recognizing GC, and Py/Im recognizing CG. Im/Im pairs are disfavored (89). Subsequent work has led to the development of covalently linked hairpin molecules that bind with high affinity and specificity (90, 91). The binding site size limitations of dimeric and covalent hairpin molecules have now been determined. Dimeric polyamides bind with high affinity and specificity to 7 base pair sites, but can be extended to 13 base pair sites by substitution of a flexible β-alanine linker (92, 93). Hairpin polyamides show a similar binding range (94). Ongoing research in the Dervan group is extending the length of this binding range by incorporating β-alanine opposite Py instead of Py/Py at a subset of AT positions. Despite current disadvantages with binding site length and AT/TA degeneracy, polyamides are an extremely powerful tool for recognizing DNA in the minor groove.

Engineered DNA binding proteins, third strand oligonucleotides, and minor groove binding polyamides represent the three most promising technologies for generalizable solutions to single site recognition in a genome. The research described in this thesis utilizes two of these classes of molecules, third strand oligonucleotides and polyamides.

TECHNIQUES FOR THE ANALYSIS OF DNA BINDING MOLECULES

In the course of analysis of DNA binding molecules, many techniques are available. In this thesis, two techniques were predominantly utilized, DNase I footprinting and gel shift analysis.

DNase I footprinting relies upon the enzyme DNase I for DNA cleavage. The enzyme associates with the minor groove of DNA and shows a slight sequence specificity in cleavage. The presence of a ligand associated with DNA in either groove is sufficient to preclude cleavage of the bound site. Radiolabeling DNA at one end of a single strand allows for detection of binding of a ligand by protection from random cleavage in contrast to a control lane lacking ligand. This is depicted in a quantitative sense in Figure 14, where titration of a ligand can be utilized to determine the equilibrium association constant (95).

Cleavage at a specific and reference site are measured for each concentration in the titration and in the control lane lacking ligand. The apparent DNA target site saturation, θ_{app}, is determined using the equation

$$\theta_{app} = 1 - (I_{tot}/I_{ref})/(I_{tot}^{0}/I_{ref}^{0})$$

where I_{tot} and I_{ref} are the integrated volumes from the target and reference sites, while I_{tot}^{0} and I_{ref}^{0} are the corresponding values for the ligand-free lane. Using the saturation fitting parameters θ_{min} and θ_{max}, K_a, the apparent equilibrium association constant is determined from fitting a Langmuir binding isotherm to the data points $[L]_{bulk}$, representing the bulk ligand concentration in solution, and θ_{app}, using the modified Hill equation

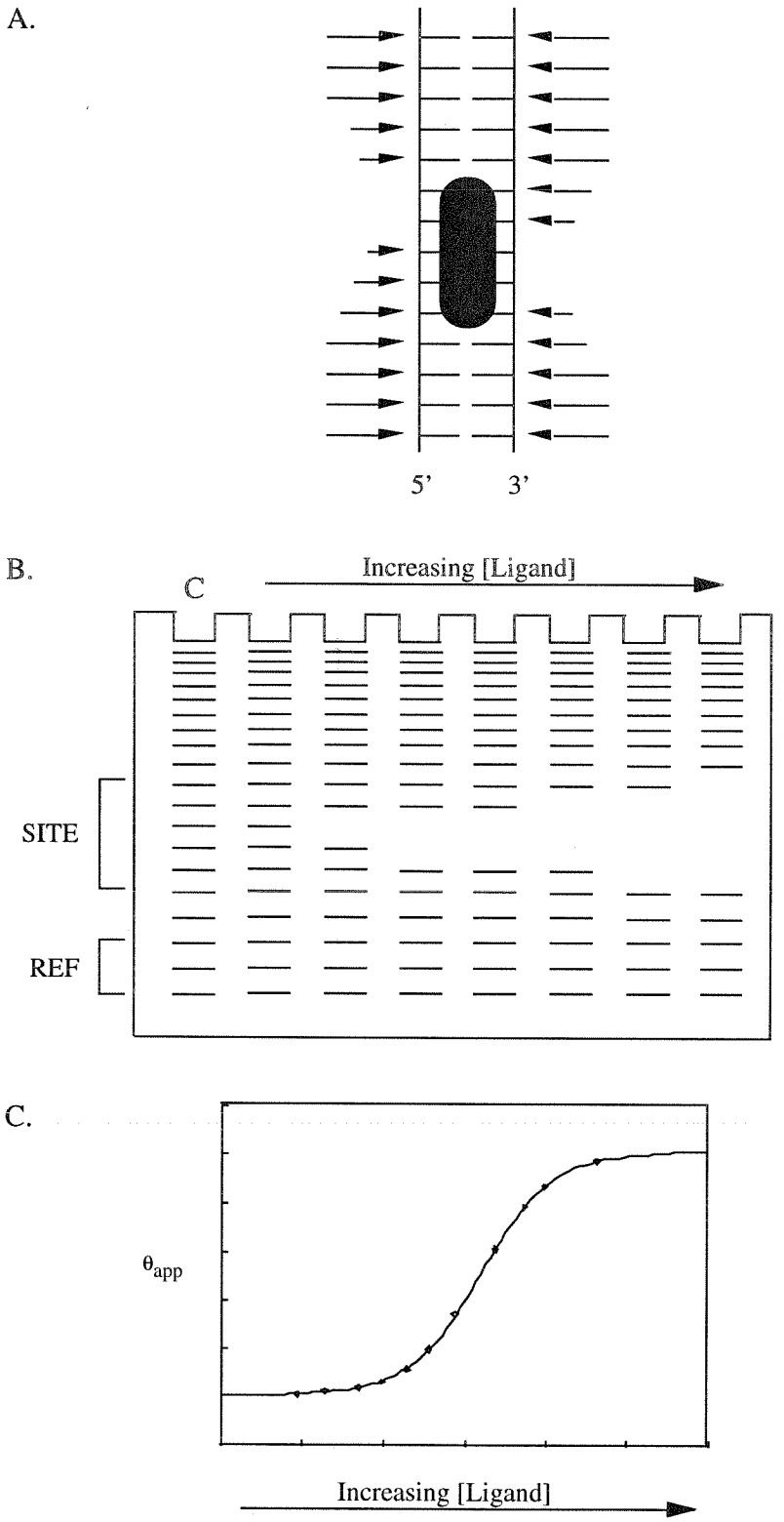


Figure 14. A DNase I footprinting reaction where occupation of a DNA site by a ligand precludes cleavage by the enzyme proportional to arrow length is shown on top. Footprinting repeated over a series of concentrations yields a gel shown in the middle. Analysis of this data as described results in an isotherm shown on the bottom.

$$\theta_{app} = \theta_{min} + (\theta_{max} - \theta_{min})\ (K_a[L]_{bulk})/(1+K_a[L]_{bulk})$$

It should be noted that this equation assumes monomeric association of ligands in the absence of cooperative effects. It also assumes that $[L]_{bulk} = [L]_{solution}$, which is only true when the labeled DNA concentration is much lower than the ligand concentration. This assumption is valid for the data presented in this thesis.

Another compound that has been used for DNA footprinting for qualitative experiments is MPE. This sequence neutral small molecule intercalator contains an EDTA moiety, which binds Fe^{++}. Fe^{++}•EDTA in the presence of a reducing agent such as dithiothreitol generates a defusable radical that cleaves DNA proximal to its generation, where the MPE is intercalated (96). Binding of a ligand precludes binding of MPE at that site and consequent hydroxyl radical generation.

A second class of experiments performed in this thesis are those involving gel shift analysis. As a DNA molecule, or DNA with a ligand bound migrates through a gel, mobility is driven by the charge. DNA separation on a gel is dependent upon additional factors such as the persistence length, the contour length, and the frictional coefficient, as well as the charge (97). Therefore, the shape of the molecule and the pore size of the gel effect mobility. This motion can be visualized microscopically as a DNA molecule reptating through pores as it traverses a gel (98, 99).

When a ligand binds to DNA, the size of the structure increases, and its frictional drag through a gel retards its mobility. A titration of a binding ligand will show a shift from a less retarded ligand free band to a more retarded band with ligand bound as the concentration is increased. This is depicted in Figure 15. The percentage shifted is themodynamically equivalent to the fractional saturation of a site in DNase I footprinting, and data can be fit with the same modified Hill equation (100).

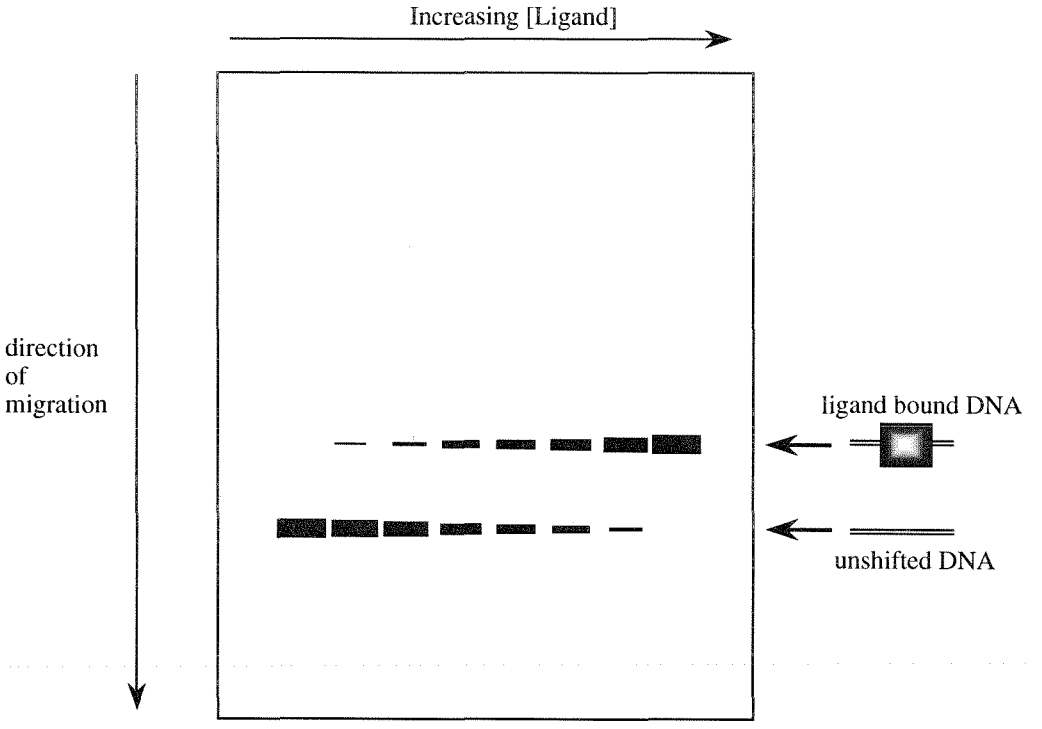
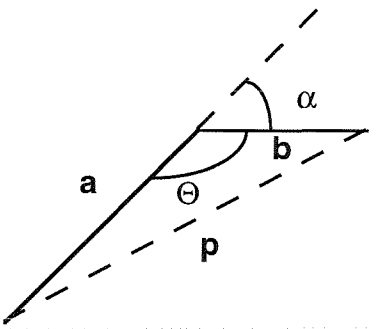


Figure 15. This depiction shows a gel shift titration with increasing concentrations of ligand, where bands represent radiolabeled DNA that has migrated through the gel dependent upon complex size.

Similarly, when DNA is bent, its mobility is reduced. The hypothesized reason for this is that DNA constrained in a bent structure can not realign with the electric field as it reptates through a gel. From this model, two electrophoretic assays have been developed that take advantage of the reduction of mobility in bent DNA.

The first assay is a circular permutation assay (13, 101). This assay is depicted in Figure 16. Here, a bent site is moved from one terminus of a restriction fragment through the center, to the other end. In doing so, the mean square end-to-end distance of the fragment is reduced for the fragment with the centered bend. Theory holds that the mean square end-to-end distance is directly related to gel mobility for fragments of the same size. Using the law of cosines as shown below, a relationship can be derived relating relative mobility to bend angle.



$$p^2=a^2+b^2-2ab\cos\Theta$$

MAXIMUM RETARDATION

$$a=b$$

$$p^2=2a^2(1-\cos\Theta)$$

since $\sin^2\Theta/2=1/2(1-\cos\Theta)$ and $\sin(\pi/2-\alpha/2)=\cos(\alpha/2)$,

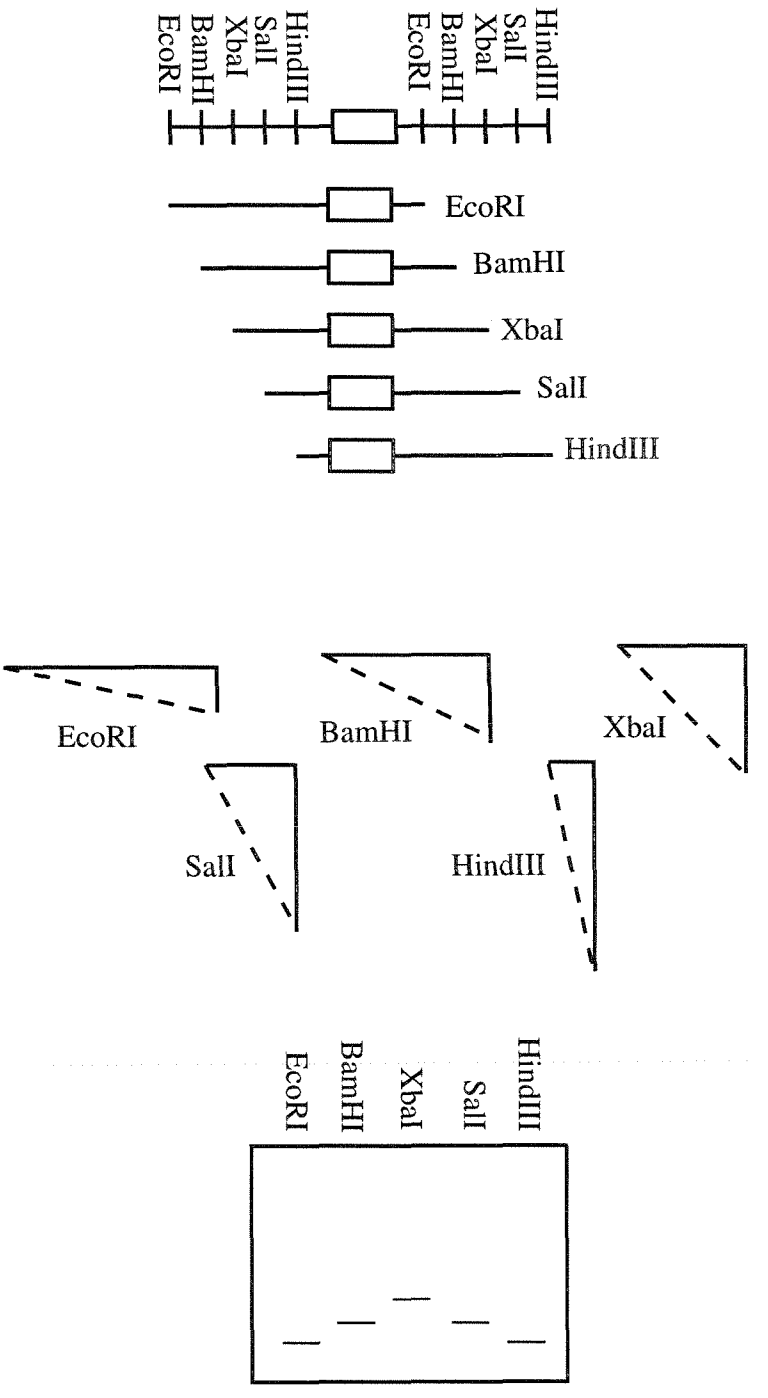


Figure 16. A circular permutation assay uses a plasmid with tandem polylinkers surrounding a ligand binding site. As the bend locus is moved from one edge of the ligand binding site through the center to the other edge, the end-to-end distance changes, resulting in an alteration in mobility as depicted. Note that bends near the center of the fragment are shifted to the largest extent.

$$p=2a\cos(\alpha/2)$$

MINIMUM RETARDATION

$$p=2a$$

BEND ANGLE

$$X_{max}/X_{min}=\cos(\alpha/2)$$

This assay has been shown to be sensitive to DNA flexibility and to the binding of ligands that are large compared to the target DNA fragment, and the value obtained from such experiments is now labeled as the flexure angle (102, 103). A second assay has been developed that provides information on bend center and direction as well as the bend angle (102, 104-107).

A phasing analysis is based upon the measurement of the bend angle of three phased A_6 tracts at 54° towards the minor groove (108). Altering the distance between the center of the A tract and the center of the designed bend locus will alter the phasing of the bends, where the periodicity of mobility should be 10.4 bp, the helical phasing of the DNA helix. This relationship is shown diagramatically in Figure 17. Mobilities are plotted against linker length and fit to the equation

$$\mu=\mu_{ave}\{A_{ph}/2[\cos(2\pi(S-S_T/P_{ph}))]+1\}$$

where μ is the electrophoretic mobility measured in apparent base pairs (R_L), A_{ph} is the phasing amplitude, S is the distance between bend centers, S_T is the out of phase bend center distance, and P_{ph} is the phasing period. The bend angle, α_B, was then determined from the geometrically derived phasing function (102, 106).

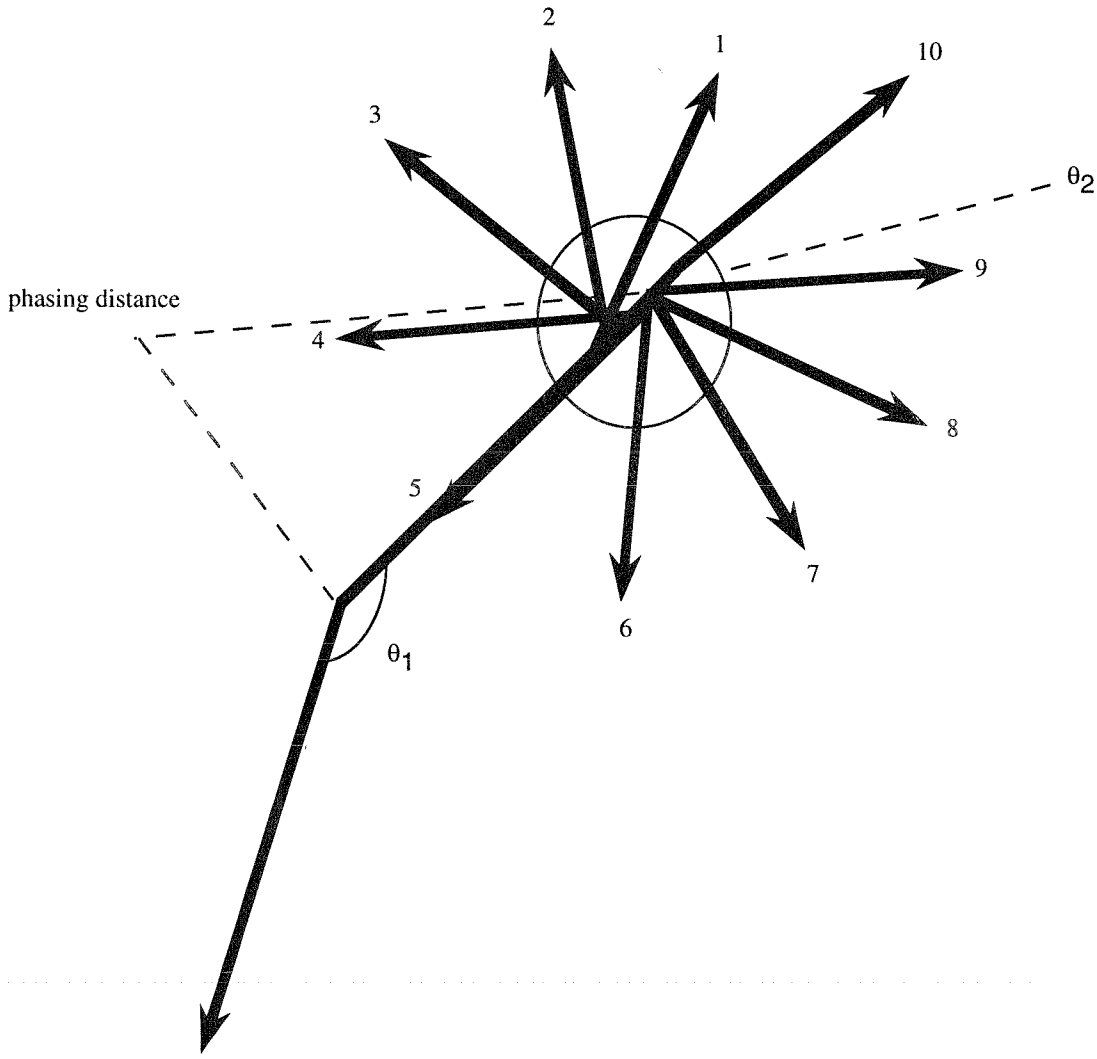


Figure 17. A diagram of a phasing analysis, where an A tract bent by angle θ_1 is separated by the phasing distance from a ligand induced bend of angle θ_2 is depicted where the DNA helix is viewed from the side. The phasing of the two angles determines the additivity of the net bend angle.

$$\tan(k\alpha_B/2)=(A_{ph}/2)/\tan(k\alpha_c/2)$$

where α_c is the A tract bend angle.

One concern with this technique is the involvement of two separated bend angles, where the theory on mobility is based upon a single bend angle. It is important to keep this distance as small as possible in relation to the overall size of the fragment, so that the migrating DNA approximates an "l" or a "v", not an "s" or a "u." Modeling of the relative end-to-end distance in relation to phasing distance shows this effect on net measured bend angle for a centered fragment in Figure 18. This s effect has also been reported experimentally (107). This technique was utilized for measurement of reported bend angles using a long fragment and short interbend distance.

THESIS WORK

The focus of this thesis is the design of ligands to sequence-specifically bind to DNA to alter its topology. Chapter 2 describes the binding properties of the novel intercalating base D_3 in the context of triple helix formation. The ability of this intercalating base to bend DNA was assayed for. Chapter 3 describes the design of triple helical DNA bending ligands using an alternative strategy, conformational restriction. Chapter 4 undertakes a further biophysical analysis of this bending system. Chapter 5 presents the design of a complementary DNA straightening polyamide that displaces bound third strand oligonucleotide from its bent conformation.

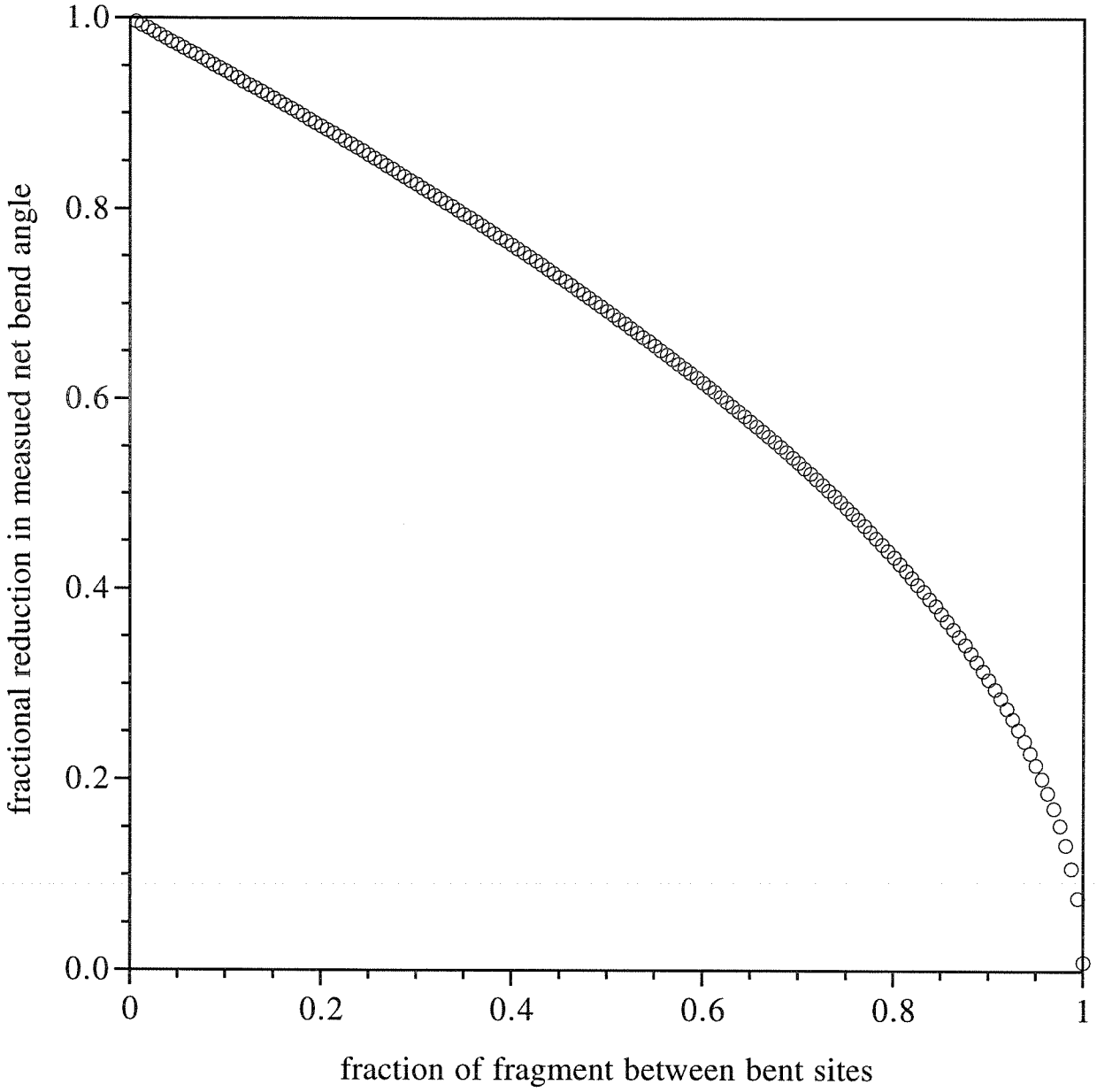


Figure 18. Based on the law of cosines for measuring the net end-to-end distance of a fragment, the effect of converting a triangle to a trapezoid is shown, where the fourth side represents the interbend angle distance. This analysis shows the importance of designing a fragment with an interbend distance that is much shorter than the fragment length.

REFERENCES

1. Watson, J. D., Hopkins, N. H., Roberts, J. W., Steitz, J. A. & Weiner, A. M. (1987) *Molecular Biology of the Gene* (The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA).
2. Saenger, W. (1984) *Principles of Nucleic Acid Structure* (Springer-Verlag, New York).
3. Westheimer, F. H. (1987) *Science* **235,** 1173-1178.
4. Switzer, C. Y., Moroney, S. E. & Benner, S. A. (1993) *Biochemistry* **32,** 10489-10496.
5. Bolshoy, A., McNamara, P., Harrington, R. E. & Trifonov, E. N. (1991) *Proc. Natl. Acad. Sci., USA* **88,** 2312-2316.
6. Strauss, J. K. & Maher, L. J. (1994) *Science* **266,** 1829-1834.
7. Manning, G. S. (1983) *Biopolymers* **22,** 689-729.
8. Bond, J. P., Anderson, C. F. & Record, M. T. (1994) *Biophysical Journal* **67,** 825-836.
9. Philpott, M. R. & Glosli, J. N. (1995) *J. Electrochem. Soc.* **142,** L25-L28.
10. Dlakic, M., Park, K., Griffith, J. D., Harvey, S. C. & Harrington, R. E. (1996) *J. Biological Chemistry* **271,** 17911-17919.
11. Dlakic, M. & Harrington, R. E. (1996) *Proc. Natl. Acad. Sci., USA* **93,** 3847-3852.
12. Haran, T. E., Kahn, J. D. & Crothers, D. M. (1994) *J. Mol. Biol.* **244,** 135-143.
13. Thompson, J. F. & Landy, A. (1988) *Nucleic Acids Research* **16,** 9687-9705.
14. Goodman, S. D., Nicholson, S. C. & Nash, H. A. (1992) *Proc. Natl. Acad. Sci., USA* **89,** 11910-11914.

15. Hallet, B., Rezsohazy, R., Mahillon, J. & Delcour, J. (1994) *Mol. Microbiol.* **14,** 131-139.

16. Goryshin, I. Y., Kil, Y. V. & Reznikoff, W. S. (1994) *Proc. Natl. Acad. Sci., USA* **91,** 10834-10838.

17. Henriquez, V., Milisavljevic, V., Kahn, J. D. & Gennaro, M. L. (1993) *Gene* **134,** 93-98.

18. Nakajima, M., Sheikh, Q. I., Yamaoka, K., Yui, Y., Kajiwara, S. & Shishido, K. (1993) *Mol. Gen. Genet.* **237,** 1-9.

19. Muller, H. P. & Varmus, H. E. (1994) *EMBO J.* **13,** 4704-4714.

20. Milot, E., Belmaaza, A., Rassart, E. & Chartrand, P. (1994) *Virology* **201,** 408-412.

21. Natesan, S. & Gilman, M. Z. (1993) *Genes Dev* **7,** 2497-2509.

22. Meacock, S., Pescini-Gobert, R., DeLamarter, J. F. & vanHuijsduijen, R. H. (1994) *J. Biological Chemistry* **269,** 31756-31762.

23. Kim, J., Klooster, S. & Shapiro, D. J. (1995) *J. Biological Chemistry* **270,** 1282-1288.

24. Perez-Martin, J., Rojo, F. & DeLorenzo, V. (1994) *Microbiological Reviews* **58,** 268-290.

25. Manning, G. S., Ebralidse, K. K., Mirzabekov, A. D. & Rich, A. (1989) *J. Biomol. Struct. Dyn.* **6,** 877-889.

26. Polaczek, P., Kwan, K., Liberles, D. A. & Campbell, J. L. (1997) , submitted.

27. Finzi, L. & Gelles, J. (1995) *Science* **267,** 378-380.

28. Parvin, J. D., McCormick, R. J., Sharp, P. A. & Fisher, D. E. (1995) *Nature* **373,** 724-727.

29. Liang, C., Ewig, C. S., Stouch, T. R. & Hagler, A. T. (1993) *J. Am. Chem. Soc.* **115,** 1537-1545.

30. Ramstein, J. & Lavery, R. (1988) *Proc. Natl. Acad. Sci., USA* **85,** 7231-7235.

31. Kim, S. C., Skowron, P. M. & Szybalski, W. (1996) *J. Molecular Biology* **258,** 638-649.

32. Kim, J. L., Nikolov, D. B. & Burley, S. K. (1993) *Nature* **365,** 520-527.

33. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. (1993) *Nature* **365,** 512-520.

34. Haqq, C. M., King, C., Ukiyama, E., Falsafi, S., Haqq, T. N., Donahoe, P. K. & Weiss, M. A. (1994) *Science* **266,** 1494-1500.

35. Mondragon, A. & Harrison, S. C. (1991) *J. Molecular Biology* **219,** 321-334.

36. Fisher, R. F. & Long, S. R. (1993) *J. Molecular Biology* **233,** 336-348.

37. Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. (1996) *Cell* **87,** 1295-1306.

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

39. Sitlani, A., Dupureur, C. M. & Barton, J. K. (1993) *J. Am. Chem. Soc.* **115,** 12589-12590.

40. Sitlani, A. & Barton, J. K. (1994) *Biochemistry* **33,** 12100-12108.

41. Park, C., Campbell, J. L. & Goddard, W. A. (1993) *Proc. Natl. Acad.Sci., USA* **90,** 4892-4896.

42. Park, C., Campbell, J. L. & Goddard, W. A. (1995) *J. Am. Chem. Soc.* **117,** 6287-6291.

43. Schmiedeskamp, M. & Klevit, R. E. (1994) *Current Opinion in Structural Biology* **4,** 28-35.

44. Suzuki, M. & Yagi, N. (1994) *Proc. Natl. Acad. Sci., USA* **91,** 12357-12361.

45. Jamieson, A. C., Kim, S. & Wells, J. A. (1994) *Biochemistry* **33,** 5689-5695.

46. Choo, Y. & A.Klug (1994) *Proc. Natl. Acad. Sci., USA* **91,** 11168-11172.

47. Choo, Y. & Klug, A. (1994) *Proc. Natl. Acad. Sci., USA* **91,** 11163-11167.

48. Rebar, E. J. & Pabo, C. O. (1994) *Science* **263,** 671-673.

49. Pomerantz, J. L., Sharp, P. A. & Pabo, C. O. (1995) *Science* **267,** 93-96.

50. Greisman, H. A. & Pabo, C. O. (1997) *Science* **275,** 657-661.

51. Kuwahara, J. & Coleman, J. E. (1990) *Biochemistry* **29,** 8627-8631.

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

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

54. Griffin, L. C. & Dervan, P. B. (1989) *Science* **245,** 967-971.

55. Radhakrishnan, I. & Patel, D. J. (1994) *Structure* **2,** 17-32.

56. LeDoan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J., Thuong, N. T., Lhomme, J. & Helene, C. (1987) *Nucleic Acids Research* **15,** 7749-7760.

57. Singleton, S. F. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 6957-6965.

58. Roberts, R. W. & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci., USA* **88,** 9397-9401.

59. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannageli, C., Garestier, T. & Helene, C. (1992) *Biochemistry* **31,** 9269-9278.

60. Fosella, J. A., Kim, Y. J., Shih, H., Richards, E. G. & Fresco, J. R. (1993) *Nucleic Acids Research* **21,** 4511-4515.

61. Best, G. C. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 1187-1193.

62. Hunziker, J., Priestley, E. S., Brunar, H. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 2661-2662.

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

64. Xodo, L. E., Manzini, G., Quadrifoglio, F., vanderMarel, G. A. & vanBoom, J. H. (1991) *Nucleic Acids Research* **19,** 5625-5631.

65. Singleton, S. F. & Dervan, P. B. (1992) *Biochemistry* **31,** 10995-11003.

66. Singleton, S. F. & Dervan, P. B. (1993) *Biochemistry* **32,** 13171-13179.

67. Singleton, S. F. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 10376-10382.

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

69. Beal, P. A. & Devan, P. B. (1992) *Nucleic Acids Research* **20,** 2773-2776.

70. Greenberg, W. A. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 5016-5022.

71. Maher, L. J., Dervan, P. B. & Wold, B. (1990) *Biochemistry* **31,** 70-81.

72. Maher, L. J., Wold, B. & Dervan, P. B. (1989) *Science* **245,** 725-730.

73. Wang, G., Seidman, M. M. & Glazer, P. M. (1996) *Science* **271,** 802-805.

74. Mcanulty, M. M. & Lippard, S. J. (1996) *Mutation Research DNA Repair* **362,** 75-86.

75. Parks, M. E. & Dervan, P. B. (1996) *Bioorganic & Medicinal Chemistry* **4,** 1045-50.

76. Griffin, L. C., Kiessling, L. L., Beal, P. A., Gillespie, P. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 7976-7982.

77. Kiessling, L. L., Griffin, L. C. & Dervan, P. B. (1992) *Biochemistry* **31,** 2829-2834.

78. Koshlap, K. M., Gillespie, P., Dervan, P. B. & Feigon, J. (1993) *J. Am. Chem. Soc.* **115,** 7908-7909.

79. Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B. & Feigon, J. (1996) *J. Molecular Biology* **257,** 1052-1069.

80. Ho, S. N., Boyer, S. H., Schreiber, S. L., Danishefsky, S. J. & Crabtree, G. R. (1994) *Proc. Natl. Acad. Sci., USA* **91,** 9203-9207.

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

82. Pelton, J. G. & Wemmer, D. E. (1989) *Proc. Natl. Acad. Sci., USA* **86,** 5723.

83. Pelton, J. G. & Wemmer, D. E. (1990) *J. Am. Chem. Soc.* **112,** 1393-1399.

84. Geierstanger, B. H., Dwyer, T. J., Bathini, Y., Lown, J. W. & Wemmer, D. E. (1993) *J. Am. Chem. Soc.* **115,** 4474-4482.

85. Geierstanger, B. H., Jacobsen, J. P., Mrksich, M., Dervan, P. B. & Wemmer, D. E. (1994) *Biochemistry* **33,** 3055-3062.

86. Geierstanger, B. H., Mrksich, M., Dervan, P. B. & Wemmer, D. E. (1994) *Science* **266,** 646-650.

87. Dwyer, T. J., Geierstanger, B. H., Mrksich, M., Dervan, P. B. & Wemmer, D. E. (1993) *J. Am. Chem. Soc.* **115,** 9900-9906.

88. Yoon, C., Prive, G. G., Goodsell, D. S. & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci., USA* **85,** 6332-6336.

89. White, S., Baird, E. E. & Dervan, P. B. *unpublished data* .

90. Mrksich, M., Parks, M. E. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 7983-7988.

91. Trauger, J. W., Baird, E. E. & Dervan, P. B. (1996) *Nature* **382,** 559-61.

92. Trauger, J. W., Baird, E. E., Mrksich, M. & Dervan, P. B. (1996) *J. Am. Chem. Soc.* **118,** 6160-6166.

93. Kelly, J. J., Baird, E. E. & Dervan, P. B. (1996) *Proc. Natl. Acad. Sci., USA* **93,** 6981-6985.

94. Turner, J. M., Baird, E. E. & Dervan, P. B. (1997) *submitted* .

95. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986) *Proc. Natl. Acad. Sci., USA* **83,** 8462-8466.

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

97. Zimm, B. H. & Levene, S. D. (1992) *Q. Rev. Biophys.* **25,** 171-204.

98. Smith, S. B., Aldridge, P. K. & Callis, J. B. (1989) *Science* **243,** 203-206.

99. Schwartz, D. C. & Koval, M. (1989) *Nature* **338,** 520-522.

100. Ferber, M. J. & Maher, L. J. (1997) *Analytical Biochemistry* **244,** 312-320.

101. Wu, H. & Crothers, D. M. (1984) *Nature* **308,** 509-513.

102. Kerppola, T. K. & Curran, T. (1991) *Science* **256,** 1210-1214.

103. Kuprash, D. V., Rice, N. R. & Nedospasov, S. A. (1995) *Nucleic Acids Research* **23,** 427-433.

104. Zinkel, S. S. & Crothers, D. M. (1987) *Nature* **328,** 178-181.

105. Zinkel, S. S. & Crothers, D. M. (1990) *Biopolymers* **29,** 29-38.

106. Kerppola, T. & Curran, T. (1993) *Mol. Cell. Biol.* **13,** 5479-5489.

107. Kerppola, T. K. (1996) *Proc. Natl. Acad. Sci., USA* **93,** 10117-10122.

108. Koo, H., Drak, J., Rice, J. A. & Crothers, D. M. (1990) *Biochemistry* **29,** 4227-4234.

CHAPTER 2

SEQUENCE SPECIFIC DNA RECOGNITION AND BENDING USING OLIGONUCLEOTIDES CONTAINING NOVEL BASE D_3

Oligonucleotide-directed triple helix formation is a general method for the sequence-specific recognition of double helical DNA. Two classes of triple helix forming oligonucleotides have been identified. In one class, pyrimidine oligonucleotides bind double helical DNA in the major groove parallel to the A-rich purine tract through the formation of specific Hoogsteen hydrogen bonds (1, 2). In the other class, purine oligonucleotides bind double helical DNA in the major groove antiparallel to the G-rich purine tract through the formation of specific reverse Hoogsteen hydrogen bonds (3-6). The stabilities of these oligonucleotides are dependent upon the length of the third strand, single base mismatches, pH, cation concentration and valence, temperature, and backbone composition (1-5, 7-17).

The high specificity of triplex forming oligonucleotides for their cognate sites (~1000 fold for a single base internal mismatch) in conjunction with the direct readout of hydrogen bond recognition, where there is no folding problem to solve, make third strand oligonucleotides extremely appealing for *in-vitro* recognition of double helical DNA (12, 18, 19). The match and mismatch binding affinities of oligonucleotides containing the four natural base pairs at a single variable position targeted to a cognate site containing each of the four base pairs at that single variable position are shown in Table 1 (12). The major drawback to the utilization of triplex forming oligonucleotides is the sequence requirement. With the natural bases, only purines can be recognized with high affinity and specificity through the formation of C•GC and T•AT triplets in the pyrimidine motif or G•GC and T•AT or A•AT triplets in the purine motif. In the

Table 1. The equilibrium association constants (M^{-1}) for the natural bases are reported. The oligonucleotide sequence is 5' TTCTCTBTCTCCTTT 3', where C is 5-methyl C and titrations were performed at 22° C, pH 7.0, with 100 mM NaCl, 10 mM Bis Tris, and 0.25 mM spermine. This data was originally reported in reference (12).

BASE	AT	CG	GC	TA
^{Me}C	8.2 (±1.9) x 10^4	1.7 (±0.6) x 10^5	**>1.4 x 10^8**	<10^4
T	**>2.1 x 10^8**	5.4 (±0.7) x 10^5	<10^4	<10^4
G	<10^4	<10^4	<5 x 10^4	1.1 (±0.3) x 10^6
A	1.6 (±0.6) x 10^5	<10^4	<10^4	<10^4

pyrimidine motif, G can recognize TA at a single internal position with 30 fold reduced affinity and a net specificity of >20 fold (12, 20). An NMR structure containing a GT•A triplet shows interactions with both the proximal and 3' base resulting in a distorted third strand χ angle in both the G•TA triplet and in the triplet 5' to it. This triplet is shown in Figure 1 (21). Thus, the design of nonnatural bases for the recognition of pyrimidines remains the major unsolved problem limiting the generality of the triple helix.

Two bases in the literature that have been reported to bind AT and CG nonspecifically are N^4-3-acetamidopropyl cytosine and N^4-(6-aminopyridyl) cytosine (22, 23). The chemical structures of these bases are shown in Figures 2a and 2b. Subsequent quantitative examination of N^4-(6-aminopyridyl) cytosine indicates that the amino-benzoyl protecting group was not removed during deprotection and that the resulting base binds with only moderate affinity and low specificity (24).

The novel nucleoside D_3 (1-(2-deoxy-β-D-ribofuranosyl)-4-(-3-benzamido) phenylimidazole) is depicted in Figure 2. It has been previously synthesized and that synthesis is shown in Figure 3 (25). Qualitative affinity cleavage indicated that the order of strength of recognition for this nucleoside was TA>CG>AT>GC (25, 26). It was also seen that the base to the 3' side in the target strand played an important role in selectivity, where an adenosine enabled stronger binding than a guanosine (26). The structure of an intramolecular triplex containing a D_3•TA triplet has been elucidated by NMR and molecular modeling (27, 28). D_3 shows an intercalative binding mode, where the benzoyl group stacks between the purine on the complementary strand and the 5' base, illuminating the sequence dependence of the neighboring bases. This structure is depicted in Figure 4.

Several studies were undertaken in the characterization of D_3 containing oligonucleotides to explore their utility as a solution to the pyrimidine recognition

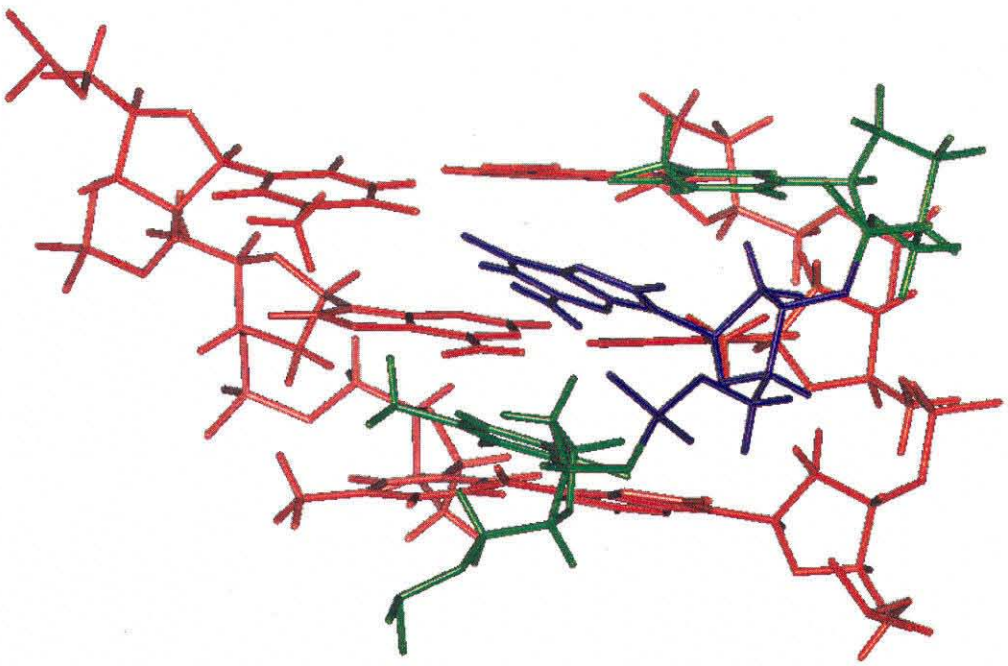


Figure 1. The G•TA triplet is shown in the context of its neighboring bases in an NMR structure (21). Here the G residue in blue is twisted and shows an altered χ angle. The register is distorted and the neighboring bases, shown in green are no longer coplanar.

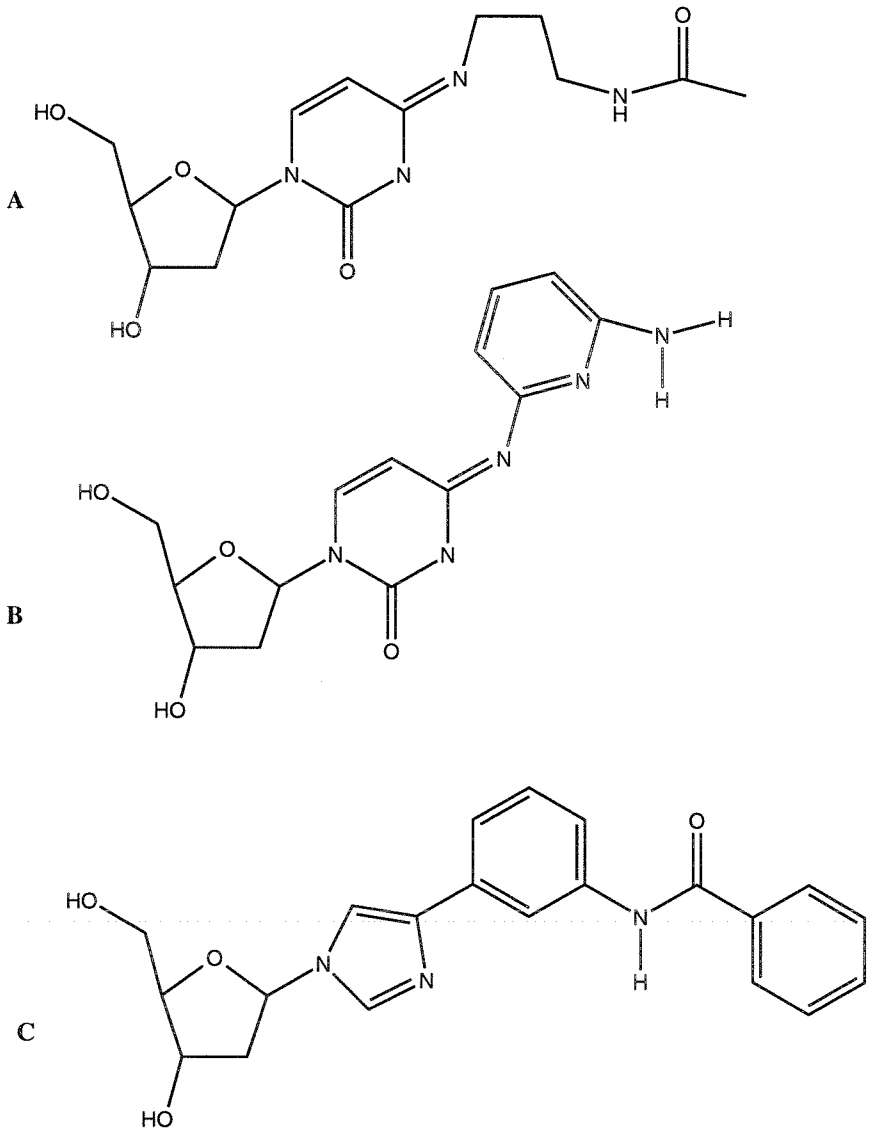


Figure 2. The chemical structures of N^4-(3-acetamidopropyl)cytosine (A), N^4-(6-aminopyridinyl)cytosine (B), and D_3 (C) are depicted.

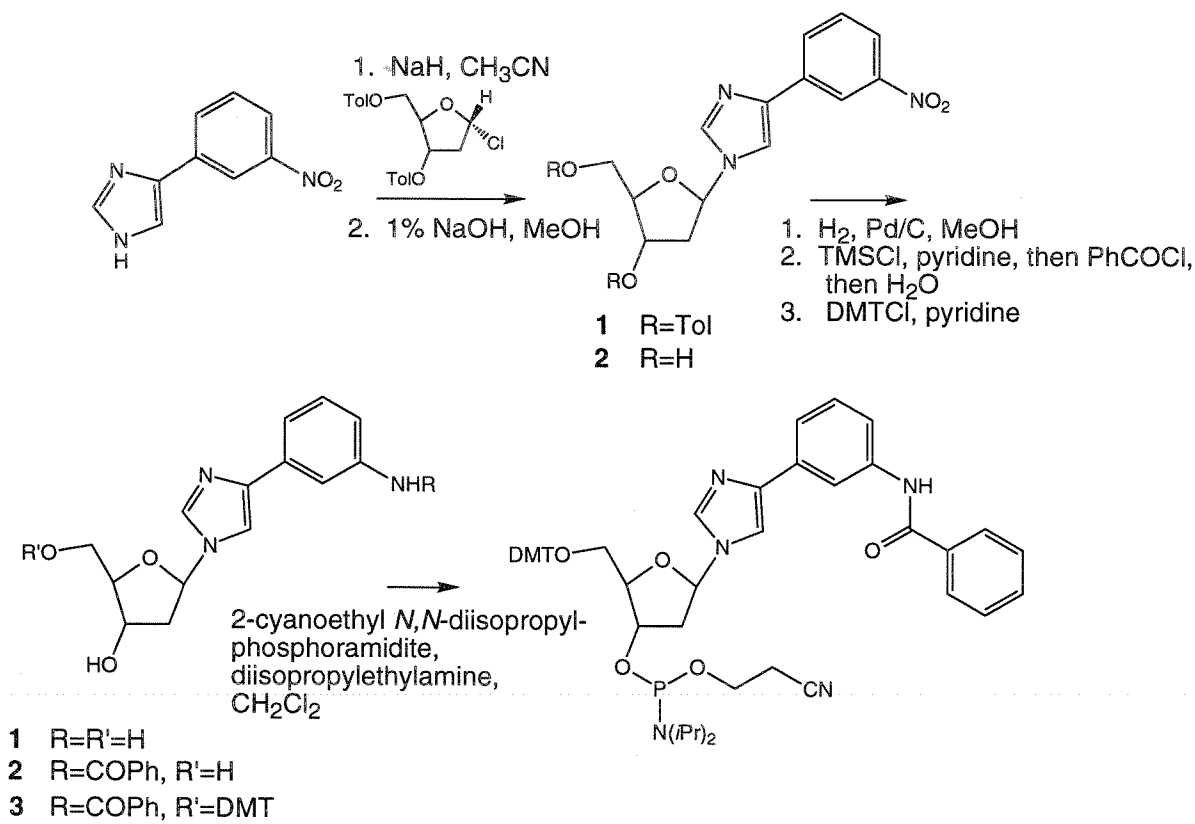


Figure 3. Scheme for the synthesis of novel nucleoside D_3, as previously reported by (25).

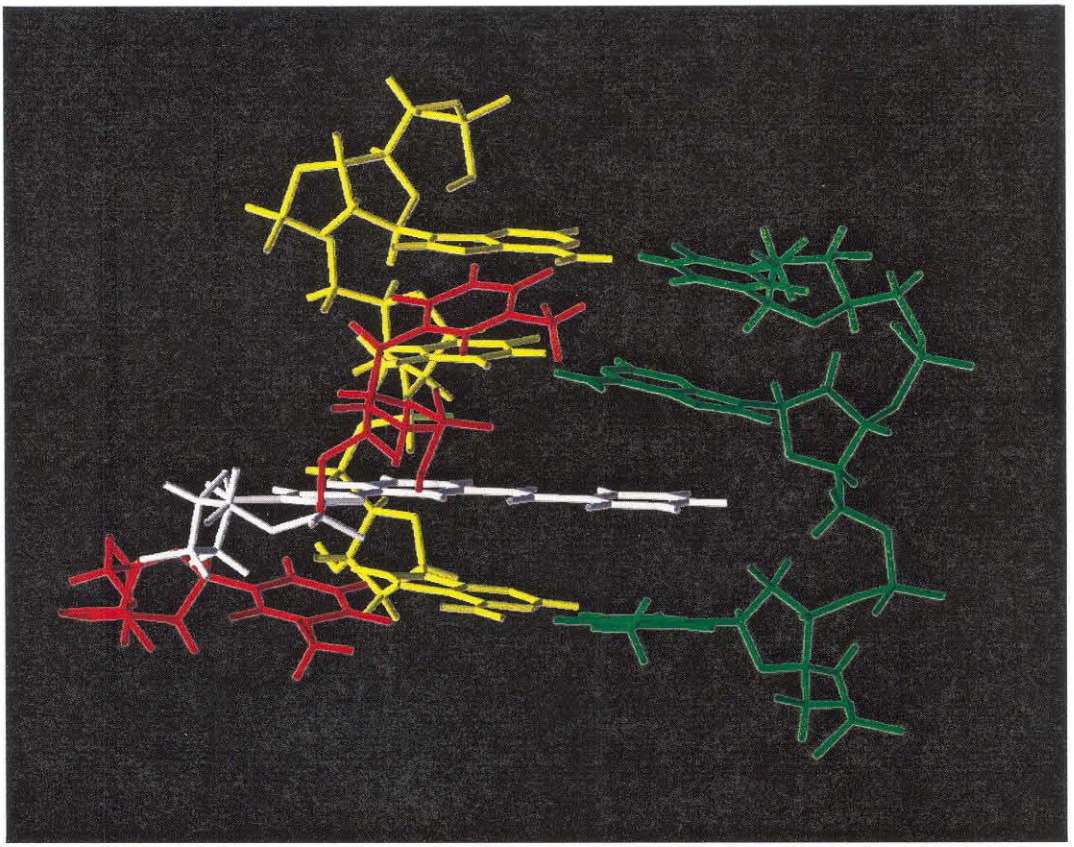


Figure 4. The NMR structure of a D_3•TA triplet within a triple helix shows the benzoyl group of the D_3 residue intercalated over the A residue of the pyrimidine rich strand (27, 28). The register of Hoogsteen base pairing is disrupted due to this intercalation event.

problem. First, the specificity of D_3 containing oligonucleotides was examined quantitatively using DNase I footprinting. Second, multiple D_3 nucleosides were incorporated into a single oligonucleotide to examine the generality of this solution. Finally, as an intercalator, the DNA bending properties of the D_3 wedge were examined.

MATERIALS AND METHODS

General. Phosphoramidites were purchased from Glen Research and Biogenex. DNA synthesis was performed using standard methods on an Applied Biosystems 380B DNA synthesizer. Oligonucleotide purification was performed on a Pharmacia FPLC using ProRPC reverse phase and mono Q anion exchange columns. Enzymes were purchased from Boehringer Mannheim with the exception of Sequenase 2.0, which was purchased from United States Biochemical. Deoxyadenosine 5'-[α-^{32}P]-triphosphate was obtained from Amersham. Storage Phosphor autoradiography was performed using a Molecular Dynamics 400S Phosphorimager and ImageQuant software.

Oligonucleotide characterization. Intact oligonucleotides were characterized by HPLC on a reverse phase column (Vydac). Oligonucleotides were also digested with snake venom phosphodiesterase and alkaline phosphate and analyzed by reverse phase HPLC. A D_3 nucleoside standard was found to coelute with the expected peak.

Preparation of labeled DNA. Single D_3 binding experiments were performed on previously described plasmids (11). Other plasmids were constructed by standard methodologies (29). Plasmids were cleaved with the appropriate restriction enzymes and simultaneously filled in with Sequenase 2.0, the appropriate 5'-[α-^{32}P]

nucleoside triphosphates, and nonradioactive deoxynucleoside triphosphates. Fragments were isolated by nondenaturing polyacrylamide gel electrophoresis, treated with proteinase K, filtered, precipitated, phenol/chloroform/isoamyl alcohol (25:24:1) extracted, and passed through a NICK size exclusion column (Pharmacia).

Quantitative DNase I footprinting. Oligonucleotide titrations were performed at 22° C, pH 7.0 in the presence of 250 μM spermine, 100 mM NaCl, 10 mM bis tris, and 100 μM (in base pairs) calf thymus DNA. Equilibrations were carried out for at least 24 hours in all cases and for at least 36 hours for the single D_3 experiments. Samples were equilibrated in a total volume of 180 μl. Workup of samples and data was performed as previously described (30). Adenine specific reaction control lanes were also prepared as previously described (31). The gels depicted are 8% denaturing polyacrylamide gels analyzed by storage phosphor densitometry and displayed with a linear dynamic range of 50. The values reported are the average of four footprint titration experiments ± SEM.

Phasing analysis. Phasing analysis was performed on a restriction fragment containing phased A tracts linked by a mixed sequence of 3, 5, or 8 residues to the target site (32). Analysis was performed by equilibrating samples for at least two hours at 22° C, pH 5.5, with 5 μM oligonucleotide, 45 mM Mes, 1 mM $MgCl_2$, and ^{32}P-3'-end-labeled restriction fragment in a volume of 10 μl. One microliter of 15% glycerol loading buffer was then added. Samples were run on a 10% polyacrylamide gel at 4° C, pH 5.5, with a 75:1 acrylamide/bis-acrylamide ratio in 45 mM Mes and 1 mM $MgCl_2$ with buffer recirculation.

RESULTS AND DISCUSSION

ENERGETICS OF OLIGONUCLEOTIDES CONTAINING D_3 AT A SINGLE POSITION

D_3 phosphoramidite was synthesized as previously described (25). The yields obtained for each step were similar to those previously presented. D_3 was incorporated into an oligonucleotide using the commonly used Moser sequence, shown in Figure 5 (1, 11). A standard set of conditions for thermodynamic comparisons was set and these conditions, 22° C, pH 7.0, 250 μM spermine, 100 mM NaCl, 10 mM bis-tris, and 100 μM (in base pairs) calf thymus DNA, were used (11, 12). Sample gels for all six determinations are shown in the following figures. Figure 6 shows a gel where the X position (the variable position in the third strand) oligonucleotide contains ^{Me}C, with GC in the YZ position (the variable positions in the target double helix). Figure 7 shows ^{Me}C•AT and Figures 8-11 show D_3 opposite AT, CG, GC, and TA. The isotherms obtained are shown in Figure 12, with the resulting equilibrium association constants in Table 2.

D_3 binding to TA in a duplex at a single position was just as strong as a ^{Me}C•GC match, and D_3•CG was almost as strong. The most disfavored mismatch was D_3•GC, which was a factor of 178 worse than D_3•TA. In contrast, the differential between T•AT and T•GC at a single position using a different sequence context under the same binding conditions was reported to be >20,000 (12). Similar ordering effects were obtained in studies with D_3 using quantitative affinity cleavage that are not reported here. The specificity ranges were reduced, as was also seen for the natural bases using

1. 5' TTTTTC̲TC̲TC̲TC̲TC̲T 3'
2. 5' TTTTTC̲TC̲TD$_3$TC̲TC̲T 3'

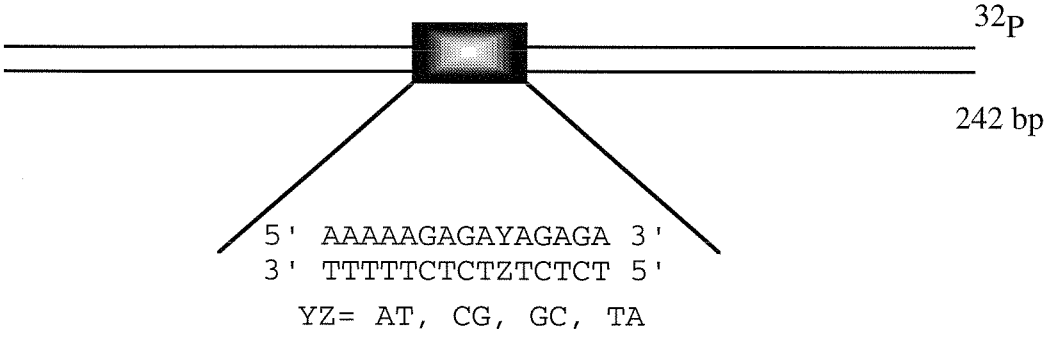


Figure 5. The oligonucleotides **1** and **2** are targeted against the fragment indicated and bind their designated sites. In the oligonucleotide sequence, C̲ indicates 5-methyl cytosine.

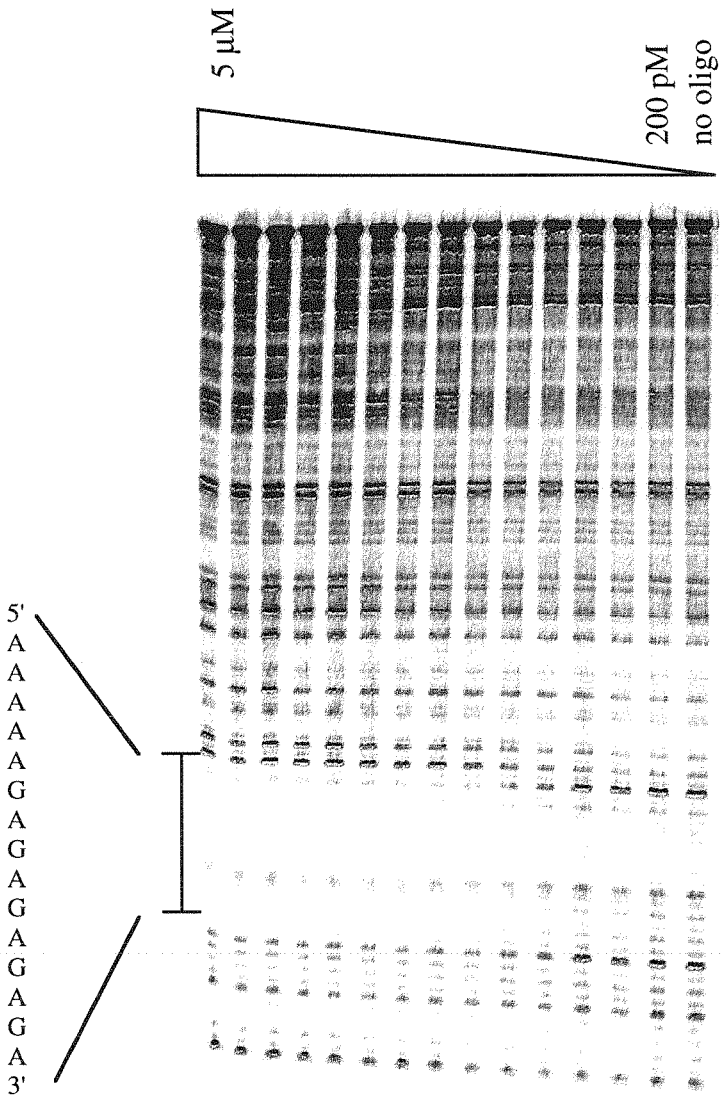


Figure 6. This sample DNase I footprint titration shows ^{Me}C binding to GC within the context of a single variable X•YZ base triplet at its cognate site.

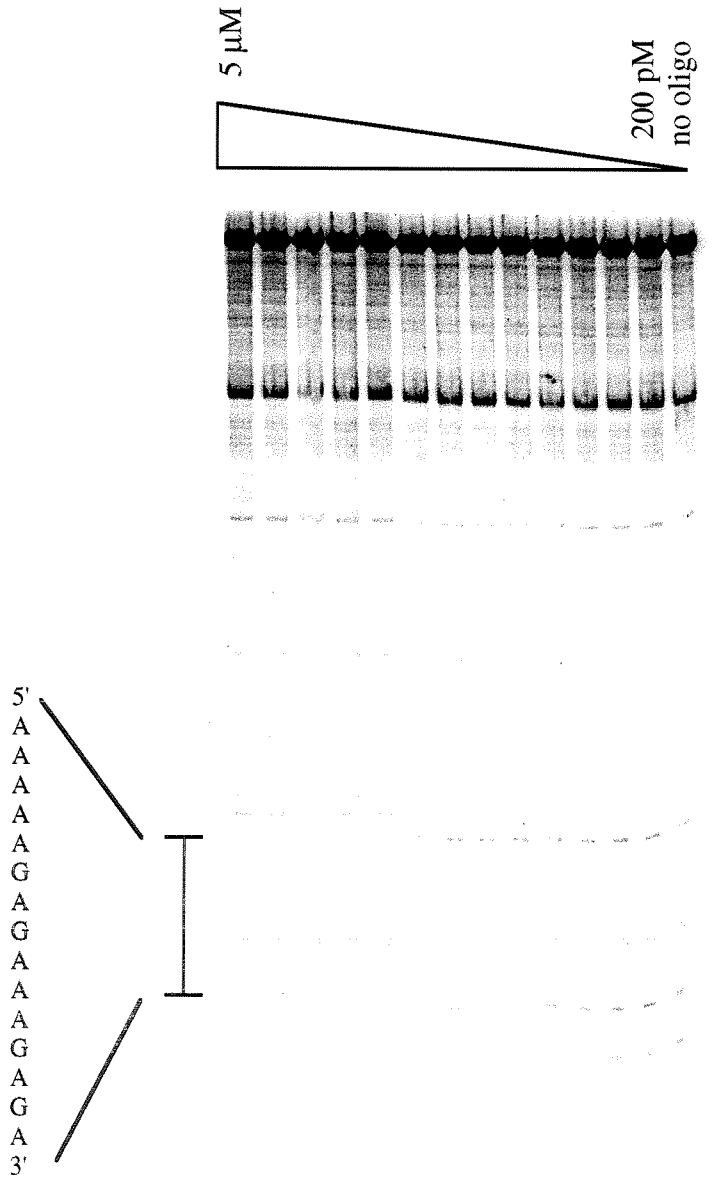


Figure 7. This sample DNase I footprint titration shows ^{Me}C binding to AT within the context of a single variable X•YZ base triplet at its cognate site.

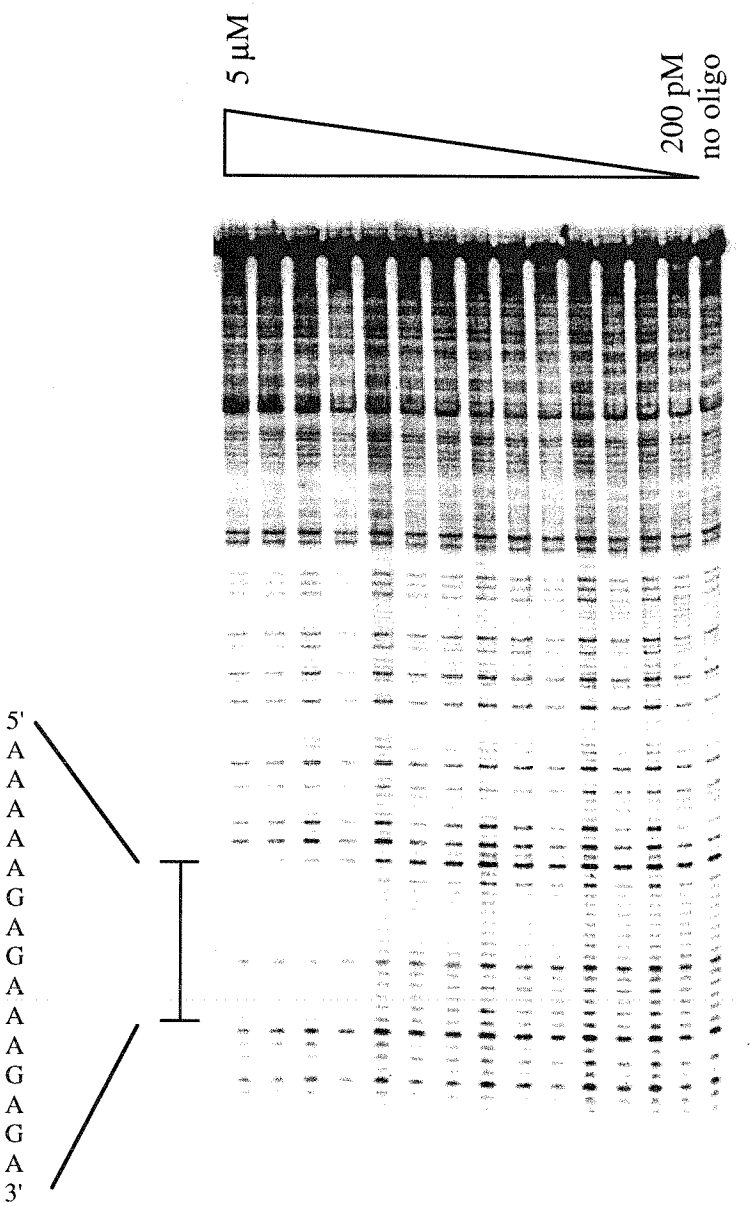


Figure 8. This sample DNase I footprint titration shows D_3 binding to AT within the context of a single variable X•YZ base triplet at its cognate site.

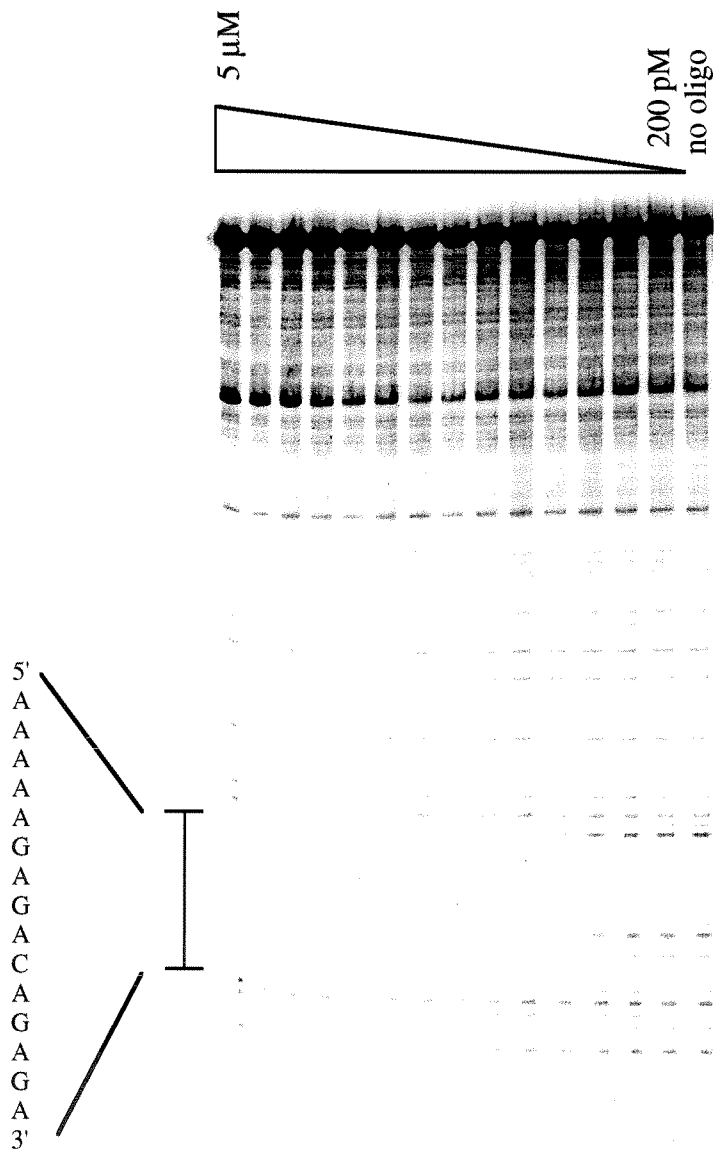


Figure 9. This sample DNase I footprint titration shows D_3 binding to CG within the context of a single variable X•YZ base triplet at its cognate site.

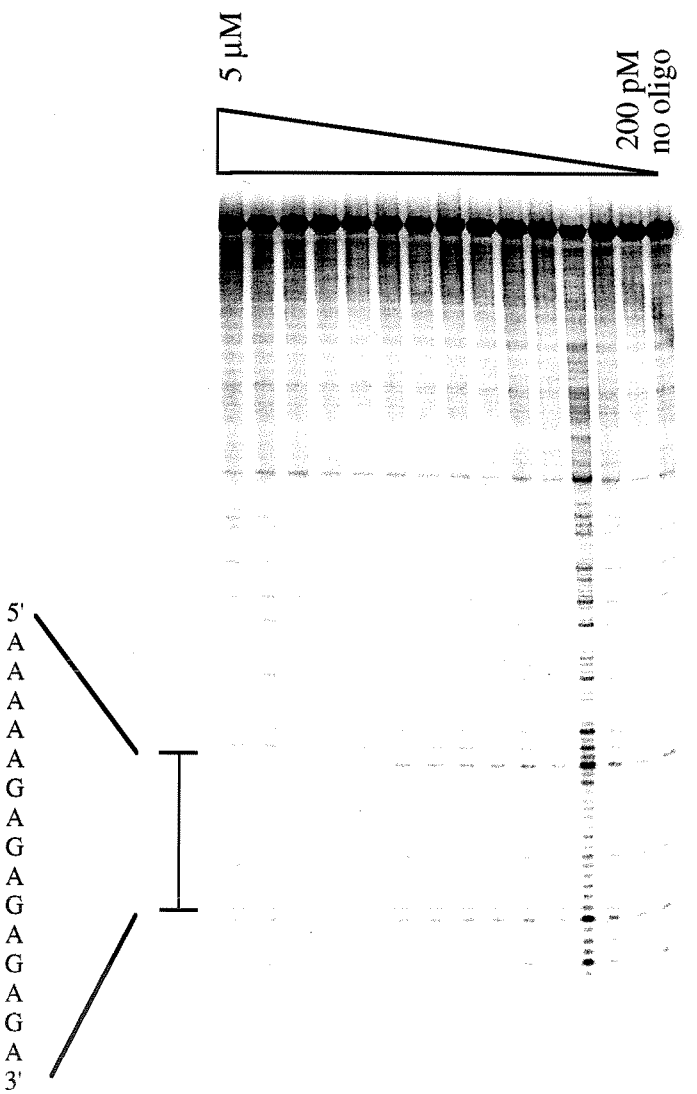


Figure 10. This sample DNase I footprint titration shows D_3 binding to GC within the context of a single variable X•YZ base triplet at its cognate site.

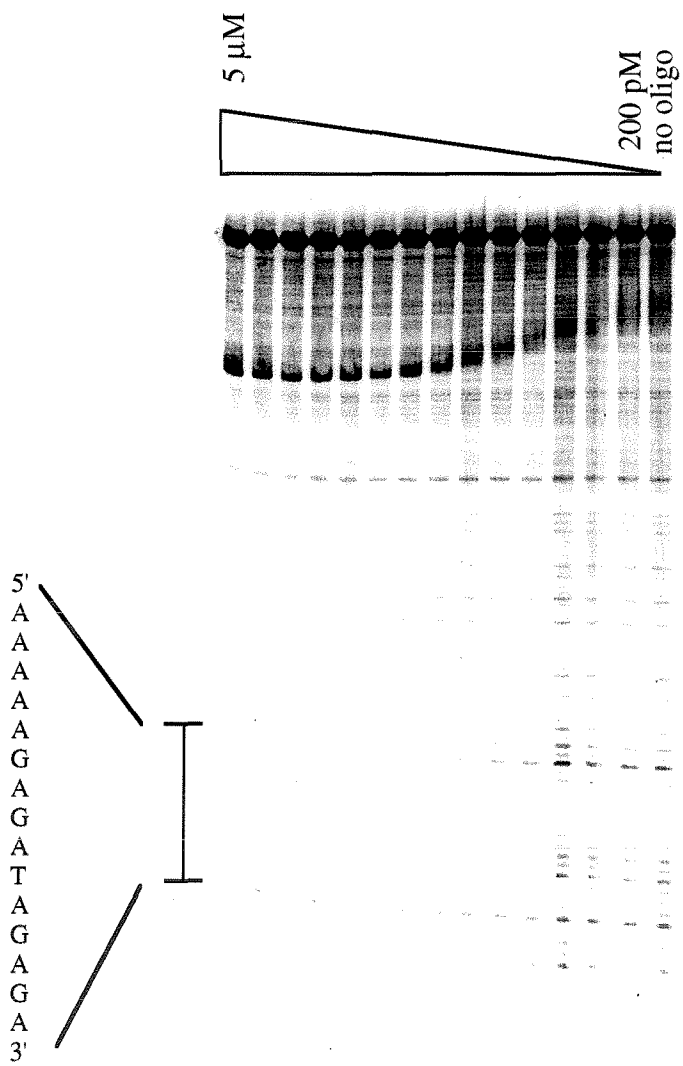


Figure 11. This sample DNase I footprint titration shows D_3 binding to TA within the context of a single variable X•YZ base triplet at its cognate site.

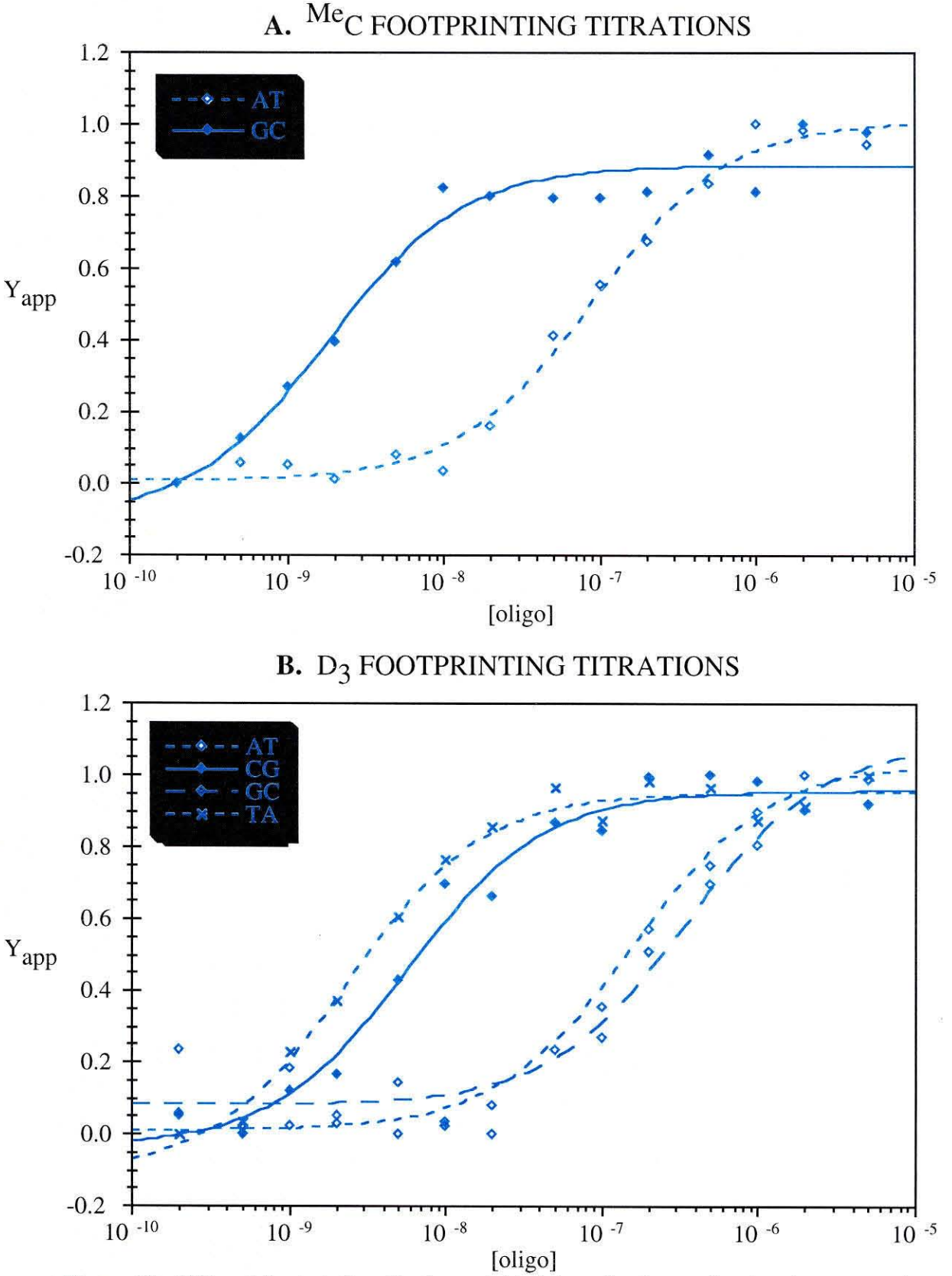


Figure 12. DNase I footprinting titrations yielded these isotherms for the six measured equilibrium association constants involving ^{Me}C containing oligonucleotides in **A** and D_3 containing oligonucleotides in **B**.

Table 2. The values obtained from quantitative DNase I footprinting titrations on D_3 containing oligonucleotides are reported. These values are the average of four experiments ± SEM.

	D_3AT	D_3CG	D_3GC	D_3TA
K_{eq} (M^{-1})	4.6 (±2.1) x 10^6	2.1 (±0.5) x 10^8	2.7 (±1.4) x 10^6	4.8 (±1.1) x 10^8
ΔG (kcal/mole)	-9.1	-11.3	-8.7	-11.8
ΔΔG (kcal/mole)	+0.1	-2.1	-0.5	-2.6

	^{Me}CAT	^{Me}CGC
K_{eq} (M^{-1})	6.3 (±2.1) X 10^6	4.7 (±1.1) X 10^8
ΔG (kcal/mole)	-9.2	-11.8
ΔΔG (kcal/mole)	0.0	-2.6

that methodology (11). The reason for measurement of a lower specificity using quantitative affinity cleavage is unclear.

The binding properties of D_3 in triple helix formation can be likened to those of inosine in double helix formation. While inosine recognition of single stranded DNA is through hydrogen bonding on the Watson-Crick face rather than through intercalation, functionally D_3 can be viewed as partially degenerate in triple helix formation when used at a single position in an analogous fashion to inosine. For inosine, the T_m of a single inosine containing duplex (13 base pairs) ranges from 43.4° C for IT to 50.9° C for IC, while those of the four natural bases range from 34.8° C for AC to 52.8° C for GC (33). D_3 may then be useful at a single position in probes targeting a single pyrimidine in a polypurine tract or a degenerate position.

MULTIPLE D_3 CONTAINING OLIGONUCLEOTIDES

D_3 was then targeted to sequences containing multiple pyrimidines. One important sequence studied in the group was that of the Huntington's disease gene repeat (34). With the cloning of the Huntington's Disease gene, it had become apparent that triplet repeat diseases, in which the expansion of trinucleotide repeats was correlated with phenotypic severity, were an increasingly common genetic phenomenon. Huntington's Disease and several other diseases bear $(CAG)_n$ repeats (35). Targeting this repeat with a third strand oligonucleotide would enable the search for homologues or other repeat diseases. A DNase I footprinting experiment was designed, as shown in Figure 13, where (CAG) repeats were targeted. This gel is depicted in Figure 14, where no binding of oligonucleotide was detected with concentrations as high as 5 μM.

The lack of binding in this experiment led to another attempt at incorporating multiple D_3 residues in an oligonucleotide. This attempt targeted a TA instead of a CG

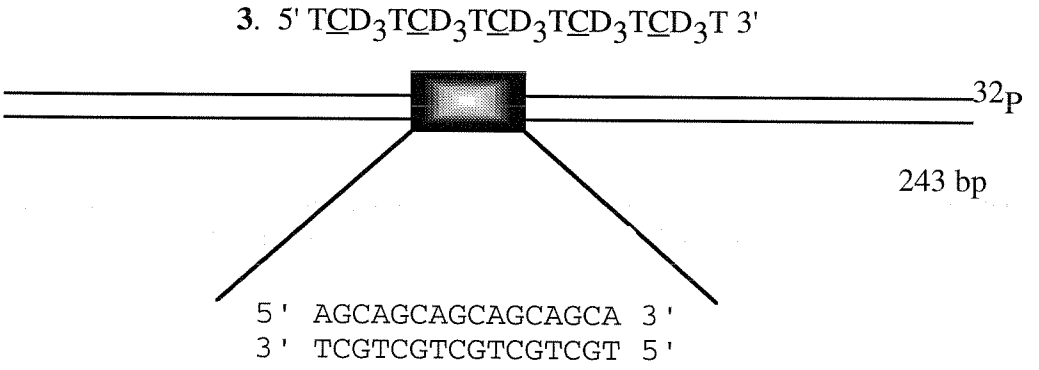


Figure 13. The oligonucleotide **3** is targeted against the fragment indicated. In the oligonucleotide sequence, C indicates 5-methyl cytosine.

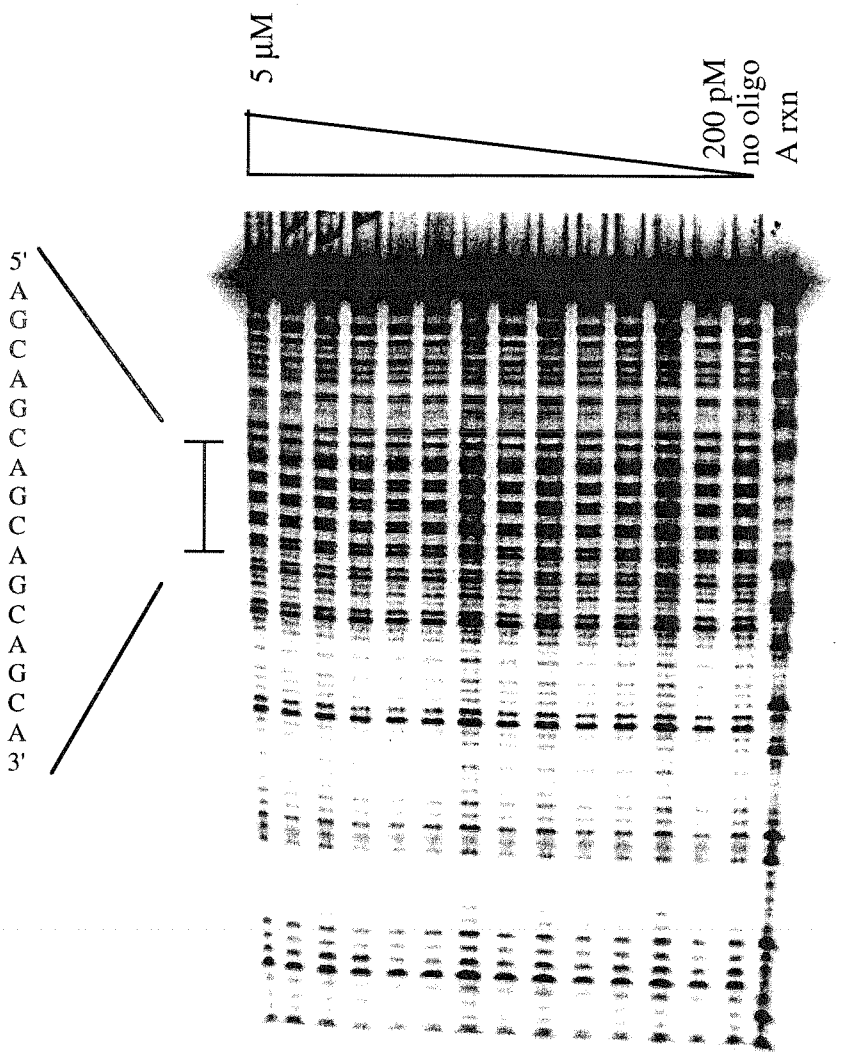


Figure 14. This sample DNase I footprint titration shows a D_3 containing third strand oligonucleotide targeted towards CAG repeats. No protection of the site is seen.

base pair, which is bound with a slightly higher affinity. Additionally, oligonucleotides were made containing three contiguous D_3 residues with the thought that helix unwinding to accommodate intercalation would only require readjusting of the register at the two junctions. This experimental setup is shown in Figure 15 with a resulting gel in Figure 16. The designed oligonucleotide showed only weak protection at the high concentration of 5 μM. In summary of these experiments, oligonucleotides containing multiple D_3 residues did not bind with high affinity to their cognate sites.

DNA BENDING USING D_3 CONTAINING OLIGONUCLEOTIDES

The binding mode of D_3 precludes the incorporation of multiple residues in a single oligonucleotide. Given the intercalative binding mode of D_3, bending was examined to see if the benzoyl group served as a wedge analogous to the proline residues in IHF (36). A phasing analysis was performed on an oligonucleotide containing a single D_3 (32). No major bend was detected using this methodology, as seen in Figure 17. No significant bend was also detected for oligonucleotides containing G•TA or ^{Me}C•GC triplets in this assay. The lack of bending for D_3 containing oligonucleotides may be due to intercalation over both strands, which causes unwinding instead of bending. While proline residues in the bent IHF structure and phenylalanine residues in the bent TATA binding protein (TBP) structure only intercalate over one strand, the designed intercalator Λ-1-Rh(MGP)$_2$phi^{5+} interacts with both strands in unwinding DNA 70° without bending it (36-39).

While it may be possible to design a DNA bending ligand based upon an intercalating base within a third strand oligonucleotide, alternative strategies for DNA bending were explored. Chapters 3 and 4 of this report explore this possibility.

4. 5' TTCTTD$_3$D$_3$D$_3$TTCCTTT 3'

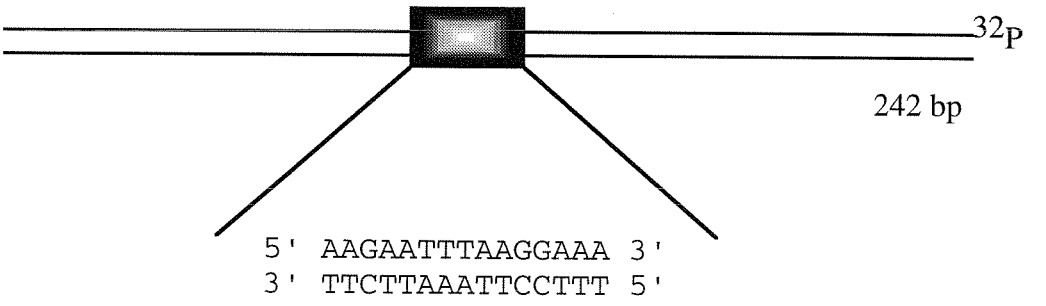


Figure 15. The oligonucleotide **4** is targeted against the fragment indicated. In the oligonucleotide sequence, C indicates 5-methyl cytosine.

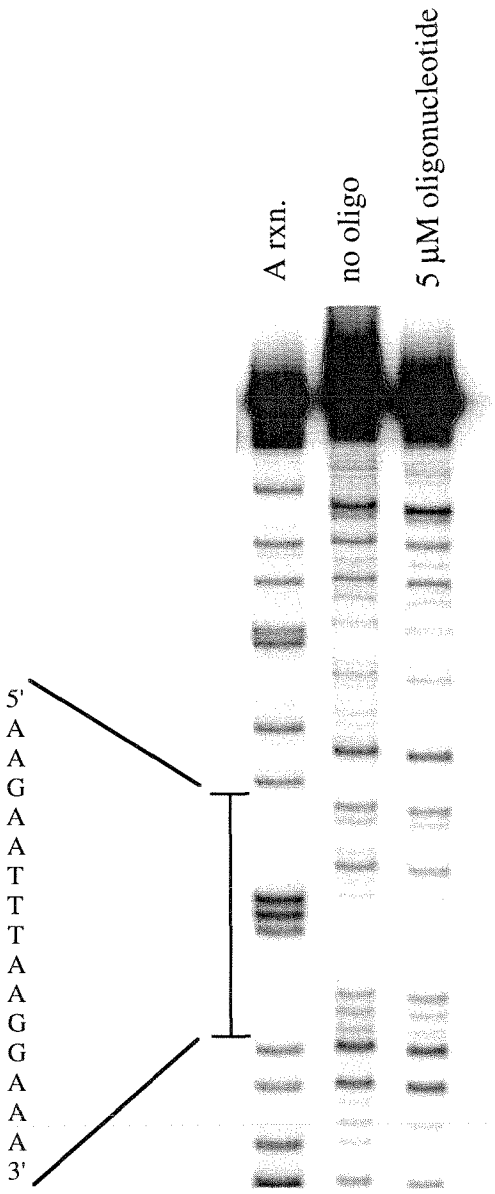


Figure 16. This sample DNase I footprint titration shows a multiple D_3 containing third strand oligonucleotide targeted towards the modified Priestley sequence. Only partial protection of the site is seen.

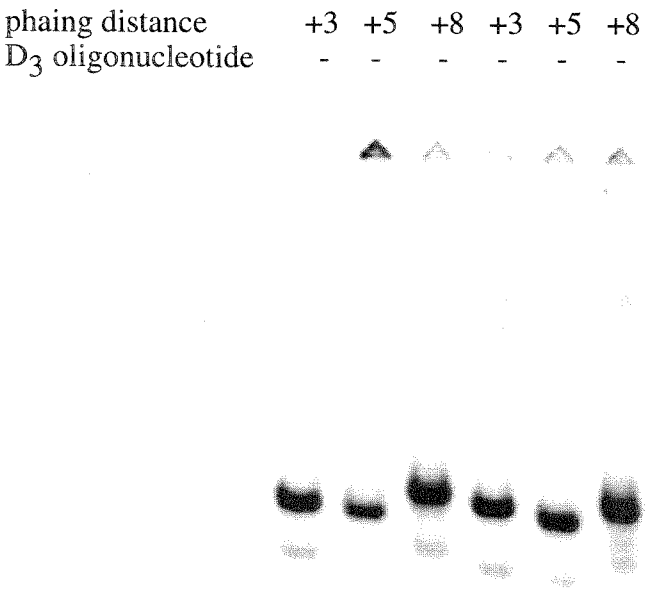


Figure 17. An oligonucleotide of sequence 5' $TT^{Me}CT^{Me}CTD_3T^{Me}CT^{Me}C^{Me}CTTT$ 3' was targeted against its cognate site containing a TA base pair at the XY position. Three plasmids were constructed, with inserts of +3, +5, and +8 separating the binding site and the phased A tracts. While the mobilities are slightly different, no significant bend is caused by the binding of this oligonucleotide.

REFERENCES

1. Moser, H. E. & Dervan, P. B. (1987) *Science* **238,** 645-650.
2. LeDoan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J., Thuong, N. T., Lhomme, J. & Helene, C. (1987) *Nucleic Acids Research* **15,** 7749-7760.
3. Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988) *Science* **241,** 456-459.
4. Beal, P. A. & Dervan, P. B. (1991) *Science* **251,** 1360-1363.
5. Beal, P. A. & Devan, P. B. (1992) *Nucleic Acids Research* **20,** 2773-2776.
6. Greenberg, W. A. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 5016-5022.
7. Singleton, S. F. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 6957-6965.
8. Roberts, R. W. & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci., USA* **88,** 9397-9401.
9. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannageli, C., Garestier, T. & Helene, C. (1992) *Biochemistry* **31,** 9269-9278.
10. Fosella, J. A., Kim, Y. J., Shih, H., Richards, E. G. & Fresco, J. R. (1993) *Nucleic Acids Research* **21,** 4511-4515.
11. Best, G. C. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 1187-1193.
12. Hunziker, J., Priestley, E. S., Brunar, H. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 2661-2662.
13. Povsic, T. J. & Dervan, P. B. (1989) *J. Am. Chem. Soc.* **111,** 3059-3061.
14. Xodo, L. E., Manzini, G., Quadrifoglio, F., vanderMarel, G. A. & vanBoom, J. H. (1991) *Nucleic Acids Research* **19,** 5625-5631.
15. Singleton, S. F. & Dervan, P. B. (1992) *Biochemistry* **31,** 10995-11003.
16. Singleton, S. F. & Dervan, P. B. (1993) *Biochemistry* **32,** 13171-13179.

17. Singleton, S. F. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 10376-10382.

18. Maher, L. J., Wold, B. & Dervan, P. B. (1989) *Science* **245,** 725-730.

19. Maher, L. J., Dervan, P. B. & Wold, B. (1990) *Biochemistry* **31,** 70-81.

20. Griffin, L. C. & Dervan, P. B. (1989) *Science* **245,** 967-971.

21. Radhakrishnan, I. & Patel, D. J. (1994) *Structure* **2,** 17-32.

22. Huang, C., Cushman, C. D. & Miller, P. S. (1993) *J. Org. Chem.* **58,** 5048-5049.

23. Huang, C. & Miller, P. S. (1993) *J. Am. Chem. Soc.* **115,** 10456-10457.

24. Greenberg, W. A. & Dervan, P. B. *unpublished results* .

25. Griffin, L. C., Kiessling, L. L., Beal, P. A., Gillespie, P. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 7976-7982.

26. Kiessling, L. L., Griffin, L. C. & Dervan, P. B. (1992) *Biochemistry* **31,** 2829-2834.

27. Koshlap, K. M., Gillespie, P., Dervan, P. B. & Feigon, J. (1993) *J. Am. Chem. Soc.* **115,** 7908-7909.

28. Wang, E., Koshlap, K. M., Gillespie, P., Dervan, P. B. & Feigon, J. (1996) *J. Molecular Biology* **257,** 1052-1069.

29. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

30. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986) *Proc. Natl. Acad. Sci., USA* **83,** 8462-8466.

31. Iverson, B. L. & Dervan, P. B. (1987) *Nucleic Acids Research* **15,** 7823-7830.

32. Zinkel, S. S. & Crothers, D. M. (1987) *Nature* **328,** 178-181.

33. Kawase, Y., Iwai, S., Inoue, H., Miura, K. & Ohtsuka, E. (1986) *Nucleic Acids Research* **14,** 7727-7736.

34. Huntington's_Disease_Collaborative_Research_Group (1993) *Cell* **72,** 971-983.

35. Winnacker, E. L. (1993) *Angew. Chem. Int. Ed. Engl.* **32,** 1415-1418.

36. Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. (1996) *Cell* **87,** 1295-1306.

37. Kim, J. L., Nikolov, D. B. & Burley, S. K. (1993) *Nature* **365,** 520-527.

38. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. (1993) *Nature* **365,** 512-520.

39. Terbrueggen, R. H. & Barton, J. K. (1995) *Biochemistry* **34,** 8227-8234.

CHAPTER 3

DESIGN OF ARTIFICIAL SEQUENCE SPECIFIC DNA BENDING LIGANDS

The text of this chapter is adapted from a published manuscript that was coauthored with Professor Peter B. Dervan.

Liberles, D.A., and Dervan, P.B. 1996. *Proc. Natl. Acad. Sci., USA* **93**, 9510-9514.

DNA bending proteins play a role in site-specific recombination, transposition, replication, viral integration, and transcription (1-12). At least three different mechanisms for protein-DNA bending have emerged. These include wedging intercalating groups into the helix, binding at two distal sites and looping the DNA around the protein, and neutralizing charge at the phosphodiester backbone (13-21). DNA bending proteins are important both for locally distorting DNA at that site and for global molding of DNA conformation to promote the binding and action of distal factors. The design of nonnatural site-specific DNA bending ligands could be envisioned through one of the strategies utilized by proteins.

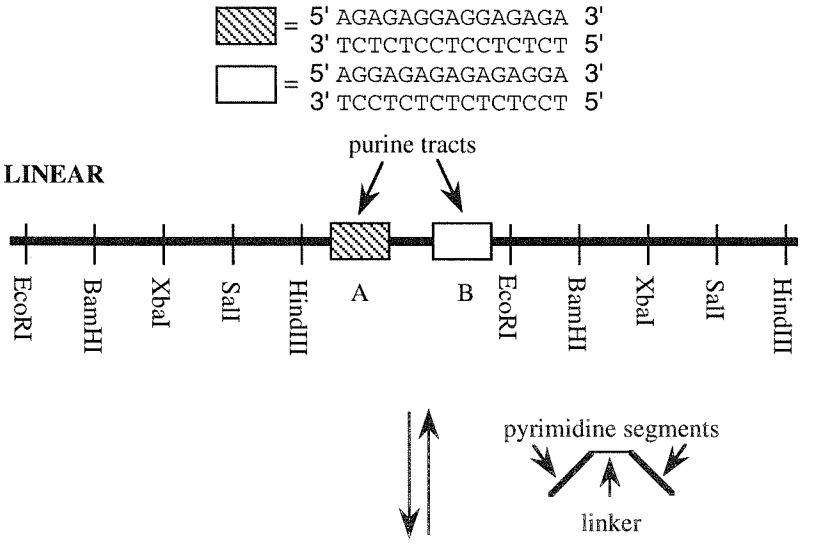
Oligonucleotide-directed triple helix formation is a general method for the sequence-specific recognition of double helical DNA. Two classes of triple helix forming oligonucleotides have been identified (22-26). In one class, pyrimidine oligonucleotides bind double helical DNA in the major groove parallel to the purine tract through the formation of specific Hoogsteen hydrogen bonds (22-36). Using triple helix formation for DNA recognition in the major groove, a bifunctional oligonucleotide was designed to mimic a DNA bending protein. DNA bending was accomplished by binding two distal DNA sites separated by 10 base pairs with oligonucleotides **1-6**

containing two DNA binding domains at the 5' and 3' ends connected by a variable spacer (Figure 1).

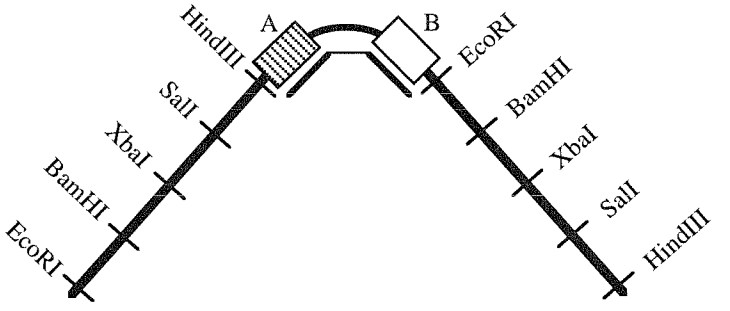
MATERIALS AND METHODS

General. Phosphoramidites were purchased from Glen Research and Biogenex. DNA synthesis was performed using standard methods on an Applied Biosystems 380B DNA synthesizer. Oligonucleotide purification was performed on a Pharmacia FPLC using ProRPC reverse phase and mono Q anion exchange columns. Enzymes were purchased from Boehringer Mannheim with the exception of Sequenase 2.0, which was purchased from United States Biochemical. Deoxyadenosine 5'-[α-^{32}P]-triphosphate was obtained from Amersham. Storage Phosphor autoradiography was performed using a Molecular Dynamics 400S Phosphorimager and ImageQuant software. Autoradiography utilized Kodak X-OMAT AR film without an intensifying screen.

Preparation of Labeled DNA. For the circular permutation assay, 101 base pair 3' end labeled restriction products were generated by digestion with the designated restriction enzyme and simultaneously filled in with Sequenase 2.0, deoxyadenosine 5'-[α-^{32}P]-triphosphate, and nonradioactive deoxynucleoside triphosphates. The 101 base pair fragment was isolated by nondenaturing 5% polyacrylamide gel electrophoresis and the purified fragment was treated with proteinase K, filtered, further extracted with phenol/chloroform, and ethanol precipitated. The 577-585 base pair fragments for phasing assays were generated by cleavage of pUC 19 derived plasmids with Nde I and Sap I (New England Biolabs) and purified as above.



BENT




```
1    5' TCTCTCCTCCTCTCTTTCCTCTCTCTCTCCT 3'
2    5' TCTCTCCTCCTCTCTTTTCCTCTCTCTCTCCT 3'
3    5' TCTCTCCTCCTCTCTTTTTCCTCTCTCTCTCCT 3'
4    5' TCTCTCCTCCTCTCTTTTTTCCTCTCTCTCTCCT 3'
5    5' TCTCTCCTCCTCTCTTTTTTTCCTCTCTCTCTCCT 3'
6    5' TCTCTCCTCCTCTCTTTTTTTTCCTCTCTCTCTCCT 3'
7    5' TCTCTCCTCCTCTCT 3'
8    5' TCCTCTCTCTCTCCT 3'
9    5' TCTCTCCTCCTCTCTTTTTTTTTTTTTTTTT 3'
10   5' TTTTTTTTTTTTTTTTCCTCTCTCTCTCCT 3'
```

Figure 1. Circular permutation analysis of ligand induced bending. The 101 base pair probes used for circular permutation analysis were generated by restriction enzyme cleavage at sites within tandem polylinkers surrounding the ligand binding site. The plasmid contains the sequence 5' GAATTCGAGCTCGGTACCCGGGGATCCTCTA GAGTCGACCTGCAGGCATGCAAGCTT**AGAGAGGAGGAGAGA**GCGGTG CGGT**AGGAGAGAGAGAGGA**TGAATTCGAGCTCGGTACCCGGGGATCCT CTAGAGTCGACCTGCAGGCATGCAAGCTT 3'. The ligands designed to bend DNA sequence specifically were two third strand oligonucleotide domains containing 5-methyl C (indicated by C) and T residues linked by a variable number of T residues.

Quantitative DNase I Footprint Titrations. The binding energy was calculated from the equilibrium association constants of the oligonucleotides 5' TCTCTCCTCCTCTCT 3' and 5' TCCTCTCTCTCTCCT 3' (where C is 5-methyl cytosine) binding to their cognate sites by quantitative DNase I footprint titrations. The values reported are the mean value of four footprint titration experiments ± SEM. Oligonucleotide titrations were performed at 22° C, pH 5.5 in the presence of 45 mM Mes, 1 mM $MgCl_2$, and ^{32}P 3' labeled target DNA and allowed to equilibrate for at least 24 hours before footprinting. Footprinting reactions and data analysis were carried out as previously described (27, 37).

Gel Shift Assay. The gel shift assay was performed by equilibrating samples for at least 1 hour at 22° C, pH 5.5 with 10 μM oligonucleotide, 45 mM Mes, 1 mM $MgCl_2$, and ^{32}P 3' labeled EcoRI cut target DNA in a volume of 10 μl. To this was added 1 μl of 15% glycerol loading buffer. Samples were run on a 10% polyacrylamide gel with a 75:1 acrylamide: bis-acrylamide ratio in 45 mM Mes, 1 mM $MgCl_2$, and pH 5.5 at 4° C with buffer recirculation.

Circular Permutation Analysis. Circular permutation analysis was performed under similar gel shift conditions. Experiments were performed with equilibrations at 22° C, pH 5.5, 45 mM Mes, and 1 mM $MgCl_2$ containing 20 μM oligonucleotide. Quantitation used previously described methods (1).

Phasing Analysis. Phasing analysis was performed on 577-585 base pair fragments where the intrinsically curved sequence 5' AAAAAACGGCAAAAAACGGGCAAAAAA 3' was linked by a variable GC tract to the target sequence. Equilibrations were performed as in the gel shift assay with an oligonucleotide concentration of 10 μM under similar electrophoretic conditions with a 7% polyacrylamide gel and a 32:1 acrylamide: bis-acrylamide ratio.

Quantitation of Bend Angle. Bend angle quantitation was based on previously described methods (21, 38-40). Briefly, the relationship between the distance migrated and fragment size for molecular weight standards was plotted and fit to a logarithmic function (41). This was used to determine the relative mobilities of the various bent fragments which were plotted and fit to the cosine function $\mu=\mu_{ave}\{A_{ph}/2[\cos(2\pi(S-S_T/P_{ph}))]+1\}$, where μ is the electrophoretic mobility measured in apparent base pairs (R_L), A_{ph} is the phasing amplitude, S is the distance between bend centers, S_T is the out of phase bend center distance, and P_{ph} is the phasing period (21, 40). The bend angle, α_B, was then determined from the geometrically derived phasing function $\tan(k\alpha_B/2)=(A_{ph}/2)/\tan(k\alpha_c/2)$ where α_c is the A tract bend angle (21, 40). A standardization value, k, for electrophoretic conditions was calculated to be 1.20 ± 0.11 using fragments which were retarded only by the inserted A tract (21, 40). The fragment, in addition to the inserted A tract, was found to have an intrinsic curvature of $18^0\pm0.4^0$, which was 2.3 ± 0.1 base pairs out of phase with the target site. The additional intrinsic fragment bend angle in the same plane was subtracted from the obtained angle to yield the bend angle of the structure.

RESULTS AND DISCUSSION

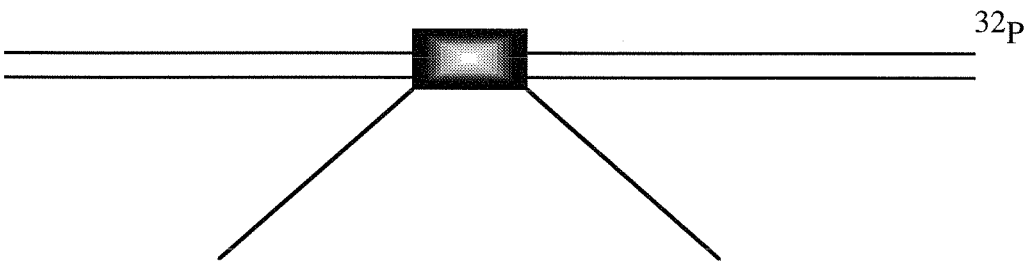
Two 15 base pair purine tracts separated by one turn of the DNA helix (10 bp) were targeted by oligonucleotides **1-6** containing two pyrimidine tracts (T and ^{Me}C) connected by a central T linker of varying size. Previous studies have demonstrated that nonadjacent purine tracts can be bound simultaneously but the influence of oligonucleotide length and sequence composition on bend angle was not examined (42, 43). To allow bending of DNA, the binding energies of the two third strand oligonucleotides must be greater than the energetic penalty for bending. These binding

energies were measured by quantitative DNase I footprint titration analysis, as shown in Figure 2. A sample DNase I footprint is demonstrated in Figure 3 for oligonucleotide **7**, with the resulting averaged isotherms in Figure 4. The equilibrium association constant, K_{eq} of 15mer oligonucleotide **7** targeting site A was measured as K_{eq}=2.2 ($\pm$0.2) x 10^7 M^{-1}, while that for 15mer oligonucleotide **8** targeted to site B was measured as K_{eq}=1.7 ($\pm$0.1) x 10^7 M^{-1}. From these equilibrium association constants, the binding energies of oligonucleotides bound at site A and site B were calculated as -10.0 $\pm$ 0.1 kcal/mole and -9.9 $\pm$ 0.1 kcal/mole, respectively.

DNA bending as a function of linker size was initially screened using a circular permutation assay (44). This assay is based upon the theory that DNA mobility is directly related to the mean square end-to-end distance for fragments of the same size (45). A bending target site was cloned between tandem polylinkers into pUC19. Cleavage with various restriction enzymes gave a series of 101 bp fragments in which the target site is located at various positions. As the bend site is moved through the fragment, the end-to-end distance changes.

The sequence specific binding ligands can bind as one oligonucleotide bound to two sites (bent), one oligonucleotide bound to one site (Y), or as two oligonucleotides bound to two sites (double Y). These binding modes are shown in Figure 5. Control oligonucleotides **9** and **10** contain a mismatch in one of the triplex sites and only bind to one of the two designated target sites. Oligonucleotide **4**, which contains two pyrimidine tracts separated by a central T_4 linker, reveals that it is not bound as a Y or as a double Y (Figure 6). Addition of control oligonucleotides **9** and **10** affords both the Y and double Y structures. Additionally, mobilities of the Y and double Y structures are unchanged between the EcoRI and XbaI fragments shown on both sides of Figure 6, indicating no induced DNA bend.

7. 5' T<u>C</u>T<u>C</u>T<u>CC</u>T<u>CC</u>T<u>C</u>T<u>C</u>T 3'
8. 5' T<u>CC</u>T<u>C</u>T<u>C</u>T<u>C</u>T<u>C</u>T<u>CC</u>T 3'



5' **AGAGAGGAGGAGAGA**GCGGTGCGGT**AGGAGAGAGAGAGGA** 3'
3' **TCTCTCCTCCTCTCT**CGCCACGCCA**TCCTCTCTCTCTCCT** 5'

Figure 2. The oligonucleotides **7** and **8** are targeted against the fragment indicated and bind their designated sites. In the oligonucleotide sequence, <u>C</u> indicates 5-methyl cytosine. Experimental conditions are 45 mM Mes, 1 mM $MgCl_2$, pH 5.5, 22 °C.

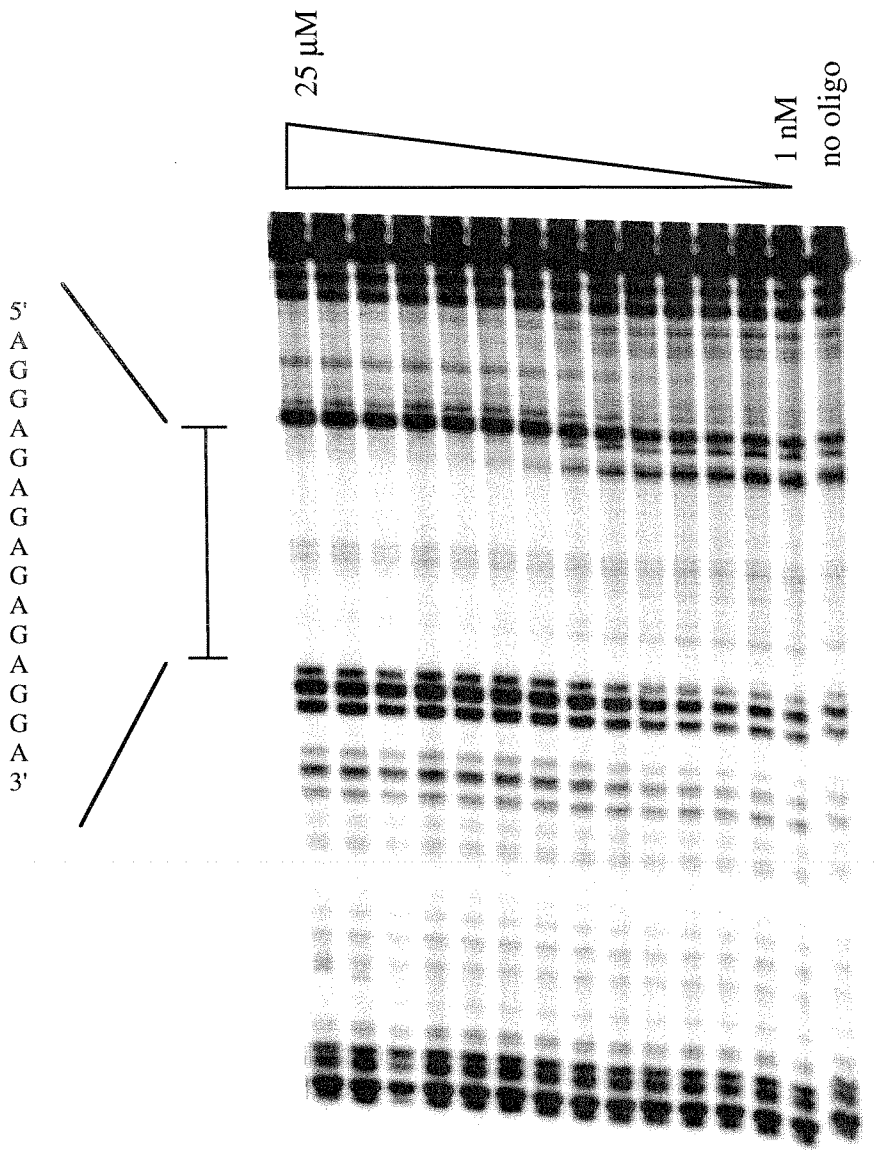


Figure 3. This sample DNase I footprint titration shows oligonucleotide **7** binding to its cognate site. A similar series of gels was run for oligonucleotide **8**.

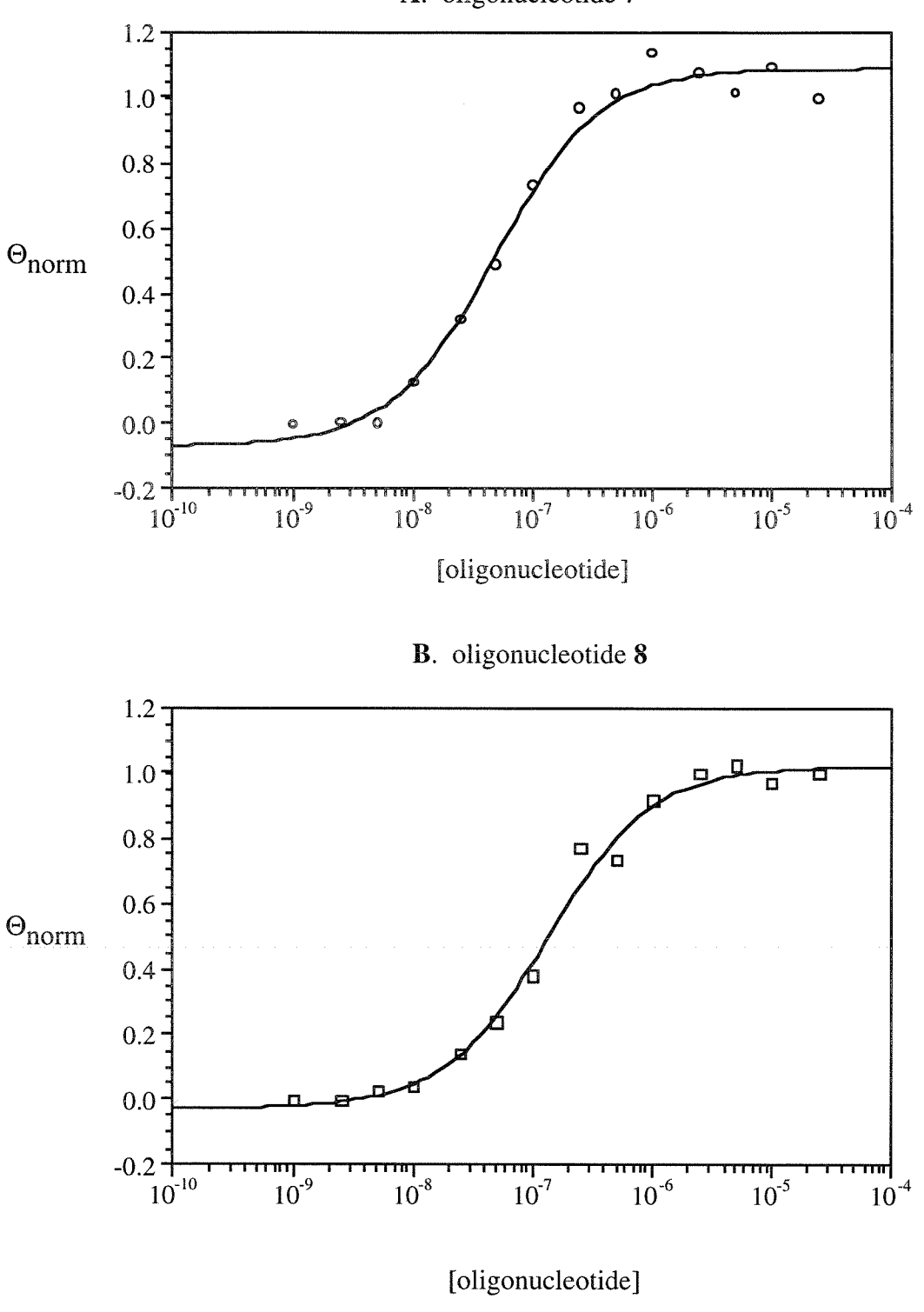


Figure 4. DNase I footprinting titrations yielded these isotherms for oligonucleotides **7** and **8**.

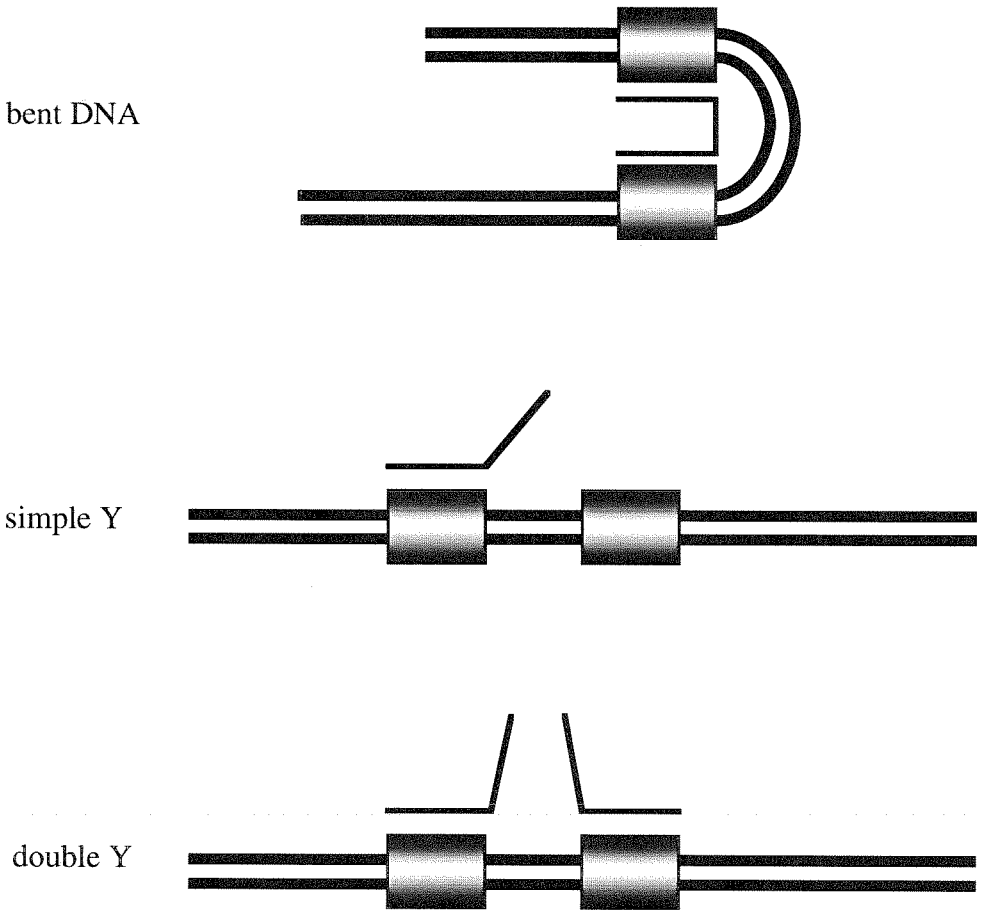


Figure 5. Possible binding modes of bidentate oligonucleotides are depicted.

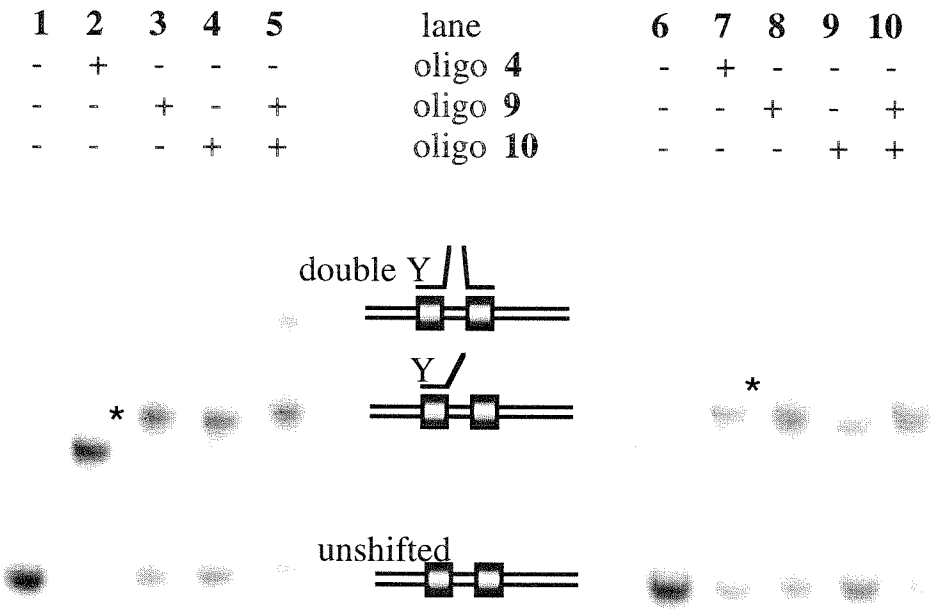


Figure 6. Phosphorimage using a linear dynamic range of 50 of a 10% (75:1 acrylamide:bis ratio) polyacrylamide gel run in 45 mM Mes, 1 mM $MgCl_2$, pH 5.5 at 4° C showing a gel shift analysis of oligonucleotides **4**, **9**, and **10**. The EcoRI and Xba I fragments from Figure 1, with the triple helix target sites near the termini and near the center respectively, were generated and 3' labeled with ^{32}P. Oligonucleotides were incubated at a concentration of 5 μM at 22° C, with 45 mM Mes, 1 mM $MgCl_2$, and pH 5.5. Lanes 1 and 6 show the EcoRI and Xba I fragments respectively incubated without oligonucleotide. Lanes 2 and 7 show the gel shift generated by oligonucleotide **4**, with the bent fragment indicated by an *. Lanes 3 and 8 depict this for oligonucleotide **9**, lanes 4 and 9 for oligonucleotide **10**, and lanes 5 and 10 for an incubation with both oligonucleotides **9** and **10**.

Circular permutation analysis of the bent structures shows significant position-dependent anomalous mobility as a function of linker length. Oligonucleotides with shorter linkers yield larger bend angles (Figure 7). A three T spacer appears to be the minimal linker where a majority of the molecular ensemble is in the bent conformation under these conditions. As the linker size is further shortened, a decreasing proportion of the molecular ensemble reported here is found as the bent structure with a concomitant increase in the populations of Y and double Y species. A range of bend angles of differing magnitude can be produced and the linker size can be tuned to yield the magnitude of the desired bend angle. The circular permutation amplitudes are shown in Figure 8, with the resulting flexure angles listed in Table 1.

The circular permutation assay has been found to overestimate bend angles where other factors such as increased DNA flexibility or binding of a ligand that is large compared with the target duplex affect migration (21, 46). While such complicating effects are not anticipated in this experimental design, a phasing analysis was utilized to measure the bend angles of oligonucleotides **4**, **5**, and **6** (Figure 9a) (38). The A tract used in this experimental design has a bend angle of 54^0 and oligonucleotides **4**, **5**, and **6** were found to provide bend angles of $61^0 \pm 2.7^0$, $50^0 \pm 1.0^0$, and $38^0 \pm 0.5^0$, respectively (Figure 9b,c and Table 1) (47). From these values, the larger bend angles of shorter linkers can be inferred. Phasing analysis also provides information on bend direction. The bend was found to be centered in the middle of the intervening duplex directed towards the minor groove. For these molecules, where no factors other than site-directed bending affect gel mobility, there is a close correlation between circular permutation measured angles and phasing analysis measured angles.

To form the bent structure, the third strand oligonucleotide most likely binds initially at a single site. The DNA then must show sufficient dynamic flexibility to allow binding of the second domain. The final structure obtained is a function of the energetic difference between the binding energy of the third strand and the requisite

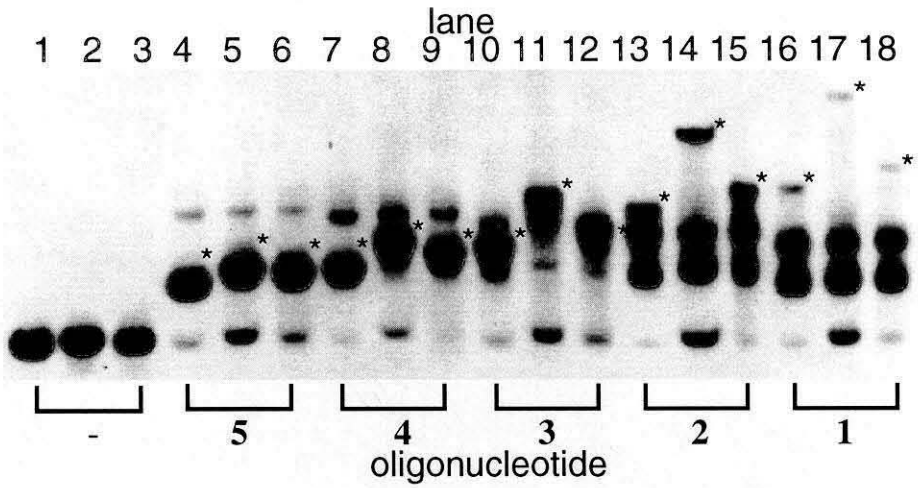


Figure 7. Autoradiogram of a 10% (75:1 acrylamide:bis ratio) polyacrylamide gel run in 45 mM Mes, 1 mM $MgCl_2$, pH 5.5 at 4° C showing circular permutation analysis of bending using oligonucleotides **1-5**. EcoRI, XbaI, and HindIII fragments from Figure 1 were generated and 3' labeled with ^{32}P. Oligonucleotides were incubated at a concentration of 20 μM at 22° C, with 45 mM Mes, 1 mM $MgCl_2$, and pH 5.5. Lanes 1-3 show the EcoRI, XbaI, and HindIII fragments respectively incubated without oligonucleotide. Lanes 4-6 show oligonucleotide **5** incubated with the EcoRI, XbaI, and HindIII fragments. Lanes 7-9 depict oligonucleotide **4** binding to the three fragments respectively, while lanes 10-12 depict this for oligonucleotide **3**, lanes 13-15 for oligonucleotide **2**, and lanes 16-18 for oligonucleotide **1**. Bent fragments are indicated with an *.

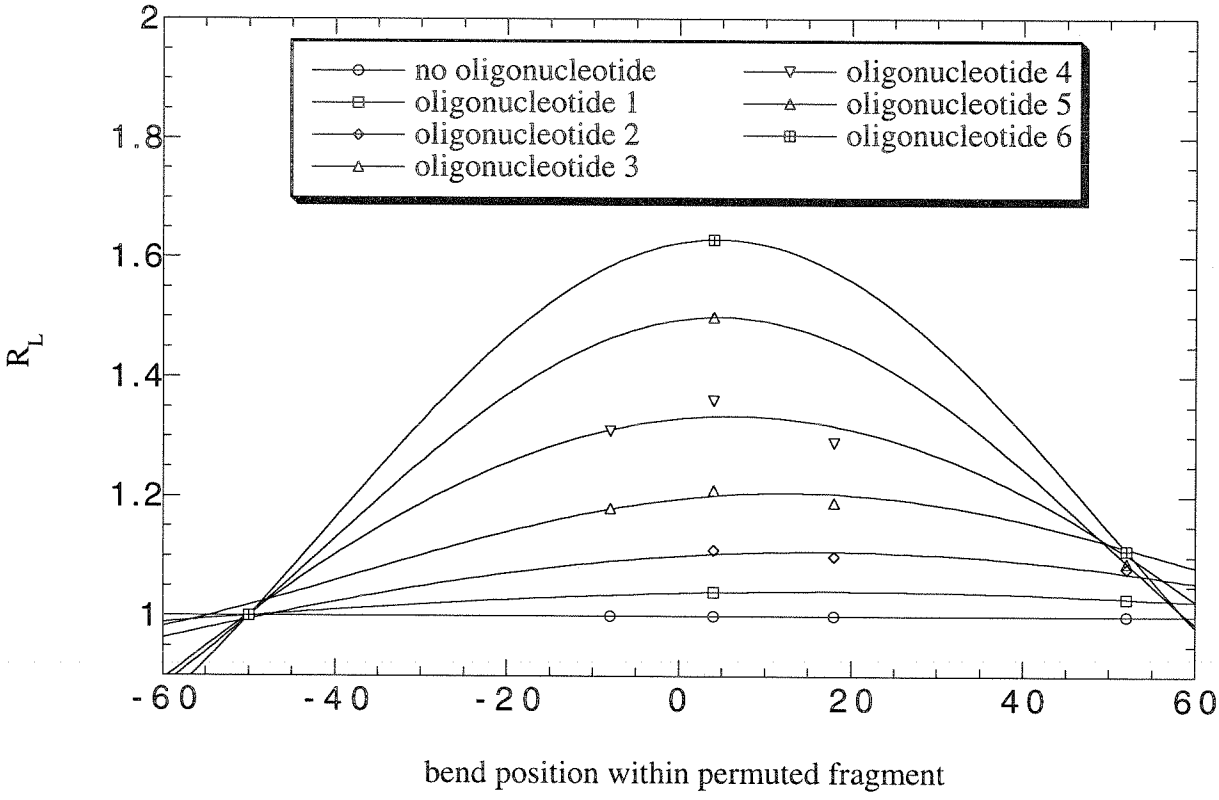


Figure 8. Circular permutation analysis of oligonucleotides (1).

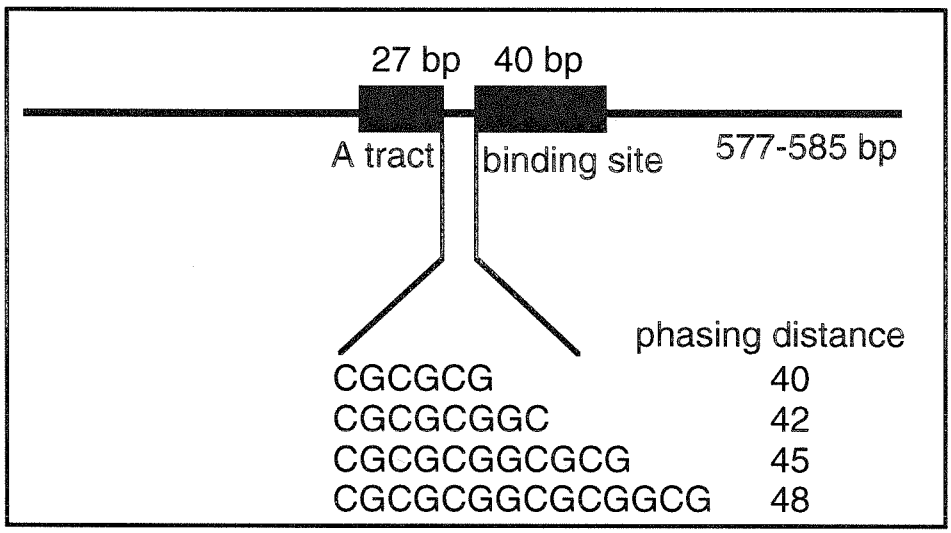


Figure 9a. General schematic for phasing analysis of DNA bending. A set of three phased A_6 tracts with total bend angle of 54° is separated by the phasing distance from a second bend.

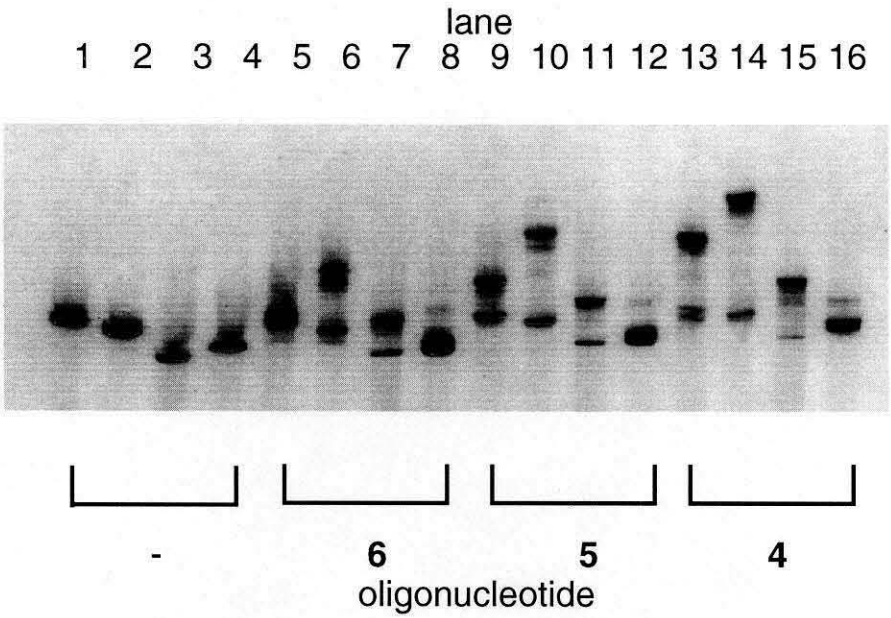


Figure 9b. Autoradiogram of a 7% (32:1 acrylamide:bis ratio) polyacrylamide gel run in 45 mM Mes, 1 mM $MgCl_2$, pH 5.5 at 4° C showing phasing analysis of bending oligonucleotides **6**, **5**, and **4**. The 577-585 base pair fragments generated by simultaneous SapI/NdeI cleavage were 3' labeled with ^{32}P. Oligonucleotides were incubated at a concentration of 10 μM at 22° C, with 45 mM Mes, 1 mM $MgCl_2$, and pH 5.5. Lanes 1-4 depict the four phasing fragments with phasing distances of 40, 42, 45, and 48 respectively with no oligonucleotide present while lanes 5-8 depict the respective fragments with oligonucleotide **6** bound, lanes 9-12 with oligonucleotide **5** bound, and lanes 13-16 with oligonucleotide **4** bound.

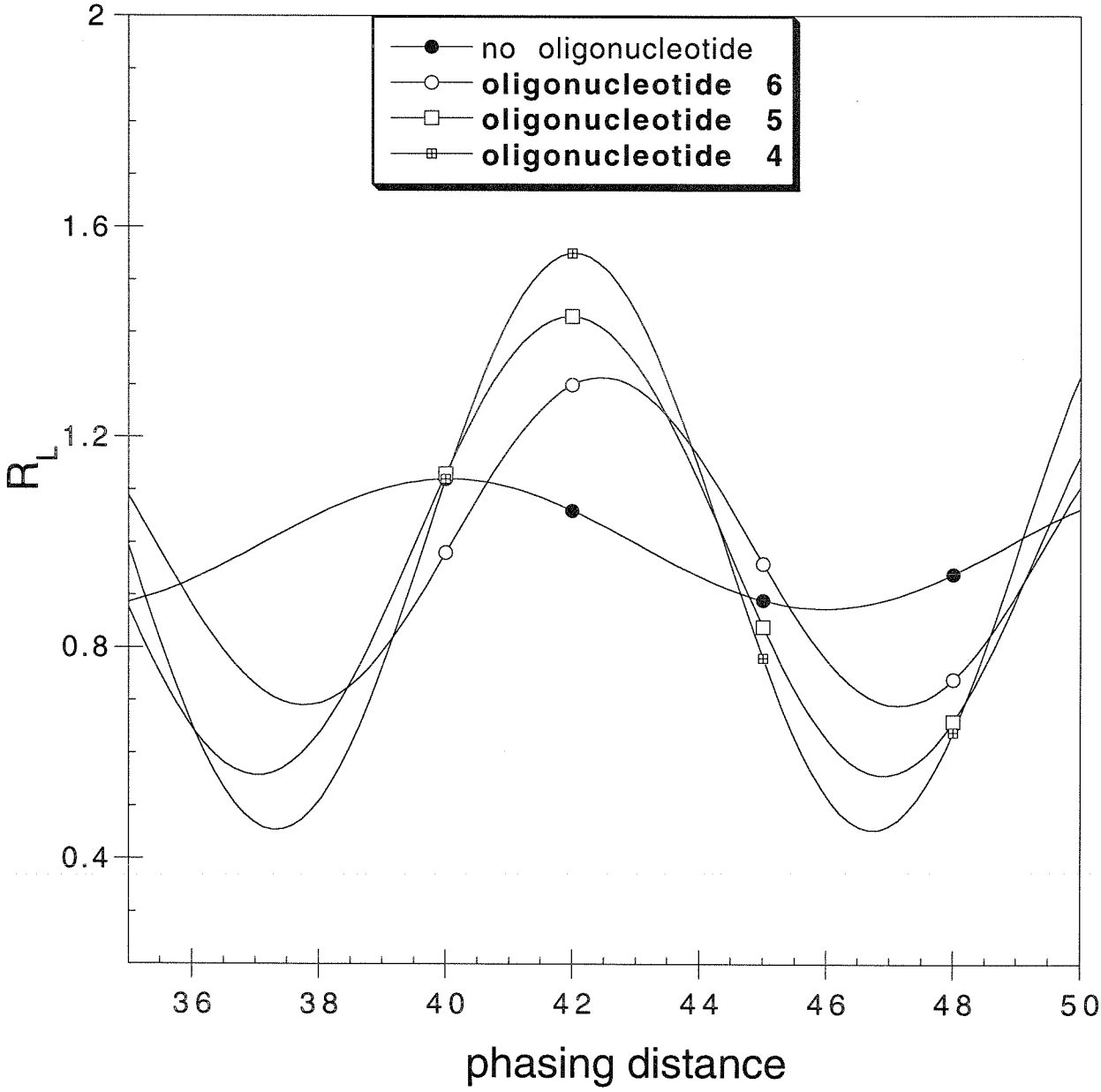


Figure 9c. Phasing diagram for oligonucleotides **6**, **5**, and **4**. The mobilities of these oligonucleotides were plotted against the phasing distance and fit to a standard cosine function (21, 40).

Table 1. Bend angles measured for triple helical structures with variable linker sizes at pH 5.5 with oligonucleotides containing 5-methyl C and T. Error values reported are the standard error of the mean. The error limits on the bend angles do not reflect the semiquantitative nature of the analysis and should be considered minimum values.

OLIGONUCLEOTIDE	LINKER SIZE	BEND ANGLE (CP)	BEND ANGLE (PHASING)
1	1	$>105^{o}$	--
2	2	$>101^{o}$	--
3	3	$>83^{o}$	--
4	4	$65^{o}\pm2^{o}$	$61^{0}\pm2.7^{0}$
5	5	$49^{o}\pm2^{o}$	$50^{0}\pm1.0^{0}$
6	6	$<30^{o}$	$38^{0}\pm0.5^{0}$

bending energy caused by constriction of the DNA. The thermodynamically favored structure occurs where this difference is greatest, with other structures populated dependent upon their relative energetic differences. The energetic cost of bending DNA can be estimated by modeling DNA as a smoothly bending wormlike chain with coulombic repulsion from phosphates placed at fixed distances (48, 49). Assuming the third strand is fully hydrogen bonded and ignoring cooperative effects, this model predicts a bend angle of 86^0 to promote equal ratios of bent and straight DNA for an oligonucleotide with a calculated binding energy of -9.9 kcal/mole. Circular permutation analysis reveals a single bent structure which decreases in population as the linker length decreases but remains populated at a length of a single T residue (Figure 7). For a linker of three T residues, the ratio of bent to straight DNA is approximately one. The bend angle of this structure was too high to measure accurately using phasing analysis and a four T linker was the smallest that was characterized quantitatively, with a bend angle of 61^0. A ribbon model that depicts a bent structure with a four base linker is shown in Figure 10. This model is bent towards the minor groove at the center of the intervening duplex, with the four T linker of the third strand pulled taut away from the duplex and unstacked. A more thorough analysis of the nature and energetics of the structure formed is presented in Chapter 4.

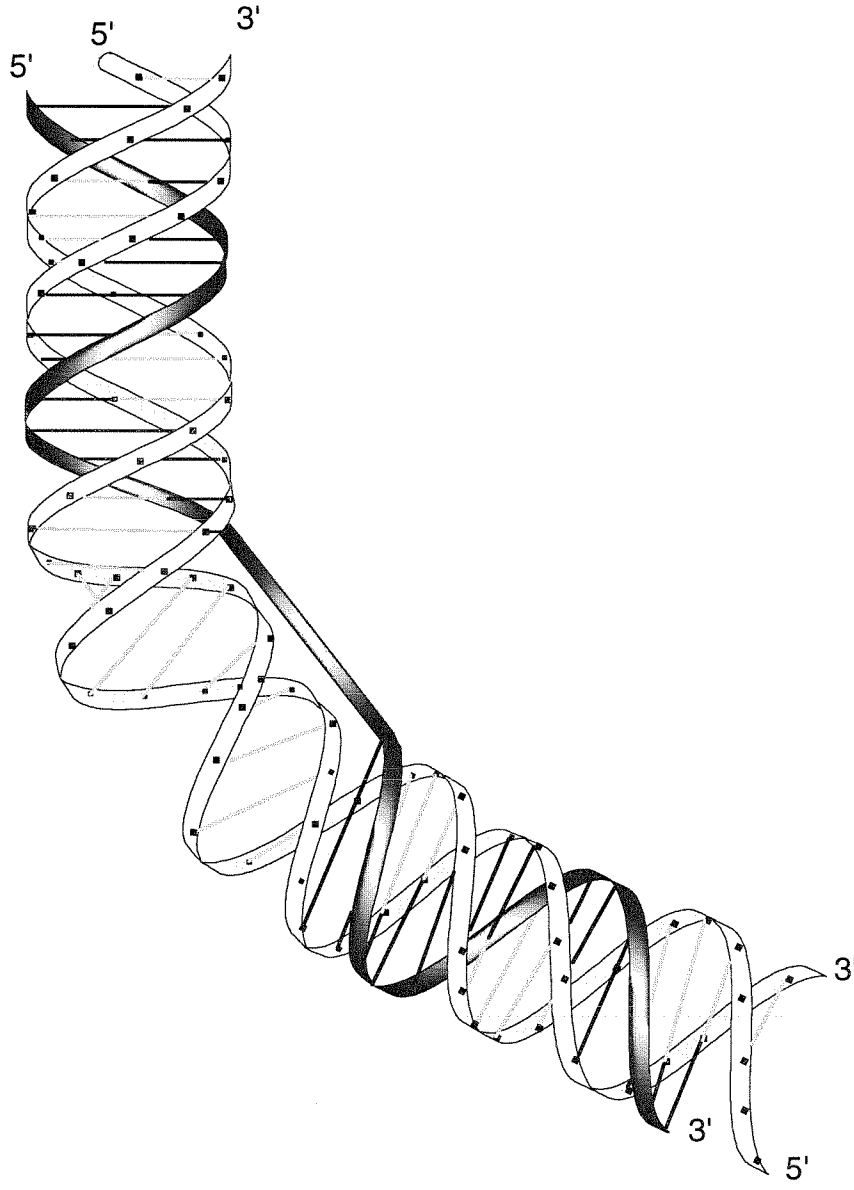


Figure 10. Ribbon model of bent structure formed by oligonucleotide **4**. Two 15 base pair triple helices are connected by a taught four T linker spanning the intervening 10 base pairs of double helix. This double helical region is bent towards the minor groove in the center.

REFERENCES

1. Thompson, J. F. & Landy, A. (1988) *Nucleic Acids Research* **16,** 9687-9705.
2. Goodman, S. D., Nicholson, S. C. & Nash, H. A. (1992) *Proc. Natl. Acad. Sci., USA* **89,** 11910-11914.
3. Hallet, B., Rezsohazy, R., Mahillon, J. & Delcour, J. (1994) *Mol. Microbiol.* **14,** 131-139.
4. Goryshin, I. Y., Kil, Y. V. & Reznikoff, W. S. (1994) *Proc. Natl. Acad. Sci., USA* **91,** 10834-10838.
5. Henriquez, V., Milisavljevic, V., Kahn, J. D. & Gennaro, M. L. (1993) *Gene* **134,** 93-98.
6. Nakajima, M., Sheikh, Q. I., Yamaoka, K., Yui, Y., Kajiwara, S. & Shishido, K. (1993) *Mol. Gen. Genet.* **237,** 1-9.
7. Muller, H. P. & Varmus, H. E. (1994) *EMBO J.* **13,** 4704-4714.
8. Milot, E., Belmaaza, A., Rassart, E. & Chartrand, P. (1994) *Virology* **201,** 408-412.
9. Natesan, S. & Gilman, M. Z. (1993) *Genes Dev* **7,** 2497-2509.
10. Meacock, S., Pescini-Gobert, R., DeLamarter, J. F. & vanHuijsduijen, R. H. (1994) *J. Biological Chemistry* **269,** 31756-31762.
11. Kim, J., Klooster, S. & Shapiro, D. J. (1995) *J. Biological Chemistry* **270,** 1282-1288.
12. Perez-Martin, J., Rojo, F. & DeLorenzo, V. (1994) *Microbiological Reviews* **58,** 268-290.
13. Kim, J. L., Nikolov, D. B. & Burley, S. K. (1993) *Nature* **365,** 520-527.
14. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. (1993) *Nature* **365,** 512-520.

15. Haqq, C. M., King, C., Ukiyama, E., Falsafi, S., Haqq, T. N., Donahoe, P. K. & Weiss, M. A. (1994) *Science* **266,** 1494-1500.

16. Mondragon, A. & Harrison, S. C. (1991) *J. Molecular Biology* **219,** 321-334.

17. Fisher, R. F. & Long, S. R. (1993) *J. Molecular Biology* **233,** 336-348.

18. Finzi, L. & Gelles, J. (1995) *Science* **267,** 378-380.

19. Manning, G. S., Ebralidse, K. K., Mirzabekov, A. D. & Rich, A. (1989) *J. Biomol. Struct. Dyn.* **6,** 877-889.

20. Strauss, J. K. & Maher, L. J. (1994) *Science* **266,** 1829-1834.

21. Kerppola, T. K. & Curran, T. (1991) *Science* **256,** 1210-1214.

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

23. LeDoan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J., Thuong, N. T., Lhomme, J. & Helene, C. (1987) *Nucleic Acids Research* **15,** 7749-7760.

24. Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988) *Science* **241,** 456-459.

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

26. Beal, P. A. & Devan, P. B. (1992) *Nucleic Acids Research* **20,** 2773-2776.

27. Singleton, S. F. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 6957-6965.

28. Roberts, R. W. & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci., USA* **88,** 9397-9401.

29. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannageli, C., Garestier, T. & Helene, C. (1992) *Biochemistry* **31,** 9269-9278.

30. Fosella, J. A., Kim, Y. J., Shih, H., Richards, E. G. & Fresco, J. R. (1993) *Nucleic Acids Research* **21,** 4511-4515.

31. Best, G. C. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 1187-1193.

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

33. Xodo, L. E., Manzini, G., Quadrifoglio, F., vanderMarel, G. A. & vanBoom, J. H. (1991) *Nucleic Acids Research* **19,** 5625-5631.

34. Singleton, S. F. & Dervan, P. B. (1992) *Biochemistry* **31,** 10995-11003.

35. Singleton, S. F. & Dervan, P. B. (1993) *Biochemistry* **32,** 13171-13179.

36. Singleton, S. F. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 10376-10382.

37. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986) *Proc. Natl. Acad. Sci., USA* **83,** 8462-8466.

38. Zinkel, S. S. & Crothers, D. M. (1987) *Nature* **328,** 178-181.

39. Zinkel, S. S. & Crothers, D. M. (1990) *Biopolymers* **29,** 29-38.

40. Kerppola, T. & Curran, T. (1993) *Mol. Cell. Biol.* **13,** 5479-5489.

41. Maniatis, T., Jeffrey, A. & vandeSande, H. (1975) *Biochemistry* **14,** 3787-3794.

42. Kessler, D. J., Pettitt, B. M., Cheng, Y., Smith, S. R., Jayaraman, K., Vu, H. M. & Hogan, M. E. (1993) *Nucleic Acids Research* **21,** 4810-4815.

43. Kiyama, R. & Oishi, M. (1995) *Nucleic Acids Research* **23,** 452-458.

44. Wu, H. & Crothers, D. M. (1984) *Nature* **308,** 509-513.

45. Zimm, B. H. & Levene, S. D. (1992) *Q. Rev. Biophys.* **25,** 171-204.

46. Kuprash, D. V., Rice, N. R. & Nedospasov, S. A. (1995) *Nucleic Acids Research* **23,** 427-433.

47. Koo, H., Drak, J., Rice, J. A. & Crothers, D. M. (1990) *Biochemistry* **29,** 4227-4234.

48. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. (1974) in *Physical Chemistry of Nucleic Acids* (Harper and Row, New York), pp. 159-166.

49. Fenley, M. O., Manning, G. S. & Olson, W. K. (1992) *J. Physical Chemistry* **96,** 3963-3969.

CHAPTER 4

A BIOPHYSICAL CHARACTERIZATION OF DNA BENDING THIRD STRAND OLIGONUCLEOTIDES

DNA bending plays a crucial role in many biological processes including gene expression. The design of nonnatural sequence specific DNA bending ligands may lead to the development of artificial regulators of these biological processes. Previously, the design of a sequence specific DNA bending ligand has been reported (1). In this study, two 15 base pair pyrimidine motif domains linked through a variable length poly-dT region bind to nonadjacent double helical sites separated by 10 base pairs. This system is complicated from an energetic standpoint and studies were undertaken to better understand third strand mediated DNA bending from a biophysical standpoint.

The binding of a third strand oligonucleotide in a triple helical complex is dependent upon many factors, including binding motif, third strand length, sequence composition, pH, and ion concentration and valence (2-17). Additionally, the composition and orientation of the linker may be important, as well as the orientation of the target site. Beyond these variables, an ensemble of bending ligand binding motifs is possible and observed dependent upon the conditions utilized. The system may show a cooperative nature to the binding of two linked independent binding domains, as has been observed previously for linked third strand oligonucleotides, adding an additional layer to the complexity of the binding energetics (18).

The studies presented here include a measurement of the binding constants for oligonucleotides with different length linkers, studies on the effects of third strand base composition, binding motif, and length, target site spacing, buffer pH and salt

concentration, and examination of the effect of orientation, electrostatics and hydrophobicity (stacking) in the linker, as well as other experiments.

MATERIALS AND METHODS

General. Solvents and reagents for synthesis were purchased from VWR. NMR spectra were obtained on a GE 300 instrument operating at 300 MHz. Chemical shifts are reported in parts per million relative to the solvent residual signal. Phosphoramidites were purchased from Glen Research and Biogenex. DNA synthesis was performed using standard methods on an Applied Biosystems 380B DNA Synthesizer. Oligonucleotide purification was performed on a Pharmacia FPLC using ProRPC reverse phase and mono Q anion exchange columns. Enzymes were purchased from Boehringer Mannheim, New England Biolabs, and United States Biochemical. Radiolabeled nucleotides were purchased from Amersham. Storage Phosphor autoradiography was performed using a Molecular Dynamics 400S Phosphorimager and ImageQuant software. Molecular modeling was performed on a Silicon Graphics work station.

SYNTHESIS

***N,N'*-di-(trifluoroacetyl)-N-(ethyl-2-O-dimethoxytrityl)-N'-2-hydroxyethyl-ethylenediamine (2).** *N,N'*-Bis(2-hydroxyethyl) ethylene diamine (1 gram) was dissolved in pyridine (20 ml) to which an excess of trifluroacetic anhydride (2.5 ml) was added through an addition funnel with constant stirring at 4° C under Ar. This mixture was stirred for 3 hours, then, 0.8 equivalents of dimethoxytrityl chloride (1.83 grams) dissolved in pyridine were added slowly through the addition funnel. The mixture was stirred for an additional 4 hours, still at 4° C.

The reaction was monitored by TLC (100% dichloromethane) and extracted with 1N sodium hydroxide and ether. The ether layer was dried ($MgSO_4$), filtered, and concentrated. The crude product was purified by flash column chromatography (3:1 hexanes:ethyl acetate with 0.2% triethylamine). The pure compound **2** was obtained in 51% yield. ^{1}H NMR (300 MHz, CD_3CN) δ 7.20-7.50 (m, 9H), 6.88 (m, 4H), 3.88 (m, 2H), 3.74 (s, 6H), 3.55 (m, 2H), 3.35 (m, 2H); HRMS (FAB) calculated for $C_{31}H_{32}N_2F_6O_6$ 642.2165, found 642.2195.

***N,N'*-di-(trifluoroacetyl)-N'-(ethyl-2-O-dimethoxytrityl)-ethylenediamino-N-ethyl-2-O-(*N''*,*N''*-diisopropyl-2'-cyanoethyl)-phosphoramidite (3).** Compound **2** (658.8 mg) was dissolved in tetrahydrofuran (6 ml) at 4° C and under Ar. To this was added an excess of 2-cyanoethyl *N,N*-diisopropyl chlorophosphoramidite (223 μl) and a stoichiometric amount of diisopropylethylamine (523 μl). This solution was stirred for 3 hours and checked by TLC (1:1 hexanes:ethyl acetate) for completion of the reaction. The material was purified by column chromatography (4:1 hexanes:ethyl acetate with 0.2% triethylamine) and concentrated. The purified compound **3** was dried resulting in a 65% yield. This material coupled under standard DNA synthesis conditions with an efficiency of 86%. ^{1}H NMR (300 MHz, CD_3CN) δ 7.26-7.52 (m, 9H), 6.89 (m, 4H), 3.60-3.90 (m, 12H), 3.35 (m, 2H), 2.64 (m, 6H), 1.57-1.63 (m, 12H). $C_{40}H_{49}N_4F_6PO_7$.

Deprotection of methyl phosphonate containing oligonucleotides. Methyl phosphonate containing oligonucleotides were deprotected using a previously reported procedure (19). Briefly, oligonucleotides were deprotected in ethanol:acetonitrile:ammonium hydroxide 45:45:10 for 30 minutes at room temperature.

This was followed by addition of an equal volume of ethylene diamine. The solution was reacted for six hours at room temperature. After workup, oligonucleotides were purified using a Pharmacia FPLC using standard reverse phase conditions.

Mass spectrometry (MALDI-TOF) of oligonucleotides. Oligonucleotides were confirmed by MALDI-TOF mass spectrometry.

oligonucleotide **15** calculated 9588 observed 9587.

oligonucleotide **17** calculated 9960 observed 9959.

oligonucleotide **18** calculated 9260 observed 9255.

oligonucleotide **19** calculated 9258 observed 9259.

Preparation of labeled DNA. Plasmid DNA was digested with the appropriate restriction endonucleases and simultaneously labeled with Sequenase 2.0, the appropriate deoxynucleoside-5'-[α-^{32}P]triphosphates, and nonradioactive deoxynucleoside triphosphates. The fragment was purified by gel electrophoresis, treated with proteinase K, filtered, further extracted with phenol/chloroform, and precipitated with ethanol.

DNase I footprint reactions. All reactions were equilibrated at 22 °C, pH 5.5 in the presence of 10 mM Mes, 1 mM $MgCl_2$, and labeled DNA for at least 24 hours. Footprinting reactions were carried out as previously described (20).

MPE footprint reactions. All reactions were equilibrated at 22 °C, pH 5.5 in the presence of 10 mM Mes, 1 mM $MgCl_2$, and labeled DNA for at least 24 hours. Footprinting reactions were carried out as previously described (21).

Gel shift assay and titration. All samples were equilibrated at 22 °C, pH 5.5 in the presence of 10 mM Mes, 1 mM $MgCl_2$, and labeled DNA for at least 12 hours, with oligonucleotide titrations equilibrated for at least 36 hours. Oligonucleotides concentrations are 10 μM, unless otherwise indicated in the figure legend. Oligonucleotide titrations were performed in the presence of 100 μM (in base pairs) of calf thymus DNA in addition to the above conditions. One tenth volume of 15% glycerol loading buffer was added and samples were run on a 10% polyacrylamide gel at 4 °C, pH 5.5, with a 75:1 acrylamide/bis-acrylamide ratio in 45 mM Mes, 1 mM $MgCl_2$ with buffer recirculation. Data were collected by storage phosphor densitometry (Molecular Dynamics 400S Phosphorimager). Quantitation of isotherms was performed by plotting the ligand concentration against the portion of labeled DNA in the bent conformation against the portion in the bent conformation and curve fit using a Langmuir binding isotherm, as previously described (22).

Molecular modeling. The bent structure formed by the oligonucleotide with two T linker residues was modeled. The bent duplex region was built using NAMOT (23). The complete structure was built and minimized using Biograf (Molecular Simulations, Inc.). The AMBER force field was used as described (24, 25). The electrostatic potential for 5-methyl cytosine was adapted from (26). Sodium counterions were adapted from (27). Minimizations were performed with an entirely free third strand including the bound region except for the last residues which were constrained. The bent region was modeled as smoothly bending and the base pair planes were maintained while the backbone conformation was minimized. Molecules were subjected to 20 ps rounds of dynamics followed by minimization, where the lowest energy structures were taken.

RESULTS AND DISCUSSION

DNA bending third strand oligonucleotides have many possible binding modes available. In trying to understand the behavior of bending ligands, an energetic assessment of the factors involved in bent structure formation is necessary. A good starting point for this biophysical evaluation is the one provided by a theoretical framework.

THEORETICAL ANALYSIS OF DNA BENDING ENERGETICS

Short pieces of DNA are relatively inflexible and can be viewed as rigid rods. As DNA elongates, it shows flexibility and bending. To approximate this, short rods are attached to each other with a random linkage angle to form a wormlike chain. For curved DNA, as the rod length is decreased, the wormlike chain assumes a smoothly bending characteristic and negative charges can be placed at fixed distances to represent phosphates. The negative charge of phosphates gives the DNA stiffness from an electrostatic standpoint. Other structural variables also contribute to the stiffness of DNA, which can be viewed as a springlike material with a spring constant or elasticity corresponding to its stiffness. This stiffness is proportional to the persistence length, which can be measured experimentally (28, 29). Modeling of DNA yields a relationship shown below,

$$\Delta G=(RTp/2l)\Theta^2$$

where p is the persistence length, about 150 base pairs for random sequence DNA, l is the contour length, and Θ is the bend angle. A derivation of this equation is presented in (28).

For bending mediated by a bidentate third strand oligonucleotide, this theoretical expression can be utilized to predict the penalty accrued by distorting DNA. This free energy value can be combined with the binding expression of a second site oligonucleotide assuming saturation of the first site. This is probably not a fully accurate model of what occurs in solution, given the thermodynamic equivalence of the two sites and the presumed need for stretching of the third strand to bind both sites followed by zipping of the third strand domains to form the final structure. Essentially the model ignores both distortion effects on the third strand binding energy and cooperative effects between the two third strand domains as well as kinetic trapping effects in double Y formation, which are expected to be concentration dependent. However, this model is a good rough starting point to view the binding event quantitatively. The resulting equation is presented below,

$$\Theta^2 = (RT\ln([Y]/[B]) + \Delta G_{site\ A})(2l/RTp)$$

where [B] and [Y] represent the proportions found in bent and Y conformations for a given concentration of third strand oligonucleotide and $\Delta G_{site\ A}$ is the measured monomer binding energy at the first site, -10 kcal/mole. The derived relationship is shown graphically in Figure 1.

With this theoretical starting point, it is expected that DNA bending to large angles should be readily accommodated and that was in fact seen experimentally. This model predicts a 1:1 ratio of bent:straight DNA at 86^0, which is the bend angle measured by circular permutation analysis for a three T linked oligonucleotide, where a

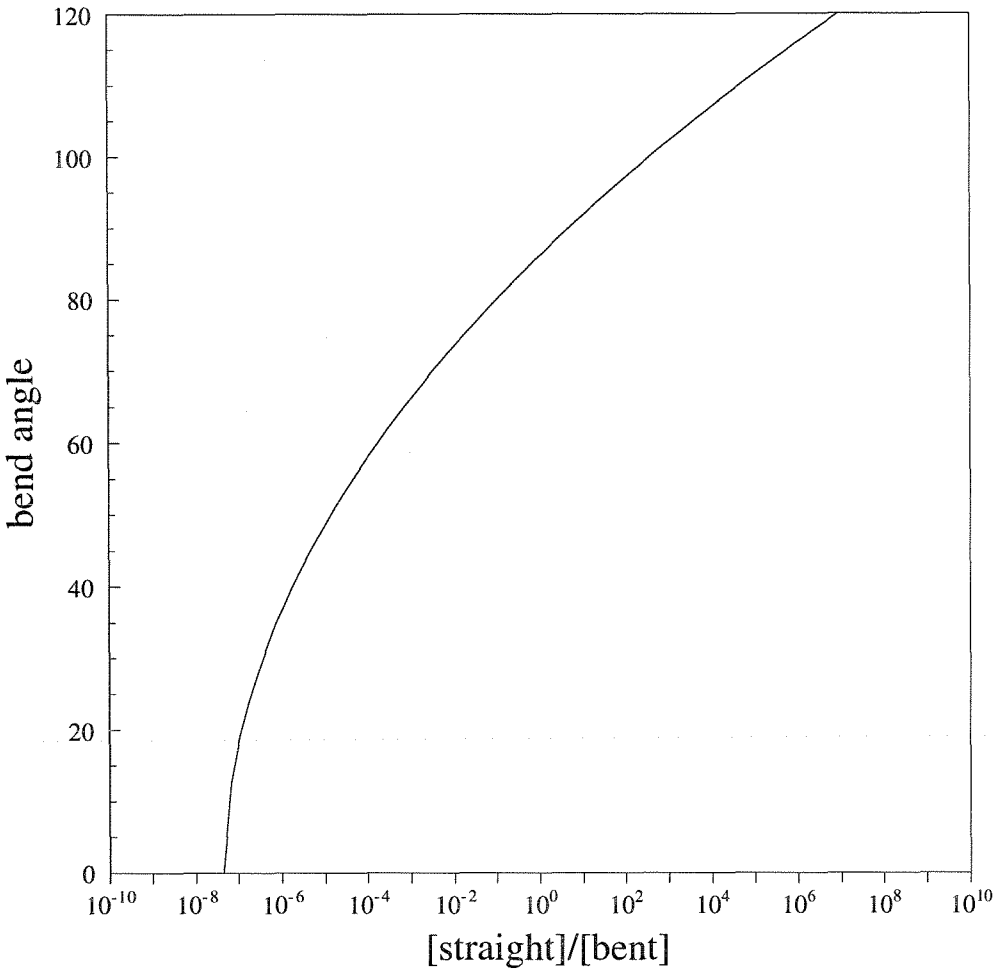


Figure 1. The relationship between bend angle and the proportion of the molecular ensemble found in the bent conformation is depicted as predicted from a theoretical model (28).

roughly 1:1 ratio of bent:straight DNA is obtained. However, this model breaks down in two ways when compared to the initial data from Figure 7 of Chapter 3. First, the ratio of bent:straight DNA is concentration dependent, indicating that kinetic trapping does play a role. Secondly and more importantly, the ratio of bent:straight in all concentrations examined showed a much flatter relationship between this ratio and bend angle, indicating both a significant degree of cooperativity and a smaller energy required for bending towards the minor groove. Given the problems with the presented theoretical model, we sought to examine the energetics of bending using the experimental system described in the previous chapter.

EXPERIMENTAL DETERMINATION OF BENDING OLIGONUCLEOTIDE BINDING AFFINITIES

Using gel shift titration analysis, a series of binding affinities for oligonucleotides with different linkers was determined. Experimental confirmation of the affinity of the unlinked monomer bound to site A, which was previously measured using quantitative DNase I footprint titration analysis, was also obtained by gel shift titration analysis. This experimental design with the sequence is depicted in Figure 2. Sample titrations depicting the binding of oligonucleotides **6-9**, **6-6**, **6-5**, **6-4**, **6-2**, and **1** are shown as Figures 3-8 (oligonucleotide sequences are shown in Table 1). The final data obtained are presented in Table 2 and shown graphically in Figure 9.

The relationship between linker length and binding affinity leads to several immediate conclusions. The effects of cooperativity are intermingled with the bending penalty and the curve shape is not what was predicted theoretically, where shorter linkers bear an energetic cost. The oligonucleotide with the highest affinity induces a measured bend angle of 61^0. Here, the ratio of bent:Y DNA seen on the circular permutation gel is close to minimal. Thus cooperative effects are much stronger than

6-9 5' TCTCTCCTCCTCTCT**TTTTTTTTT**TCCTCTCTCTCTCCT 3'
6-6 5' TCTCTCCTCCTCTCT**TTTTTT**TCCTCTCTCTCTCCT 3'
6-5 5' TCTCTCCTCCTCTCT**TTTTT**TCCTCTCTCTCTCCT 3'
6-4 5' TCTCTCCTCCTCTCT**TTTT**TCCTCTCTCTCTCCT 3'
6-2 5' TCTCTCCTCCTCTCT**TT**TCCTCTCTCTCTCCT 3'
1 5' TCTCTCCTCCTCTCT 3'

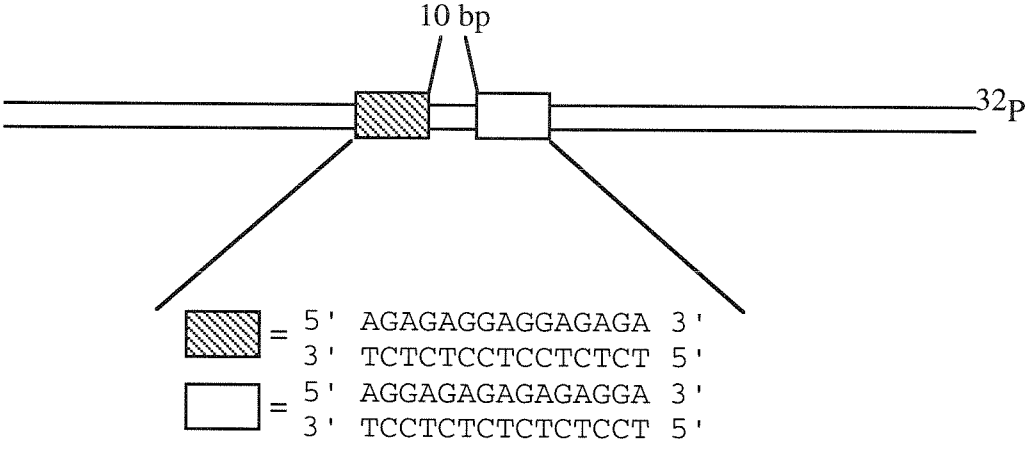


Figure 2. The oligonucleotides presented are targeted against the restriction fragment indicated by gel shift titration analysis. In the oligonucleotide sequence, C indicates 5-methyl cytosine.

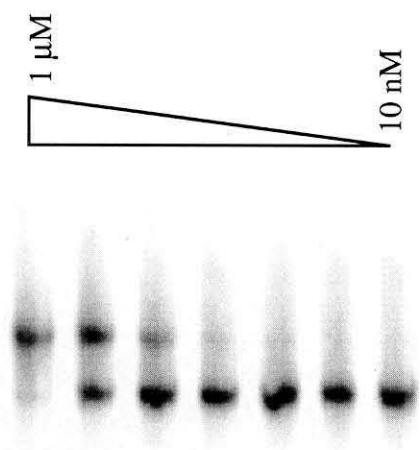


Figure 3. A phosphorimager generated image (linear dynamic range of 50) of a gel shift titration is shown for oligonucleotide **6-9**. The concentration range on this gel is from 10 nM to 1 μM.

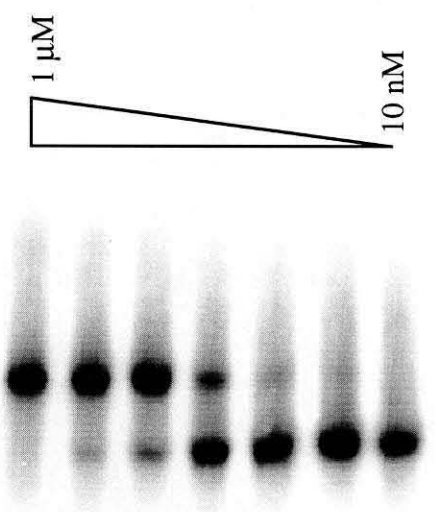


Figure 4. A phosphorimager generated image (linear dynamic range of 50) of a gel shift titration is shown for oligonucleotide **6-6**. The concentration range on this gel is from 5 nM to 500 nM.

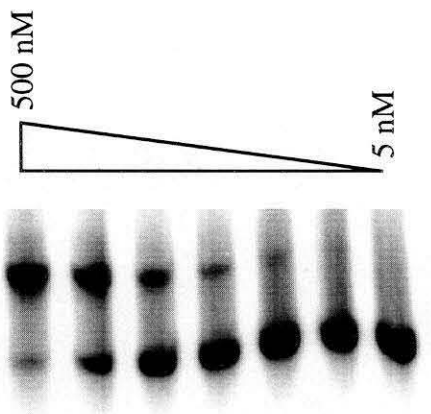


Figure 5. A phosphorimager generated image (linear dynamic range of 50) of a gel shift titration is shown for oligonucleotide **6-5**. The concentration range on this gel is from 10 nM to 1 μM.

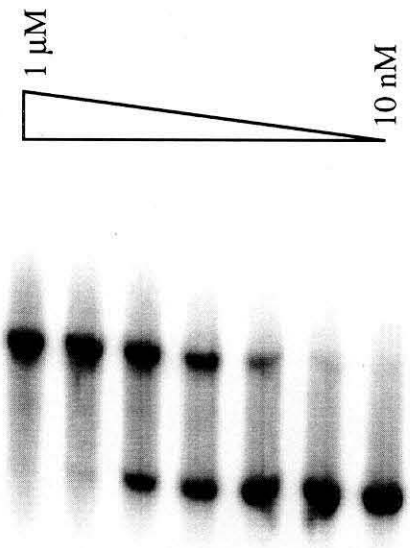


Figure 6. A phosphorimager generated image (linear dynamic range of 50) of a gel shift titration is shown for oligonucleotide **6-4**. The concentration range on this gel is from 10 nM to 1 μM.

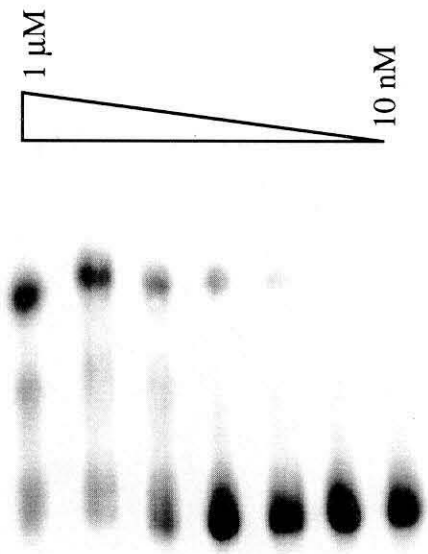


Figure 7. A phosphorimager generated image (linear dynamic range of 50) of a gel shift titration is shown for oligonucleotide **6-2**. The concentration range on this gel is from 10 nM to 1 μM.

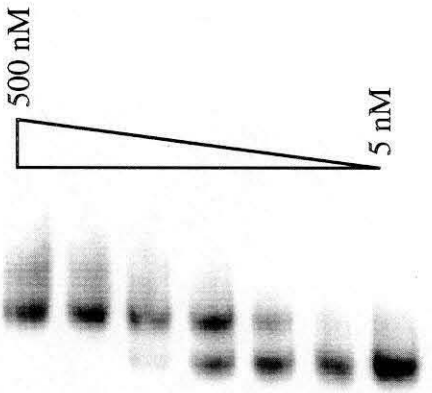


Figure 8. A phosphorimager generated image (linear dynamic range of 50) of a gel shift titration is shown for oligonucleotide **1**. The concentration range on this gel is from 5 nM to 500 nM.

Table 1. This table indicates the oligonucleotides and buffers used for the experiments in this chapter. Here, U is 5-propynyl U, C is 5-methyl C, ϕ is abasic, ^{Me}pT is a methyl phosphonate linkage, EG is polyethylene glycol, and PA is polyamine.

ODN #	OLIGONUCLEOTIDE SEQUENCE
1	TCTCTCCTCCTCTCT
2	TCCTCTCTCTCTCCT
3	UCUUCUUUCUCUT
4	UUCUUUCUCUT
5	TCTTCTTTCTCTT
6	TCTCTCCTCCTCTCT$(T)_{0\text{-}9}$TCCTCTCTCTCTCCT
7	TCCTCTCTCTCTCCT$(T)_{5,10,15}$TCTCTCCTCCTCTCT
8	TCTTCTTTCTCTT$(T)_{3,5}$TCTTCTTTCTCTT
9	TTCTTTCTCTT$(T)_{3,5}$TTCTTTCTCTT
10	UCUCUCCUCCUCUCU$(T)_{3,5}$UCCUCUCUCUCUCCT
11	UCUUCUUUCUCUU$(T)_{3,5}$UCUUCUUUCUCUT
12	UUCUUUCUCUU$(T)_{3,5}$UUCUUUCUCUT
13	TGGTGTGTGTGTGGT$(T)_3$TGTGTGGTGGTGTGT
14	TGGGTGGGGTGGGGT$(T)_3$TGGGTGGGGTGGGGT
15	TCTCTCCTCCTCTCT$(\phi)_3$TCCTCTCTCTCTCCT
16	TCTCTCCTCCTCTCT$(G)_3$TCCTCTCTCTCTCCT
17	TCTCTCCTCCTCTCT$(^{Me}pT)_3$TCCTCTCTCTCTCCT
18	TCTCTCCTCCTCTCT$(EG)_2$TCCTCTCTCTCTCCT
19	TCTCTCCTCCTCTCT$(PA)_2$TCCTCTCTCTCTCCT

Buffers:

A-45 mM Mes, 1 mM $MgCl_2$, pH 5.5
B-45 mM Bis-Tris, 1 mM $MgCl_2$, pH 7.1
C-50 mM Bis-Tris, 140 mM KCl, 10 mM NaCl, 1 mM $MgCl_2$, 1 mM Spermine, pH 7.1
D-45 mM Bis-Tris, 100 mM NaCl, 10 mM $MgCl_2$, pH 7.1
E-10 mM Bis-Tris, 10 mM NaCl, 10 mM $MgCl_2$, pH 7.1
F-45 mM Mes, 1 mM $MnCl_2$, pH 5.5

Table 2. A summary of the measured binding affinities for bending oligonucleotides is reported. Standard error of the mean from four gels is reported.

OLIGONUCLEOTIDE	GEL SHIFT AFFINITY (M^{-1})	FOOTPRINTING AFFINITY (M^{-1})
6-9	2.4 ($\pm$0.3) x 10^6	
6-6	5.7 ($\pm$0.7) x 10^6	
6-5	6.0 ($\pm$1.3) x 10^6	
6-4	7.0 ($\pm$0.9) x 10^6	
6-2	4.0 ($\pm$0.7) x 10^6	
1	1.4 ($\pm$0.7) x 10^7	2.2 ($\pm$0.1) x 10^7

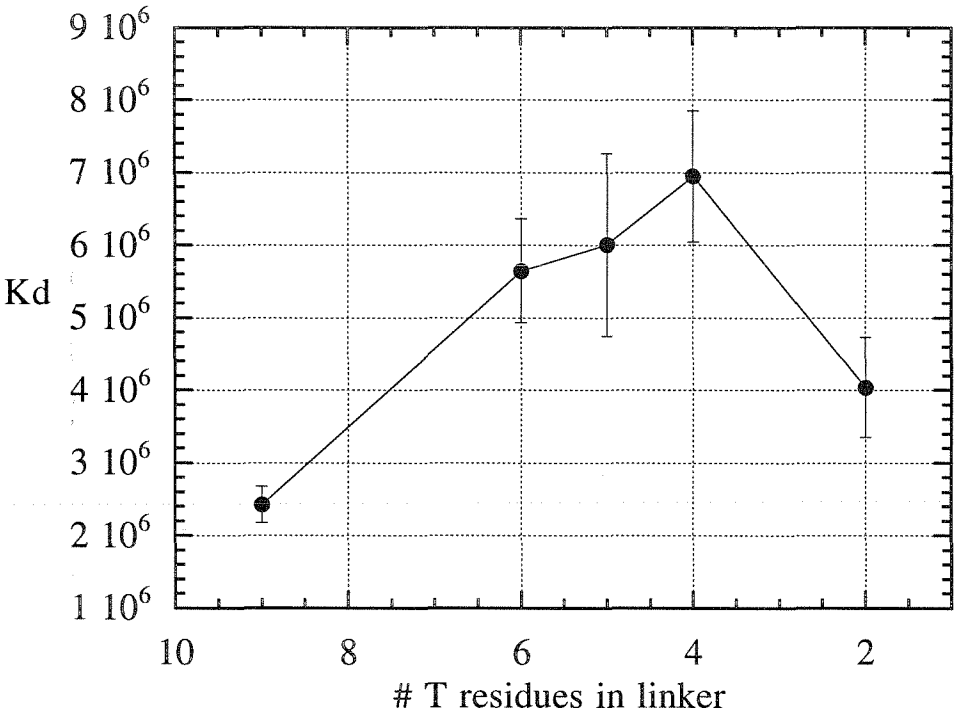


Figure 9. This graphical representation of the data from Table 2 shows the relationship between linker length and measured binding affinity (M^{-1}).

the bending penalty for this significantly bent oligonucleotide. Cooperative effects can be estimated physically by viewing the binding event at the second site as initiating at either the proximal or distal end of the second oligonucleotide. Given the need for oligonucleotide unstacking and elongation, initial binding at the distal end appears to be a better model. Calculated effective molarities for both models are presented in Table 3, where significant increases in effective concentration are predicted for all oligonucleotides studied. The values presented in Table 3 are obtained from geometrical modeling of the half sphere on the DNA face accessible to the third strand oligonucleotide bound at one site.

Additionally, the energy required to bend DNA towards the minor groove appears to be minimal, as viewed by the small energetic differences between the oligonucleotide binding affinities. This result has also been observed for a similar poly-ethylene glycol system designed in the purine motif (30, 31). The data obtained by Akiyama and Hogan are plotted in Figure 10. While the curve shape is different, the net effect is similar in the small observed penalty for bending DNA. In that study, bending was observed even with poly-ethylene glycol linkers that from geometrical measurement should connect the two domains without requiring DNA bending.

Another interesting observation was that the unlinked third strand oligonucleotides bore higher affinities than any of the linked compounds. Thus, there appears to be an energetic penalty for a large dangling end, which is probably entropic in nature. This contrasts with short dangling ends in RNA, which enhance a duplex binding affinity through base stacking (32).

Table 3. The upper table depicts a rough calculated estimate of the effective molarity of the second site oligonucleotide about its site following binding by the first and assuming association from the proximal end. The lower table depicts this assuming association from the distal end and may be the preferred model.

OLIGO	~LENGTH/T RESIDUE (Å)	RADIUS OF LINKER (Å)	HALF SPHERE VOL. (cc)	EFFECTIVE MOLARITY (M)
6-9	5	45	1.9×10^{-19}	8.7×10^{-3}
6-6	5	30	5.7×10^{-20}	2.9×10^{-2}
6-5	5	25	3.3×10^{-20}	5.1×10^{-2}
6-4	5	20	1.7×10^{-20}	9.9×10^{-2}
6-3	5	15	7.1×10^{-21}	2.4×10^{-1}
6-2	5	10	2.1×10^{-21}	7.9×10^{-1}

OLIGO	~LENGTH/T RESIDUE (Å)	RADIUS OF LINKER (Å)	HALF SPHERE VOL. (cc)	EFFECTIVE MOLARITY (M)
6-9	5	120	3.6×10^{-18}	4.6×10^{-4}
6-6	5	105	2.4×10^{-18}	6.9×10^{-4}
6-5	5	100	2.1×10^{-18}	7.9×10^{-4}
6-4	5	95	1.8×10^{-18}	9.3×10^{-4}
6-3	5	90	1.5×10^{-18}	1.1×10^{-3}
6-2	5	85	1.3×10^{-18}	1.3×10^{-3}

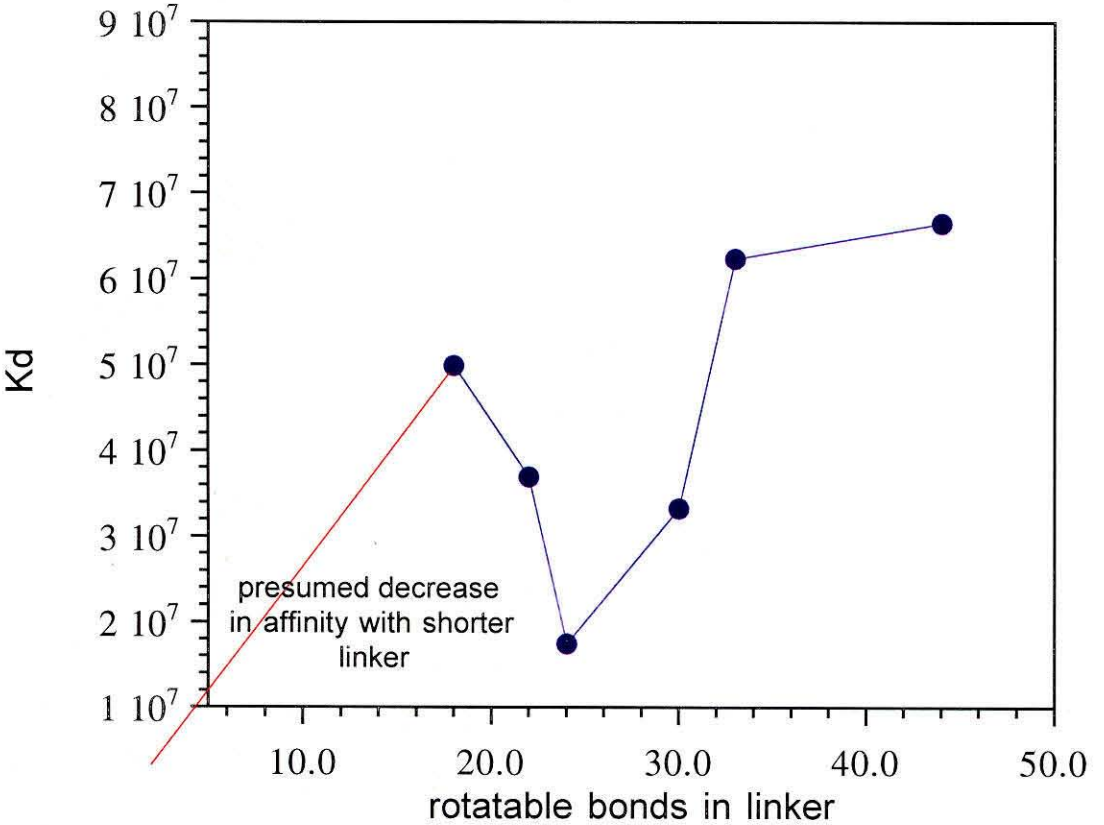


Figure 10. This is a graphical representation of the relationship between linker length and measured binding affinity (M^{-1}) for a purine motif oligonucleotide linked by polyethylene glycol as reported (31).

OLIGONUCLEOTIDES CONTAINING 5-PROPYNYL U AND THE EFFECT OF OLIGONUCLEOTIDE LENGTH

Other studies have also been undertaken to analyze the plasticity of the system including the effects of oligonucleotide length, sequence composition, pH, salt, linker orientation and composition, and target site size. Oligonucleotide and buffer conditions utilized are displayed in Table 1. Each oligonucleotide binding domain in the initial experiment (oligonucleotides **1** and **2**) has a binding affinity of approximately 2 x 10^7 M^{-1} for its cognate site and large bend angles are observed with bidentate oligonucleotide **6** (Tables 1, 4). In an attempt to find a shorter oligonucleotide with comparable binding affinity, the target sequence, salt conditions, and composition of the third strand were altered. 5-propynyl U has been shown to increase the affinity of third stand oligonucleotides due to its enhanced stacking capabilities (33). 11 base and 13 base oligonucleotides containing 5-propynyl U (oligonucleotides **3** and **4**) were designed with binding affinities in the same range as oligonucleotides **1** and **2** (Table 1).

Natural base pyrimidine motif oligonucleotides were designed to bind the same target sequences as the propynyl oligonucleotides. Oligonucleotide **5** also binds with an affinity in the same range as the original 15mer as gauged from a single gel (Table 4). Polyacrylamide gels to observe bending were run on natural base oligonucleotides **8** and **9** as well as analogues with 5-propynyl U. A natural base 13mer oligonucleotide (oligonucleotide **8**) binds to yield a bent structure with a smaller bend angle than the originally designed 15mer (oligonucleotide **6**) with approximately the same affinity (Table 5).

5-propynyl U containing oligonucleotides also do not bind to yield the expected large bend angles (Table 6). Oligonucleotide **10**, which presumably binds with a higher affinity, shows a similar amount of shifted DNA compared to oligonucleotide **6**,

Table 4. The affinities of monomer oligonucleotides are reported here. Generally, measured numbers are reported from an average of three or more gels. Numbers where fewer gels have been run are indicated, but the values should be considered with the above reservation.

OLIGONUCLEOTIDE	BUFFER	AFFINITY (M^{-1})	NUMBER GELS
1	A	$2.2\pm0.2\times10^{7}$	3
2	A	$1.7\pm0.1\times10^{7}$	3
3	D	7.6×10^{6}	1
3	E	$4.2\pm0.5\times10^{7}$	2
4	E	$2.2\pm0.3\times10^{7}$	3
5	E	1.6×10^{7}	1

Table 5. A summary of the results obtained with natural base oligonucleotides with T lnkers is reported.

OLIGONUCLEOTIDE	DUPLEX BP	BUFFER	RESULT
6	10	A	bending at no linker->50% bending at T=3
6	10	B	bending detected for T=3
6	10	C	bending observed at T=4, significant seen bending at T=5
7	10	A	discrete structure formed by pulled back third strand
8	10	A	bent to small angle
8	10	D	bent to small angle
9	10	D	no bent structures

Table 6. A summary of the results obtained with 5-propynyl U containing oligonucleotides in the pyrimidine motif is reported.

OLIGONUCLEOTIDE	DUPLEX BP	BUFFER	RESULT
6	10	D	discrete bent structure
10	10	D	same bent structure plus other foldback structures
8	10	A	bent to small angle
8	10	D	bent to small angle
11	10	A	bent to larger but still small angle
11	10	D	bent to larger but still small angle
9	10	D	no bent structures
12	10	D	bent to small angle

but unlike the natural base analog, it does not form one discrete highly bent structure as seen by relative gel retardation of the bend site in phase with an A tract (Figure 11). This assay is a simple gel shift, where the degree of retardation corresponds with bend angle and multiple less retarded bands are seen for oligonucleotide **10**. The shorter propynyl oligonucleotides (oligonucleotides **11** and **12**) exhibit comparable affinities to oligonucleotide **6** but do not bend DNA to the same degree. Still, they are more effective than the lower affinity natural base oligonucleotides of the same size (oligonucleotides **8** and **9**). In general, DNA solutions containing propynyl oligonucleotides at 4 °C precipitate more readily than T containing oligonucleotides and precipitation may be due to their increased hydrophobicity, although the reason for this is not clear. Overall, attempts to improve binding affinity with 5-propynyl U bending oligonucleotides were unsuccessful.

THE EFFECT OF BINDING SITE SPACING AND pH

In addition to changing the size of the oligonucleotide, the spacing of the intervening double helix was increased from 10 base pairs to 20 base pairs. Increasing the number of intervening base pairs should enable larger bend angles with the same monomer binding energy because the distortion to each individual base pair is lessened for the same net bend angle. Significantly bent structures were in fact observed by electrophoresis as visualized by relative retardation of the target fragment by oligonucleotide **6** (Table 7, Figure 12). Here, a decreasing proportion of duplex appears in an increasingly retarded discrete band, which is detectable for a 3T linker at pH 5.5 and a 7 T linker at pH 7.1. This experiment also partially demonstrates the generalizability of this technique for distal site recognition and bending by triple helix formation for varying spacer distances.

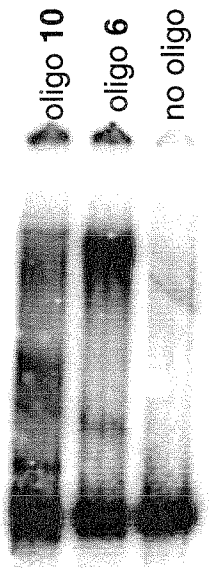


Figure 11. A phosphorimager generated image (linear dynamic range of 50) of a gel shift contrasting the binding properties of linked natural base and 5-propynyl U containing oligonucleotides.

Table 7. A summary of the results obtained by varying the spacing of the target site is reported.

OLIGONUCLEOTIDE	DUPLEX BP	BUFFER	RESULT
6	10	A	bending at no linker->50% bending at T=3
6	10	B	bending bending detected for T=3
6	20	A	bending at T=3, significant bending detected at T=5
6	20	B	bending at T=7, significant bending detected at T=9

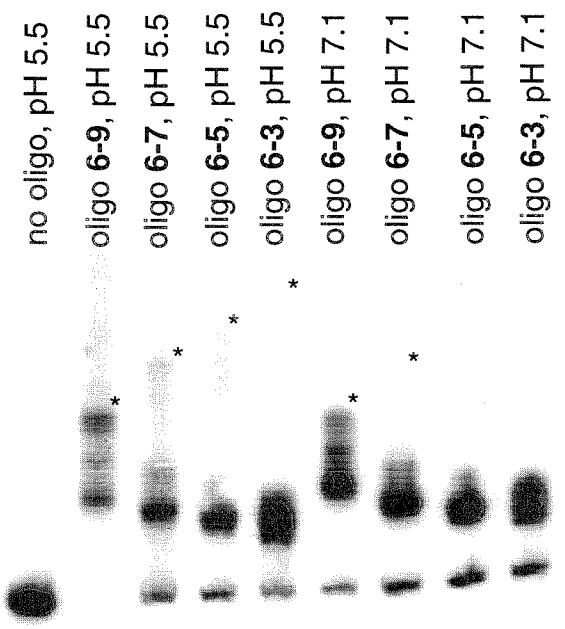


Figure 12. A phosphorimager generated image (linear dynamic range of 50) of a gel shift analysis for the effect of pH on generation of a bent structure for oligonucleotides separated by two turns of the helix instead of the original design involving one turn.

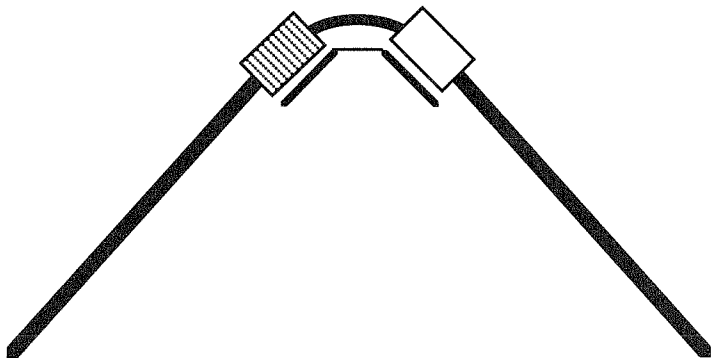
REVERSE LINKED OLIGONUCLEOTIDES

Another set of oligonucleotides was designed in which the spacer was moved from a position proximal to the intervening double helix to the other end of the third strand oligonucleotide (Figure 13). This design was envisioned as a method to remove the protection of the minor groove of the intervening bend, enabling chemistry at that site. Furthermore, if the bend can take place over a larger number of base pairs, the base pair distortion for each step should be lower, enabling larger net bend angles. In fact, this reverse linker gave smaller bend angles and electrophoretic mobilities in a circular permutation assay consistent with pulled back structures, where the ends of the third strand adjacent to the linker were not hydrogen bonded (Table 5, Figure 14). MPE footprinting of these structures gave results that were consistent with this model of a pulled back structure where the proximal ends of each binding site were not protected (Figure 15). While the bend is forming over a larger intervening region, the triple helix may be too rigid to accommodate a bent structure.

THE EFFECT OF pH AND SALT

Oligonucleotide **6** was equilibrated with the target DNA at pH 7.1 under low ionic strength and mock physiological conditions. Under these conditions, the oligonucleotide binds with lower affinity and longer linkers were required to obtain bending. While not as dramatic as the results at pH 5.5, these ligands did show binding to bent conformations at pH 7.1 (Table 5, Figure 16). Here, retarded bands are visualized for linkers of three and five T residues, although they represent a smaller fraction of the molecular ensemble than the respective structures at pH 5.5.

A.



B.

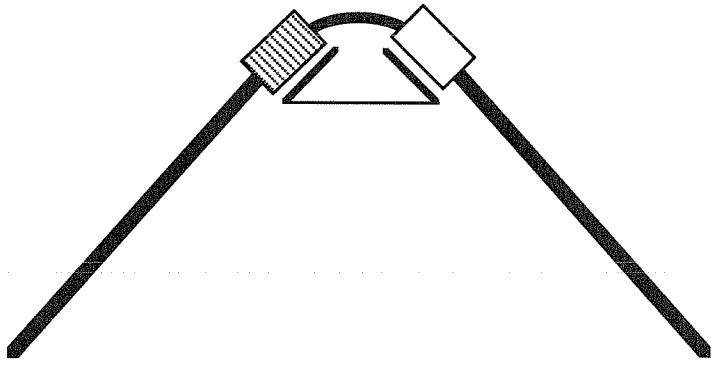


Figure 13. Reverse linked oligonucleotides, with distally linked oligonucleotides, are contrasted diagramatically with the original experimental design, where linkages are at proximal ends.

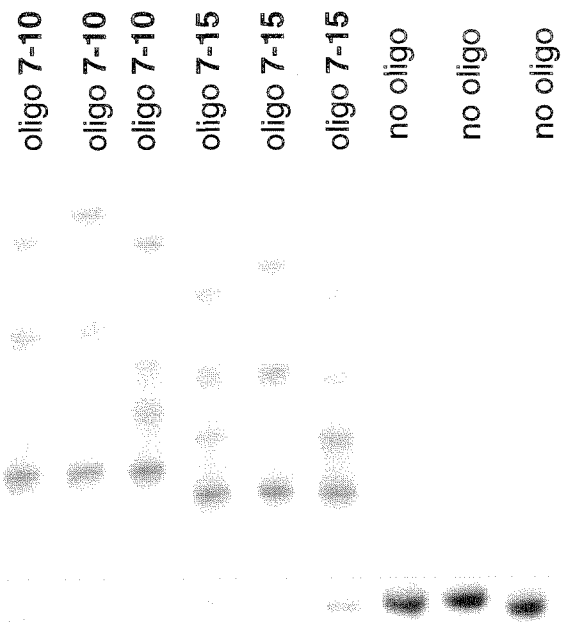


Figure 14. A phosphorimager generated image (linear dynamic range of 50) of a gel shift showing the bending properties of reverse linked oligonucleotides.

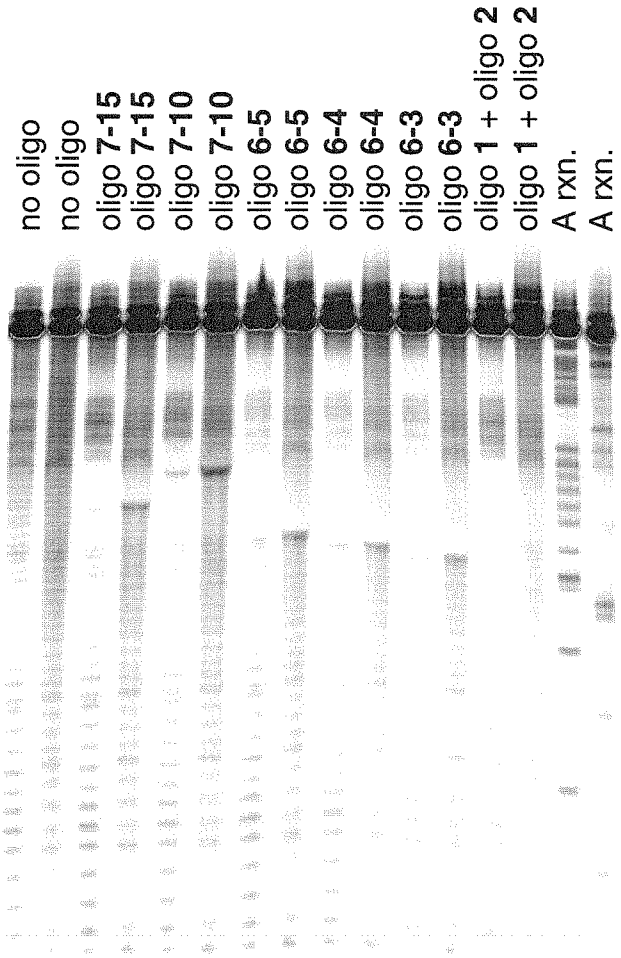


Figure 15. A phosphorimager generated image (linear dynamic range of 50) of an MPE footprinting gel contrasting the binding of forward and reverse linked oligonucleotides. The reverse linked oligonucleotides show pulled back structures which do not protect the entire third strand binding sites.

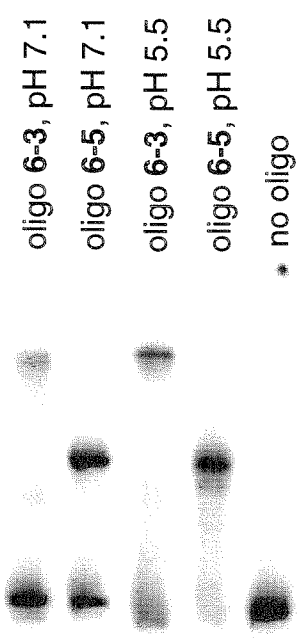


Figure 16. A phosphorimager generated image (linear dynamic range of 50) of a gel shift analysis for the effect of pH on generation of a bent structure for oligonucleotides separated by one turn of the helix.

PURINE MOTIF OLIGONUCLEOTIDES

To obtain oligonucleotides that bind with higher affinity at pH 7.1, the purine motif was explored as an alternative design. The same sequence initially targeted with the pyrimidine motif (oligonucleotide **6**) was targeted with the purine motif (oligonucleotide **13**). This oligonucleotide did not bind to give a bent structure under identical binding conditions (Table 8). A tighter binding sequence for purine motif triple helix formation was designed based upon quantitative research pursued in the Dervan group (oligonucleotide **14**) (5, 6, 34). Under two sets of conditions with moderate salt concentrations, oligonucleotide **14** gave multiple structures, some of which may be pulled back structures (Figure 17). More retarded structures were obtained as the concentration of monovalent and divalent cations was increased. This result is similar to the result obtained with oligonucleotide **10** where a high affinity oligonucleotide does not bind to give a majority of bent DNA. A similar result was seen in a literature study on purine motif bending oligonucleotides with a poly-ethylene glycol linker. Shorter linkers with 24, 22, or 18 rotatable bonds were observed to form multiple structures by gel electrophoresis (30). These poly-ethylene glycol linkers are comparable to the three T linker used in this study.

EFFECTS OF LINKER STACKING AND CHARGE

The effects of variation of stacking and charge in the linker were also examined. The initial design utilized a T residue. Alternative linkers were designed to examine the effect of stacking; these are depicted in Figure 18. Replacing each T residue with an abasic site (oligonucleotide **15**) produced an identically bound and bent species in a circular permutation assay, implying that the T residues do not stack (Table 9, Figure 19). This is seen in a calculated model for linkers of two and four T residues and in a

Table 8. A summary of results obtained with purine motif oligonucleotides is reported.

OLIGONUCLEOTIDE	DUPLEX BP	BUFFER	RESULT
13	10	A	no bent structures
13	10	F	no bent structures
14	10	A	bent to small angle
14	10	D	multiple structures-some very bent

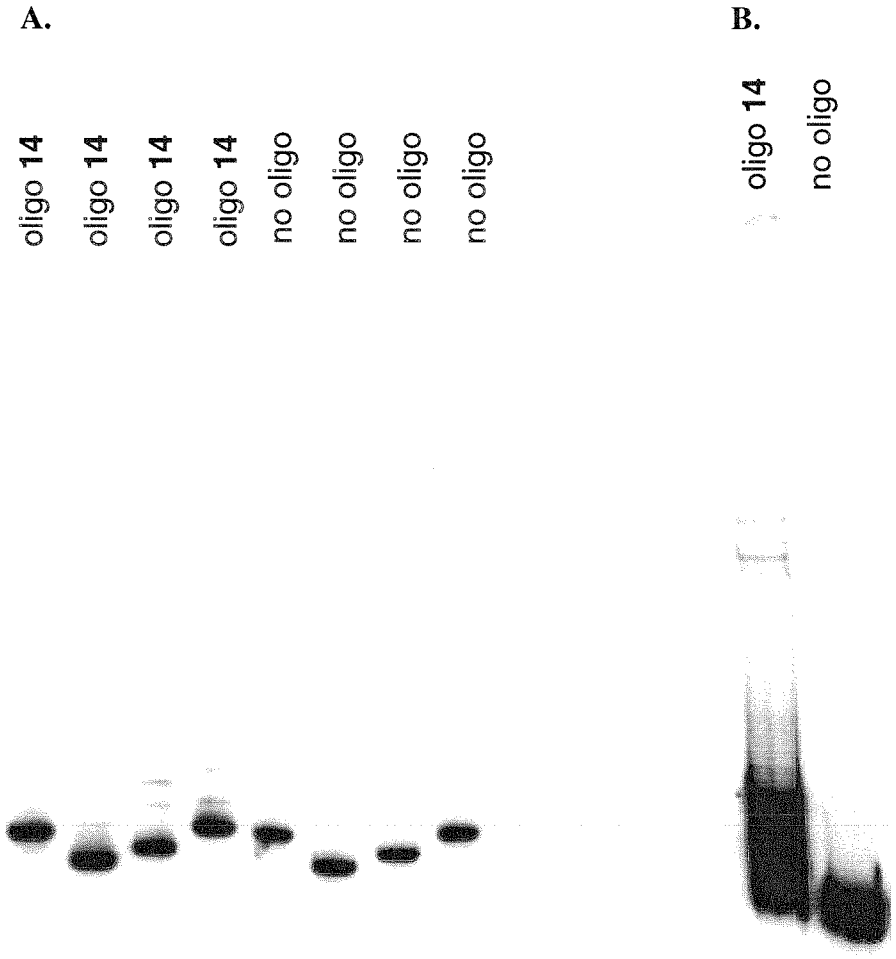


Figure 17. A phosphorimager generated image (linear dynamic range of 50) of a gel shift analysis for the binding of purine motif oligonucleotides linked by 3 T residues under varying salt conditions.

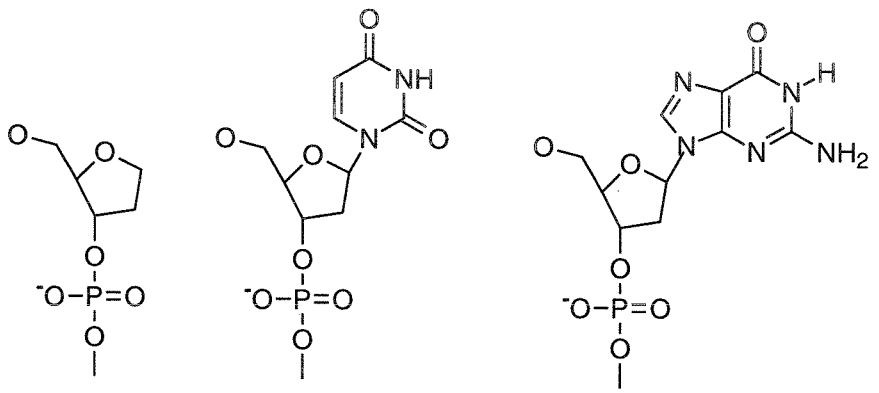


Figure 18. The abasic, T, and G residues used in linkers to analyze for stacking are presented.

Table 9. A summary of results where the effect of base stacking in the linker region was analyzed is reported.

OLIGONUCLEOTIDE	DUPLEX BP	BUFFER	RESULT
15	10	A	discrete bent structure
6	10	A	same discrete bent structure
16	10	A	some shifted to same large bend plus population in less bent conformation

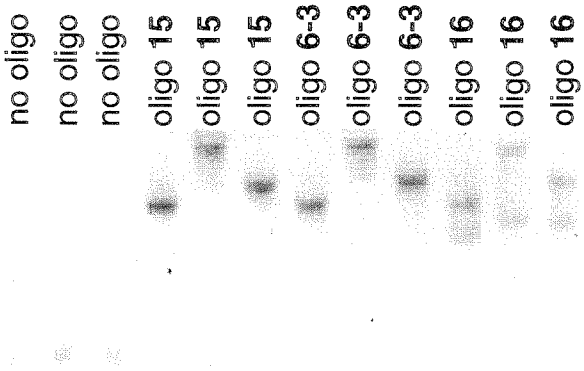


Figure 19. A phosphorimager generated image (linear dynamic range of 50) of a gel shift analysis of the binding of oligonucleotides containing a linker composed of abasic sites, T residues, or G residues. This analysis shows the importance of unstacking the linker in forming the bent structure.

crystal structure of single stranded poly-dT DNA, where the bases do not stack (Figure 20-22) (35). G residues have a larger stacking energy than T residues and they must be unstacked at an energetic price (reported to be -2.7 kcal/mol/base for cooperative triple helix formation) to yield the same bend as for a T residue (36). This energetic price leads to an alternative less bent structure also being populated electrophoretically with oligonucleotide **16** (Figure 19). The alternative conformation of a fully stacked 4-G strand is reduced from 28 Å to 20 Å and results in a much larger bend angle with the requisite energetic penalty. A model of a four G linker for comparison to the four T linker is presented in Figure 23. This structure has the G residues partially stacked with small distortion to the triple helical domains. The two third strand oligonucleotides are overlaid in Figure 24. While the T residues are unstacked and splayed away from the double helix and each other, the G residues are partially stacked, and although the phosphate backbone is somewhat stretched in the linker region, it is also necessary to distort the sugar and terminal two bases at the proximal ends of the triple helix to accommodate the intervening distance.

The effect of charge in the linker domain was also examined (Table 10). Novel designed linkers are depicted in Figure 25. With a nucleotide linker, changing from a phosphodiester (oligonucleotide **6**) to an uncharged methyl phosphonate (oligonucleotide **17**) had little effect on the bend angle or proportion of the ensemble in a bent conformation in a circular permutation assay (Figure 26). Similarly, with a freely rotating linker, changing from a neutral PEG (oligonucleotide **18**) to a positively charged polyamine (oligonucleotide **19**) had little effect on the bend angle or proportion of the ensemble in a bent conformation (Figure 27). The polyamine linker was a novel synthesized compound and its synthesis is depicted in Figure 28.

The absence of charge effects can be explained by considering charge screening. The Debye screening length is thought to be only several angstroms around duplex DNA and even shorter around triplex DNA (37, 38). In the modeled structure,

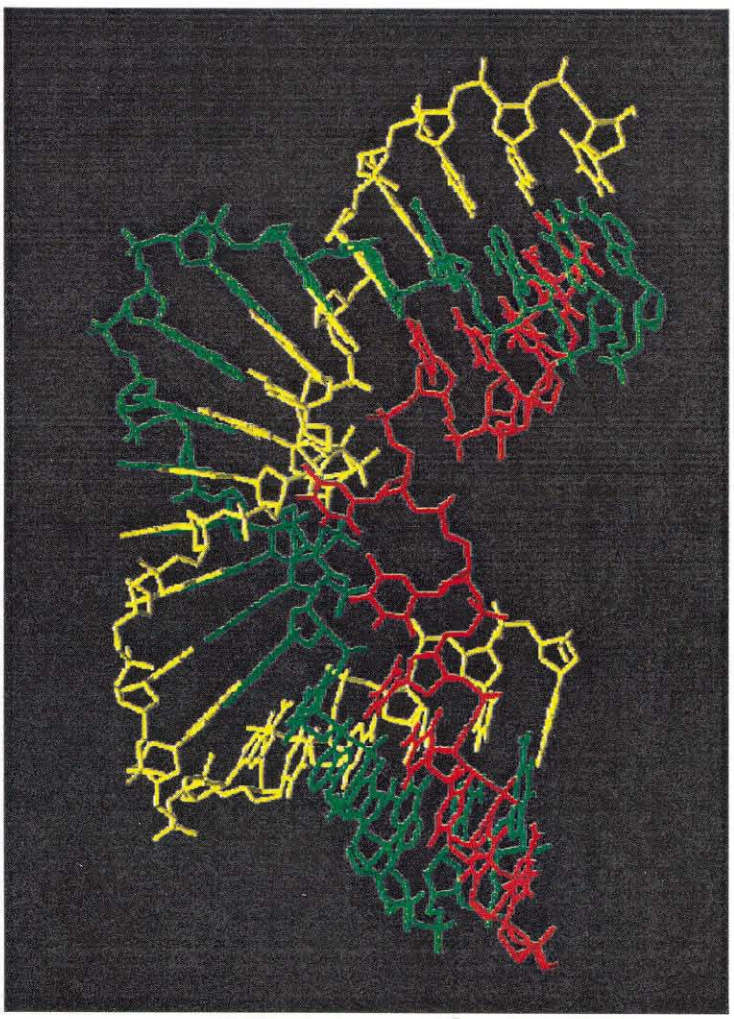


Figure 20. Modeled structure of bent complex formed by oligonucleotide **6-2**. Bent DNA was built using NAMOT and modeling was performed in Biograf (23). The pyrimidine strand is in yellow, the purine strand in green, and the third strand in red. This structure shows a large bend towards the minor groove. The DNA was modeled as smoothly bending and the base pair planes were maintained while the backbone conformation was minimized. The last residues of the triple helix were unconstrained to move away from their Hoogsteen bonds but remained bonded. The linker here shows minimal stacking interactions.

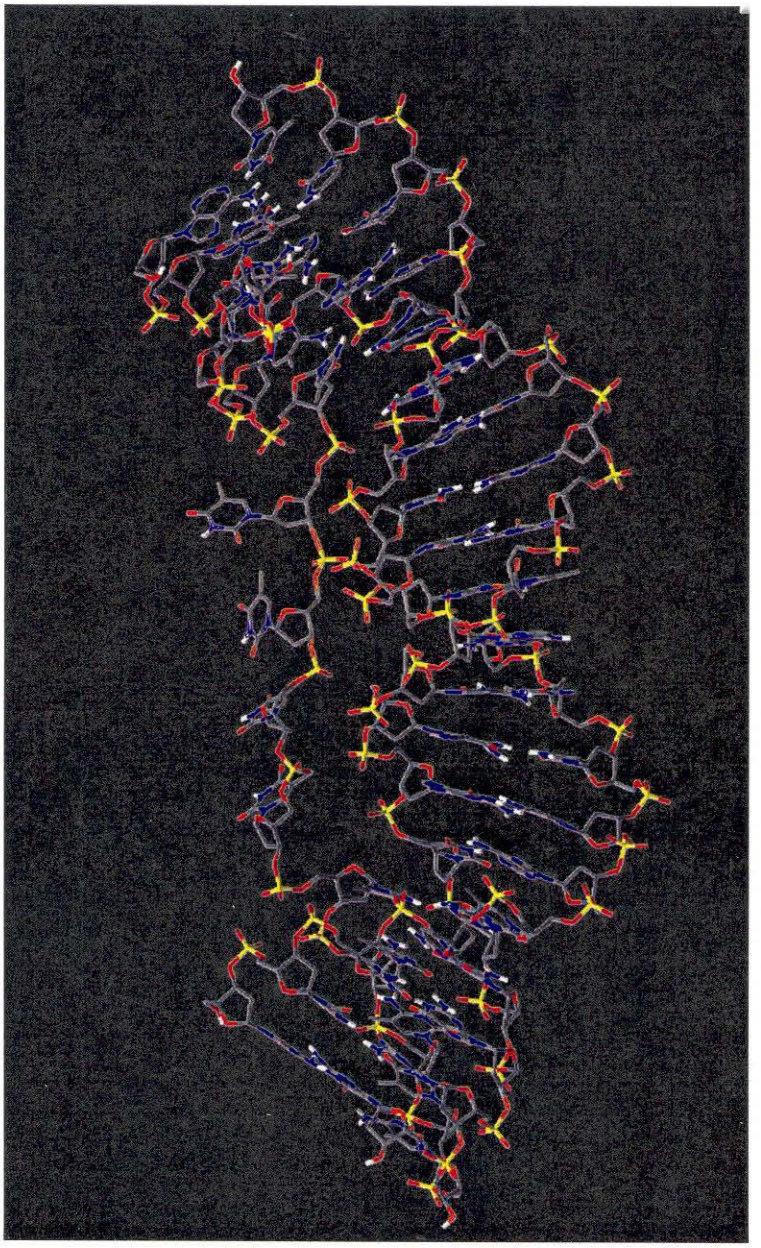


Figure 21. Modeled structure of bent complex formed by oligonucleotide **6-4**. Bent DNA was built using NAMOT and modeling was performed in Biograf (23). This structure shows a 61° bend towards the minor groove. The DNA was modeled as smoothly bending and the base pair planes were maintained while the backbone conformation was minimized. All residues of the triple helix except for the termini were unconstrained to move away from their Hoogsteen bonds but remained bonded. The linker here shows minimal stacking interactions.

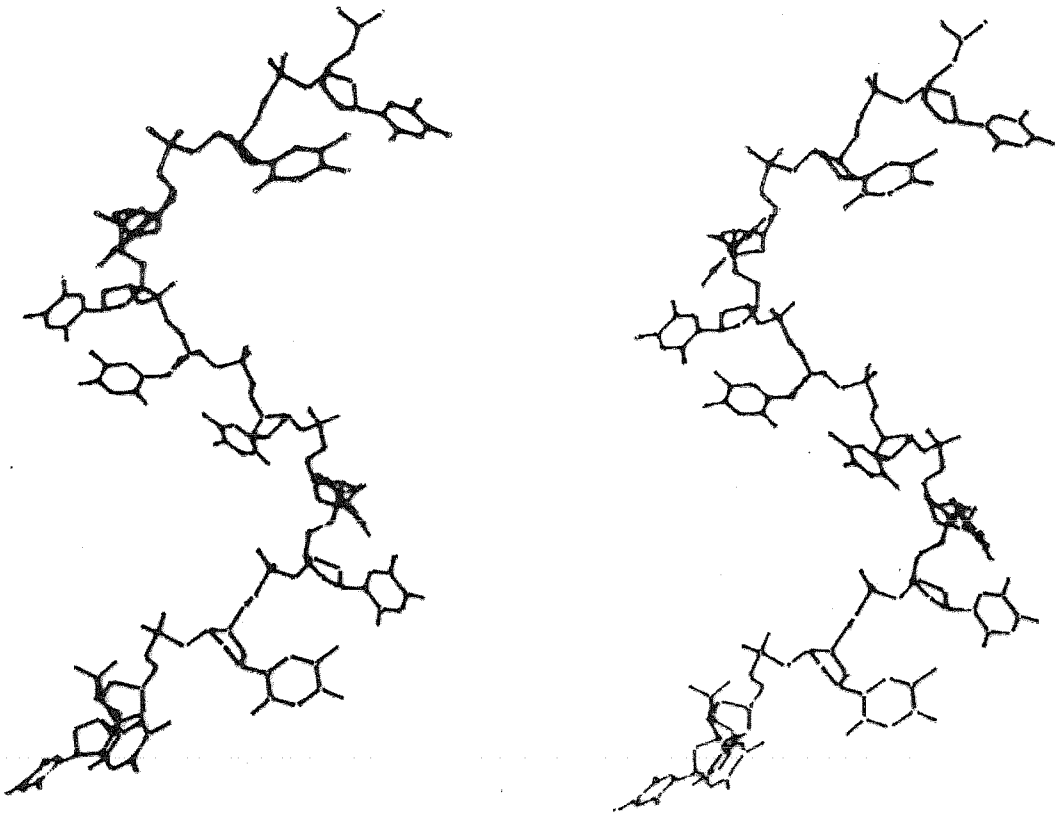


Figure 22. A crystal structure of oligo-dT shows T residues unstacked preferentially (35).

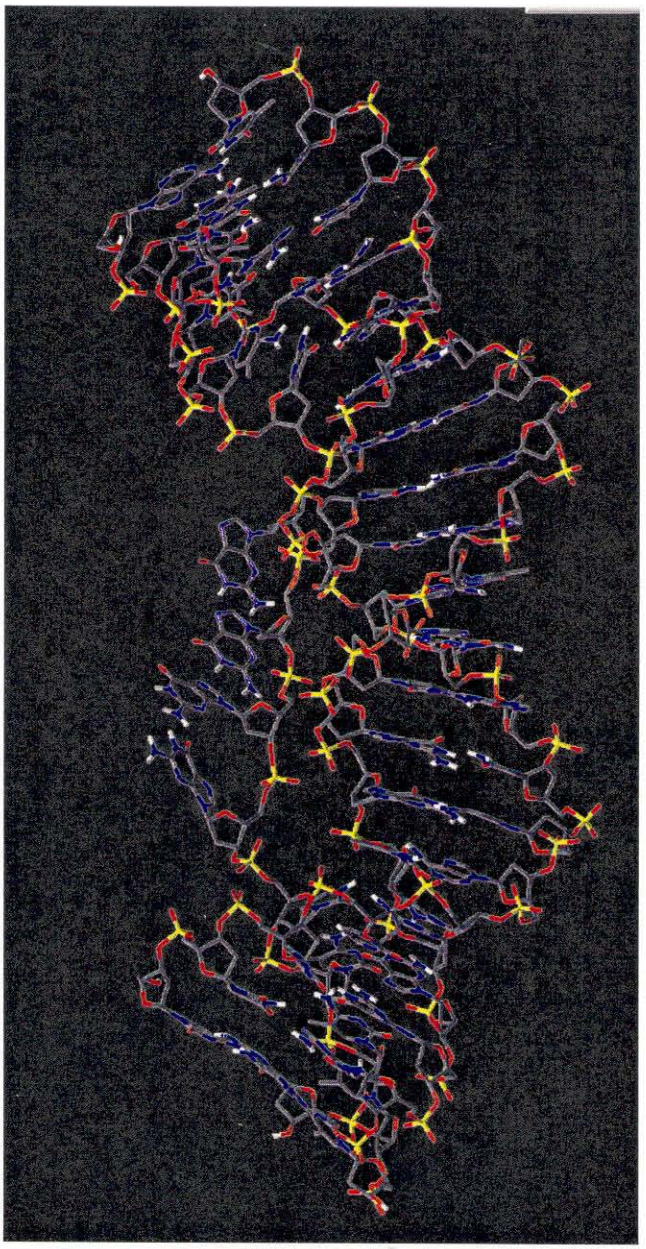


Figure 23. Modeled structure of bent complex formed by oligonucleotide **16**. Bent DNA was built using NAMOT and modeling was performed in Biograf (23). This structure shows a 61° bend towards the minor groove. The DNA was modeled as smoothly bending and the base pair planes were maintained while the backbone conformation was minimized. All residues of the triple helix except for the termini were unconstrained to move away from their Hoogsteen bonds but remained bonded. The linker here shows minimal stacking interactions.

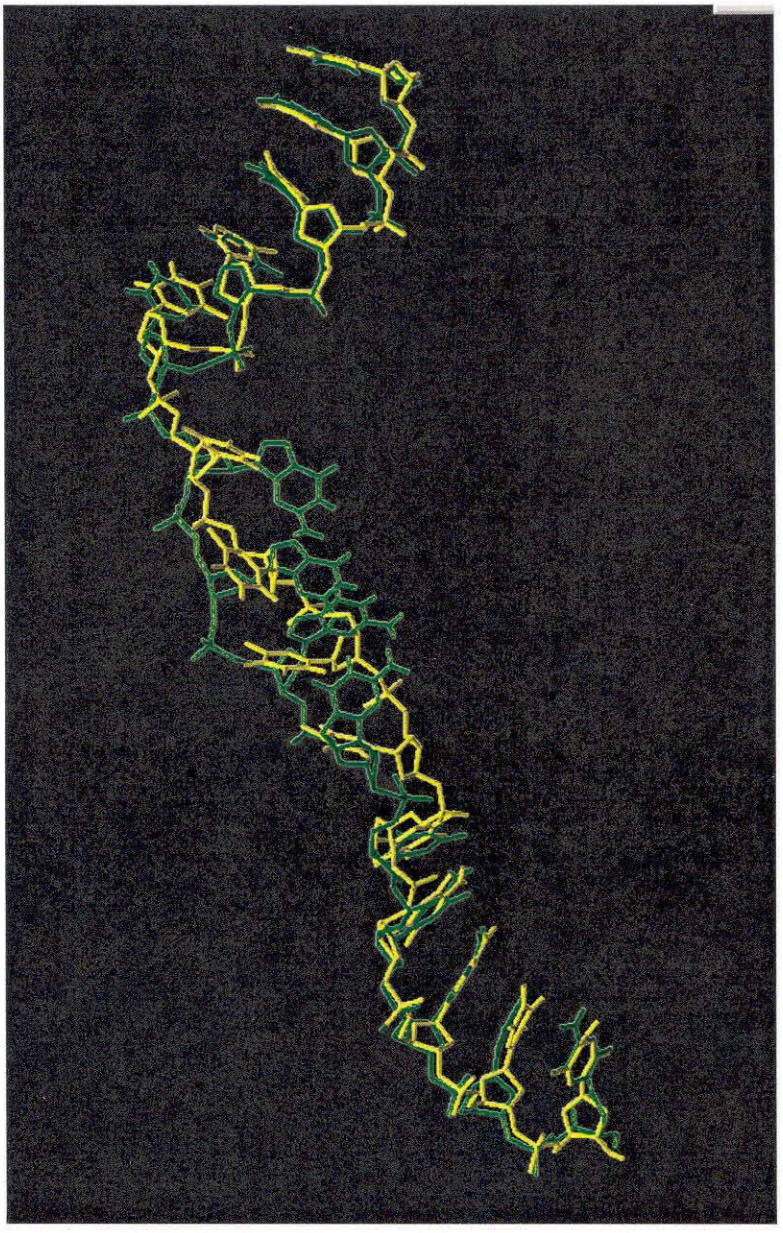
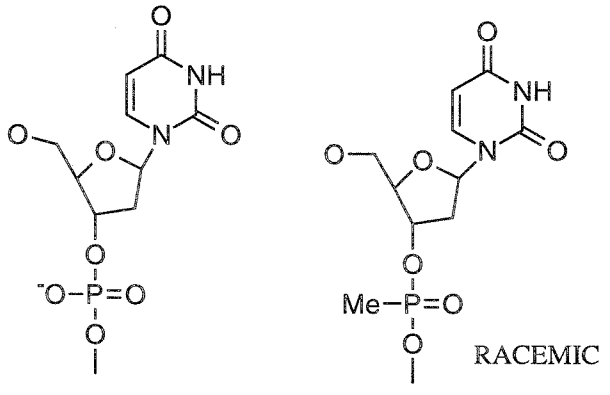


Figure 24. The third strands from Figure 21 and Figure 23 are overlaid to contrast the differences in stacking between T and G residues according to this model.

Table 10. A summary of results where the effect of linker charge on bending was analyzed is reported.

OLIGONUCLEOTIDE	DUPLEX BP	BUFFER	RESULT
6	10	A	discrete bent structure
17	10	A	same large discrete bend as **6**
18	10	A	large discrete bend
19	10	A	same large discrete bend as **18**

A.



B.

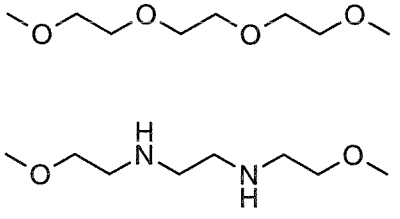


Figure 25. The residues used to analyze for the effects of charge are indicated here. A shows the nucleotidic linkages containing either a phosphodiester or racemic methyl phosphonate. B shows the freely rotating linkages of poly-ethylene glycol, and a novel polyamine.

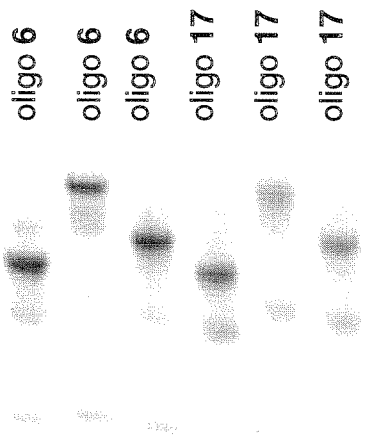


Figure 26. A phosphorimager generated image (linear dynamic range of 50) of a gel shift analysis for the effect of linker charge on the formation of a bent structure. Here, a methyl phosphonate linkage is contrasted with a phosphodiester linkage.

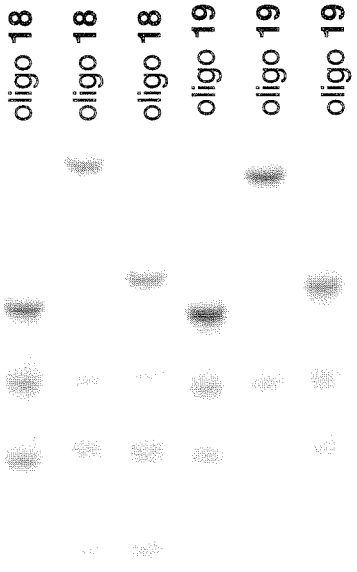


Figure 27. A phosphorimager generated image (linear dynamic range of 50) of a gel shift analysis for the effect of linker charge on the formation of a bent structure. Here, a poly-ethylene glycol linkage is contrasted with a novel polyamine linkage.

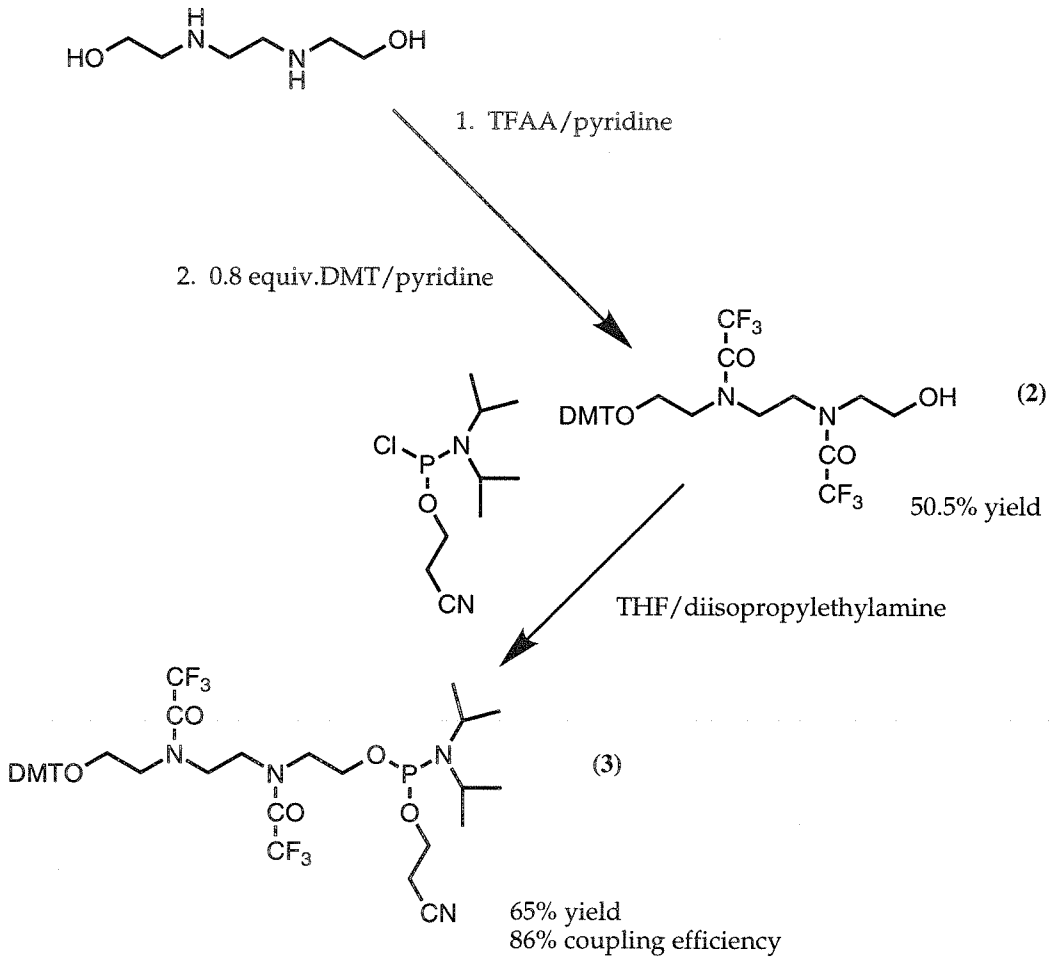


Figure 28. The synthetic scheme and yields for the synthesis of the novel polyamine utilized are shown.

a taut backbone bears phosphates that are no closer than 7 Å to duplex phosphates and the backbone itself is never closer than 4 Å to any duplex backbone position (Figures 20-21, 23). From the estimate above, it seems reasonable that the charge is largely screened by solution counterions. This is consistent with a literature report on purine-rich oligonucleotide mediated bending, where formation of the bent structure was largely independent of bulk [Mg^{++}].

The failure of the propynyl-modified or purine-rich oligonucleotides to form discrete bent structures may be tied to their increased rigidity in triple helix formation. It is not known if the proximal regions of the third strand are fully hydrogen bonded in a canonical B-like triple helical structure. If they are distorted, then these triple helical domains must pay a more substantial energetic penalty than a pyrimidine-rich triple helix, which does not base stack as well. This is similar to the case of the poly-G linker, where stacking energy creates a less flexible structure. One kinetic model for the formation of these bent structures holds that binding at one or both sites is dependent upon nucleation at the distal end(s) of the binding site. An oligonucleotide must stretch to accommodate binding from the distal end, and this is more easily facilitated for oligonucleotides lacking purine and propynyl oligonucleotides. The alternative model of binding at the proximal end requires more substantial oligonucleotide stretching or inherent duplex flexibility in conjunction with oligonucleotide stretching.

OBSERVATIONS, APPLICATIONS, AND CONCLUSIONS

A hypothetical free energy diagram depicting the binding event for oligonucleotide **6** is shown in Figure 29. However, a model for a more rigid triple helix induced by purine-rich or propynyl oligonucleotide binding, results in a change of

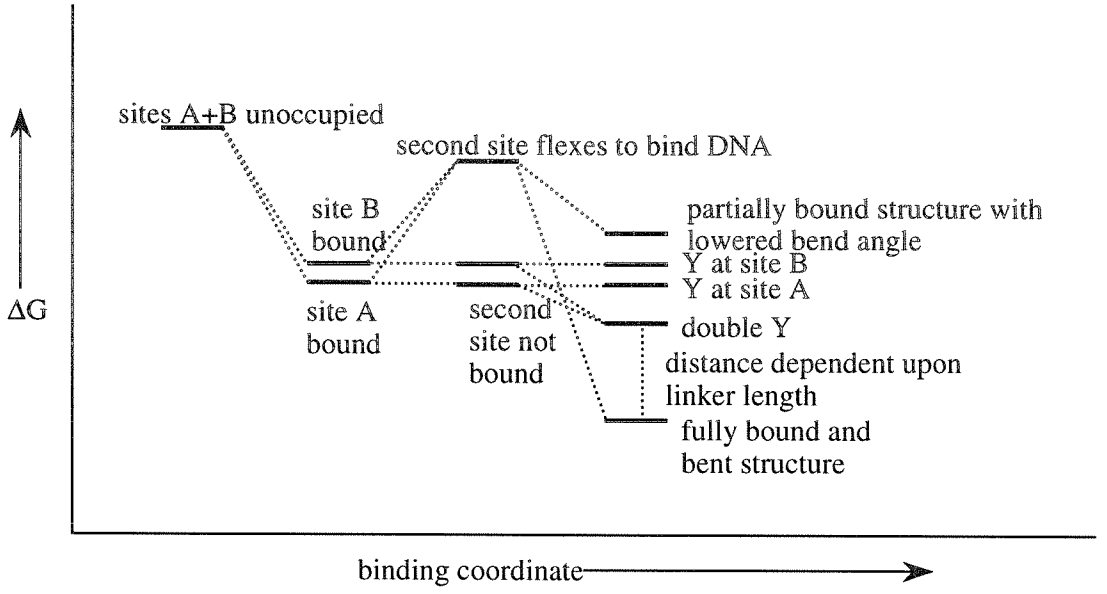


Figure 29. A hypothetical free energy diagram depicting the binding of flexible natural base pyrimidine motif third strand oligonucleotides to form bent structures is depicted.

the relative energies of the final states and is shown in Figure 30. In this case, multiple pulled back structures seem to be the most favorable thermodynamically.

An additional experiment was performed on this system. Although a full statistical analysis was not performed, transmission electron microscopy was used as a direct physical technique to visualize bent structures. Using oligonucleotide **6**, DNA with large bend angles was visualized.

To study potential applications of triple helix mediated DNA bending, studies were performed with HIV-1 integrase. HIV-1 integrase is known to be sensitive to DNA curvature and bending (39, 40). Research in the Dervan group sought to direct integration of mock HIV genomes using HIV-1 integrase and oligonucleotides **6** and **7**. However, these studies showed protection of the bent region from integration, probably from the linker in the minor groove (41). The failure of these studies may indicate the utility of these ligands towards DNA looping or major groove binding molecules, as opposed to large proteins that interact with the bent minor groove.

In a parallel experimental design, minor groove binding polyamides were appended by a flexible linker to create molecules to bend DNA towards the major groove. However, these molecules did not appreciably bend DNA in a phasing analysis (42). Two differences are apparent between polyamide mediated bending and triple helical bending. The dissociation kinetics of polyamides are much faster than for third strand oligonucleotides and the electrophoretic assay may have measured an average of associated and dissociated polyamide (43, 44). More fundamentally, DNA bending towards the major groove, which is where the bend center is for minor groove bound polyamides, may be more costly energetically than bending towards the minor groove.

Sequence specific DNA bending mediated by bidentate triple helix formation is a complex process. The studies outlined in this chapter provide a biophysical characterization of the nature of ligand binding. The plasticity of this DNA binding

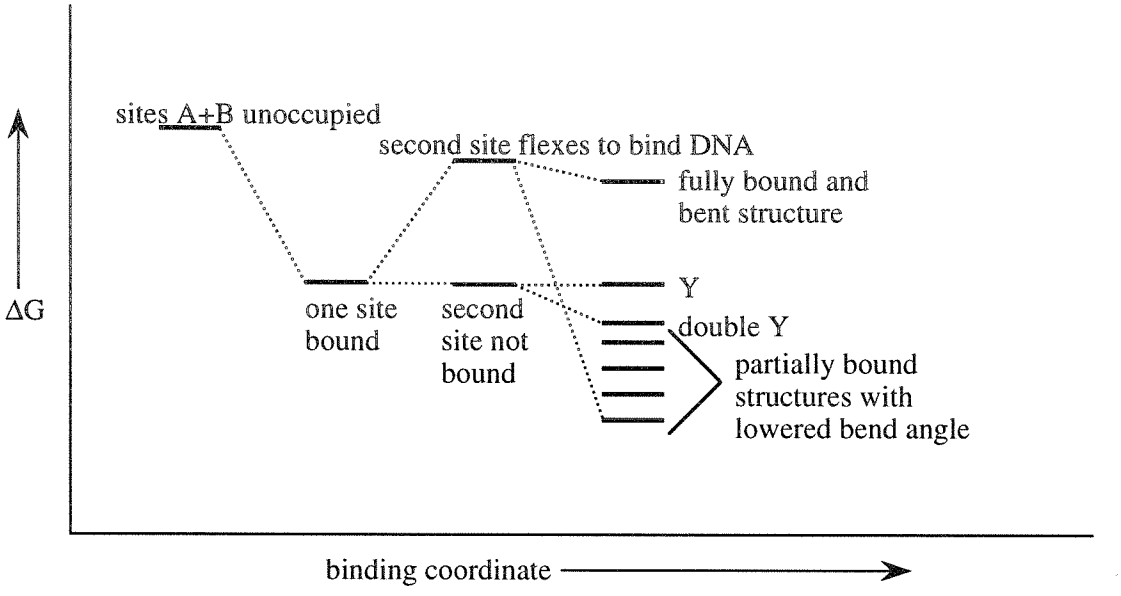


Figure 30. A hypothetical free energy diagram depicting the binding of strongly stacking purine motif and pyrimidine motif oligonucleotides containing 5-propynyl U is shown.

motif has been delineated along with its limitations. From here, additional potential applications of these ligands can be pursued for both biological, medical, and biophysical purposes.

REFERENCES

1. Liberles, D. A. & Dervan, P. B. (1996) *Proc. Natl. Acad. Sci., USA* **93,** 9510-4.
2. Moser, H. E. & Dervan, P. B. (1987) *Science* **238,** 645-650.
3. LeDoan, T., Perrouault, L., Praseuth, D., Habhoub, N., Decout, J., Thuong, N. T., Lhomme, J. & Helene, C. (1987) *Nucleic Acids Research* **15,** 7749-7760.
4. Cooney, M., Czernuszewicz, G., Postel, E. H., Flint, S. J. & Hogan, M. E. (1988) *Science* **241,** 456-459.
5. Beal, P. A. & Dervan, P. B. (1991) *Science* **251,** 1360-1363.
6. Beal, P. A. & Devan, P. B. (1992) *Nucleic Acids Research* **20,** 2773-2776.
7. Singleton, S. F. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 6957-6965.
8. Roberts, R. W. & Crothers, D. M. (1991) *Proc. Natl. Acad. Sci., USA* **88,** 9397-9401.
9. Rougee, M., Faucon, B., Mergny, J. L., Barcelo, F., Giovannageli, C., Garestier, T. & Helene, C. (1992) *Biochemistry* **31,** 9269-9278.
10. Fosella, J. A., Kim, Y. J., Shih, H., Richards, E. G. & Fresco, J. R. (1993) *Nucleic Acids Research* **21,** 4511-4515.
11. Best, G. C. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 1187-1193.
12. Hunziker, J., Priestley, E. S., Brunar, H. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 2661-2662.
13. Povsic, T. J. & Dervan, P. B. (1989) *J. Am. Chem. Soc.* **111,** 3059-3061.
14. Xodo, L. E., Manzini, G., Quadrifoglio, F., vanderMarel, G. A. & vanBoom, J. H. (1991) *Nucleic Acids Research* **19,** 5625-5631.
15. Singleton, S. F. & Dervan, P. B. (1992) *Biochemistry* **31,** 10995-11003.
16. Singleton, S. F. & Dervan, P. B. (1993) *Biochemistry* **32,** 13171-13179.

17. Singleton, S. F. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 10376-10382.

18. Distefano, M. D. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 11006-11007.

19. Hogrefe, R. I., Vaghefi, M. M., Reynolds, M. A., Young, K. M. & Arnold, L. J. (1993) *Nucleic Acids Research* **21,** 2031-2038.

20. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986) *Proc. Natl. Acad. Sci., USA* **83,** 8462-8466.

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

22. Ferber, M. J. & Maher, L. J. (1997) *Analytical Biochemistry* **244,** 312-320.

23. Tung, C. & Carter, E. S. (1994) *Cabios* **10,** 427-433.

24. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S. & Weiner, P. (1984) *J. Am. Chem. Soc.* **106,** 765-784.

25. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. (1986) *J. Computational Chemistry* **7,** 230-252.

26. Pearlman, D. A. & Kollman, P. A. (1990) *Biopolymers* **29,** 1193-1209.

27. Singh, U. C., Weiner, S. J. & Kollman, P. (1985) *Proc. Natl. Acad. Sci., USA* **82,** 755-759.

28. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. (1974) in *Physical Chemistry of Nucleic Acids* (Harper and Row, New York), pp. 159-166.

29. Cantor, C. R. & Schimmel, P. R. (1980) *Biophysical Chemistry* (W.H. Freeman and Company, New York).

30. Akiyama, T. & Hogan, M. E. (1996) *Proc. Natl. Acad. Sci., USA* **93,** 12122-12127.

31. Akiyama, T. & Hogan, M. E. (1996) *J. Biological Chemistry* **271,** 29126-29135.

32. Sugimoto, N., Kierzek, R. & Turner, D. H. (1987) *Biochemistry* **26,** 4554-4558.

33. Colocci, N. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 785-786.

34. Greenberg, W. A. & Dervan, P. B. (1995) *J. Am. Chem. Soc.* **117,** 5016-5022.

35. Camerman, N., Fawcett, J. K. & Camerman, A. (1976) *J. Molecular Biology* **107,** 601-621.

36. Colocci, N. & Dervan, P. B. , unpublished results.

37. Bond, J. P., Anderson, C. F. & Record, M. T. (1994) *Biophysical Journal* **67,** 825-836.

38. Philpott, M. R. & Glosli, J. N. (1995) *J. Electrochem. Soc.* **142,** L25-L28.

39. Muller, H. P. & Varmus, H. E. (1994) *EMBO J.* **13,** 4704-4714.

40. Milot, E., Belmaaza, A., Rassart, E. & Chartrand, P. (1994) *Virology* **201,** 408-412.

41. Mouscadet, J. F., Liberles, D. A. & Dervan, P. B. , unpublished results.

42. Albert, J. S. & Dervan, P. B. , unpublished results.

43. Albert, J. S., Baird, E. E. & Dervan, P. B. (1997) , submitted.

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

CHAPTER 5

SEQUENCE SPECIFIC INHIBITION OF LIGAND INDUCED DNA BENDING BY POLYAMIDES

Intrinsic DNA curvature and protein induced DNA bending are important in the regulation of gene activation, replication initiation, and other processes (1, 2). DNA is an inherently flexible polymer and neutral backbone analogs of DNA curve, where some rigidity is maintained in natural DNA by coulombic repulsion between phosphates on the same strand (3, 4). Sequence dependent curvature of DNA is caused both by differential solvation in the minor groove and differential base stacking leading to variation of roll and tilt values (5, 6). Proteins and other ligands that bend DNA alter the stacking of the bases by intercalation of hydrophobic groups, alter the effective Debye length of the surface through charge screening, or bend through energetic compensation for tight binding events. An example of a protein that seems to work through all three mechanisms in bending DNA by greater than 160^0 is integration host factor (IHF) (7).

Previously, it has been shown that artificial sequence specific DNA bending ligands can be designed that utilize bidentate tight binding third strand oligonucleotides to constrict the intervening duplex and bend DNA (8-11). Here, we present the design of complementary nonnatural architectural factors to straighten DNA. Compounds that bind in the minor groove such as distamycin and DAPI have been shown to alter DNA rigidity (12-14). However, compounds such as DAPI bind with low relative affinity and specificity, forming few specific contacts where binding is dominated by their positive charge. Polyamide analogs of distamycin composed of pyrrole and imidazole carboxamides have been designed that form specific high affinity structures with DNA in the minor groove. Sequence specificity is determined by the sequence of side-by-

side amino acid pairings, where imidazole (Im) opposite pyrrole (Py) recognizes a GC base pair, Py-Im recognizes CG, Py-Py is degenerate for AT or TA, while Im-Im pairing is strongly disfavored. The generality of this recognition motif has been demonstrated for a large number of sequences and is directly supported by NMR structural data (see (15) and references therein).

Previously it has been shown that polyamides bound in the minor groove can support simultaneous binding of major groove ligands that do not distort DNA such as a third strand oligonucleotide and GCN4 (16, 17). However, proteins that distort DNA such as TATA binding protein (TBP) and homeobox proteins are displaced by distamycin at very high concentrations (18-20). Furthermore, intrinsically curved A tracts are straightened by distamycin, although the mechanism of this may be through disruption of solvation in the minor groove as well as through direct rigidification (5, 13, 14). Here, it is demonstrated that a DNA bending ligand can be displaced through binding of 100 pM polyamide to a bent region not contacted by the ligand. Furthermore, the generalizability of polyamide binding may allow straightening of nearly any desirable sequence.

MATERIALS AND METHODS

Synthesis. Oligonucleotides were synthesized and purified as previously described (8). Polyamides were also synthesized as previously described (21). Briefly, synthesis was performed in a stepwise manner on Boc-β-Alanine-Pam resin using Boc protected monomers. Polyamides were then cleaved by reaction with ((dimethylamino)propyl)-amine and purified by reverse phase HPLC.

Preparation of Labeled DNA. Plasmid DNA was digested with HindIII and EcoRI for gel shift analysis, or PvuII and EcoRI for footprinting analysis, and

simultaneously labeled with Sequenase 2.0, deoxyadenosine 5'-[α-^{32}P]triphosphate, thymidine 5'-[α-^{32}P]triphosphate, and nonradioactive deoxynucleoside triphosphates. The fragment was purified by gel electrophoresis, treated with proteinase K, filtered, further extracted with phenol/chloroform, and precipitated with ethanol.

DNase I Footprint Reactions. All reactions were equilibrated at 22° C, pH 5.5, in the presence of 10 mM Mes, 1 mM $MgCl_2$, and labeled DNA for at least 24 hours. Footprinting reactions were carried out as previously described (22).

Gel Shift Assay and Titration. All samples were equilibrated as above for at least 24 hours. One-tenth volume of 15% glycerol loading buffer was added and samples were run on a 10% polyacrylamide gel at 4° C, pH 5.5, with a 75:1 acrylamide/bis-acrylamide ratio in 45 mM Mes and 1 mM $MgCl_2$ with buffer recirculation. Data were collected by storage phosphor densitometry (Molecular Dynamics 400S Phosphorimager). Quantitation of isotherms was performed by plotting the ligand concentration against the portion of labeled DNA in the bent conformation and curve fit using a Langmuir binding isotherm, as previously described (22, 23).

RESULTS AND DISCUSSION

Two 15 bp purine tracts separated by one turn of the DNA helix (10 bp) were targeted by oligonucleotides containing two pyrimidine tracts (T and ^{Me}C) connected by a central poly-T linker of 2-9 nucleotides. It has been previously shown that such oligonucleotides bend DNA to a varying degree dependent upon the linker length (8). The intervening 10 bp not targeted by the third strand oligonucleotide can be bound by

a polyamide specifically designed for that sequence, as depicted in Figure 1. The binding affinity of the polyamide for its target sequence has previously been determined under similar conditions of low ionic strength and reported to be K_{eq}= 3.7 x 10^{10} M^{-1} for **PA1** and K_{eq}= 5.0 x 10^{8} M^{-1} for **PA2**, both much stronger than the affinity of either the unlinked 15mers for their sites, or the linked bidentate oligonucleotides which bind with K_{eq}=~10^{7} M^{-1} (8, 15, 24). Given the demonstrated ability of other minor groove binding ligands to straighten DNA, here we demonstrate that targeting a polyamide to the intervening duplex of bent DNA can straighten it, displacing a ligand bound several base pairs distally, as depicted in Figure 2.

PA1 at a concentration of 100 pM can displace a DNA bending third strand oligonucleotide present at a concentration of 1 μM. This effect is clearly seen in Figure 3 for oligonucleotides with linkers of 2 and 3 T residues, which bend DNA by greater than 60^{0}. In these lanes, the more retarded bent structure is shifted to a less retarded structure. This mobility is similar to a double Y structure, where two oligonucleotides are bound by a single target DNA molecule with the polyamide presumably bound between them, but is also consistent with binding by a single oligonucleotide and dissociation of the ends proximal to the polyamide binding site in a D loop to result both in a lower net bend angle and a smaller bend angle per base pair step. For oligonucleotides with smaller bend angles, the polyamide does not displace the third strand oligonucleotide and the DNA distortion (bend angle/base pair step) is probably not sufficient to preclude efficient polyamide binding. The sequence specificity of this effect is seen in Figures 4a and 4b, where only **PA1**, but not **PA2** can displace oligonucleotide 2 at a concentration of 100 pM, while neither can displace oligonucleotide 9. Similarly, polyamides do not displace ligands such as third strand

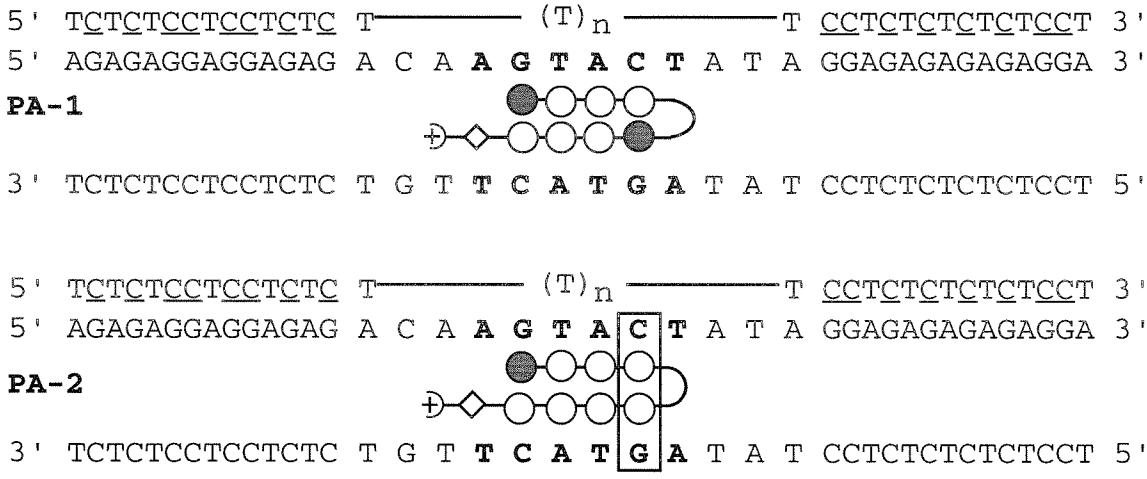


Figure 1a. A third strand oligonucleotide with the indicated sequence binding domains (C is 5-methyl Cytosine) linker by a variable T linker binds in the major groove of DNA, with the two binding domains recognizing sequences 10 bp apart. The center of this 10 bp region is bent towards the minor groove by the third strand oligonucleotide. This site is targeted by a polyamide to straighten it and displace the third strand oligonucleotide from its bidentate form. Polyamide sequences are indicated, where a dark circle represents an imidazole, a light circle a pyrrole, the curved line represents γ-aminobutyric acid, and the diamond, β-alanine. The mismatch in **PA2** is indicated by a box.

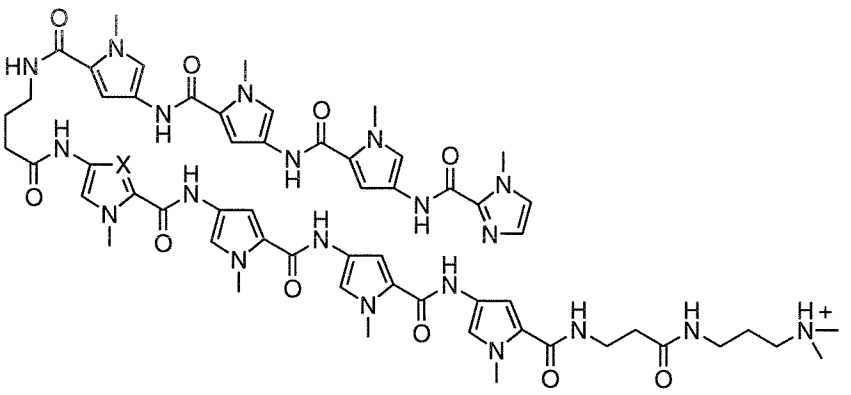


Figure 1b. The chemical structures of **PA1** and **PA2** are presented, where X=N for PA1 and CH for PA2.

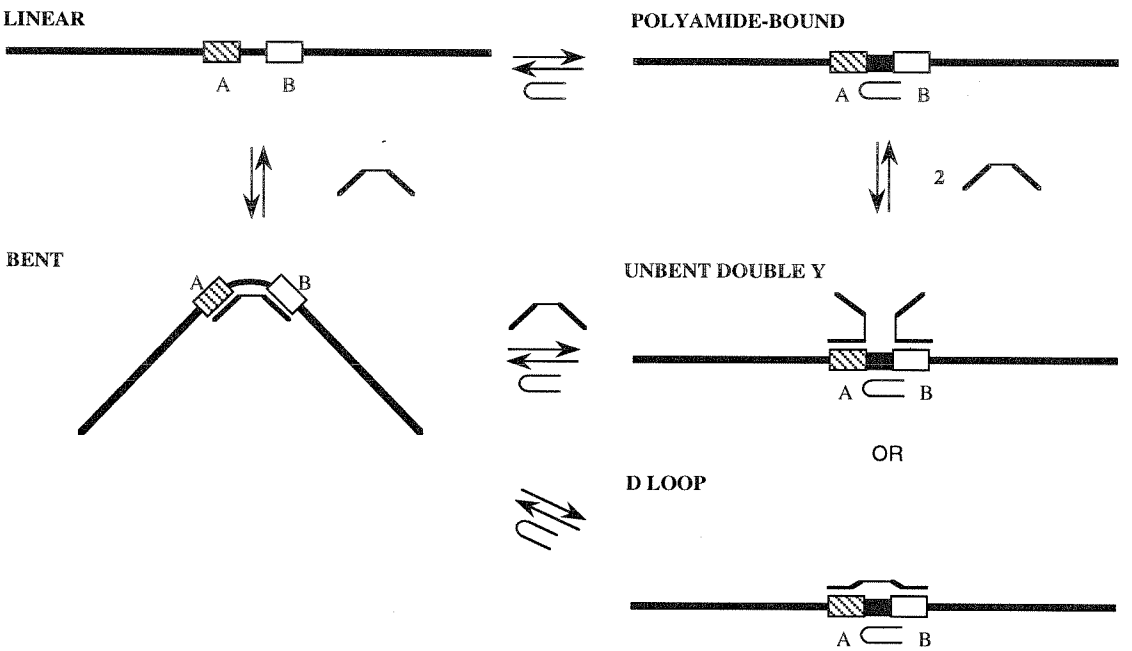


Figure 2. This depiction of the experimental design shows the binding of the bending oligonucleotide on the left, where polyamide binding to the right displaces the bidentate oligonucleotide from one of its binding sites to form either a double Y structure, where two third strand oligonucleotides bind to one DNA strand with a polyamide bound in the intervening 10 bp of DNA or a D loop with one oligonucleotide bound in a straight conformation.

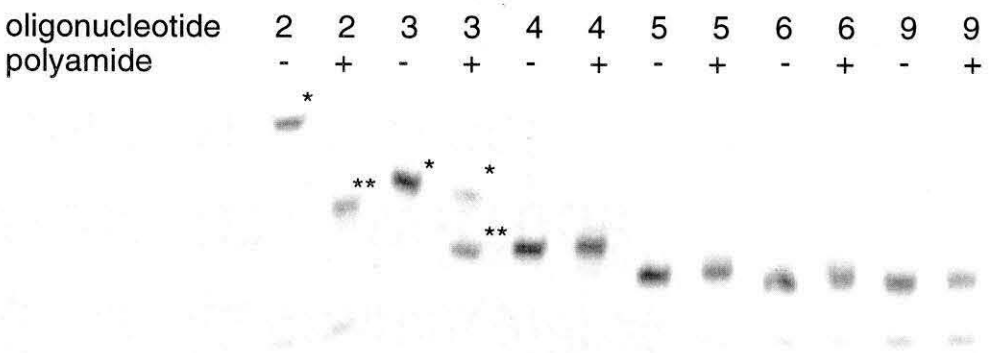


Figure 3. PhosphorImage using a linear dynamic range of 50 of a 10% (75:1 acrylamide/bis ratio) polyacrylamide gel run in 45 mM Mes, 1 mM $MgCl_2$, pH 5.5 at 4° C showing a gel shift analysis of oligonucleotides **2-6**, **9** on a 3' ^{32}P end-labeled restriction fragment generated with EcoRI and HindIII. Incubations included 1 μM oligonucleotide and 100 pM **PA1**, as indicated at 22° C at pH 5.5 with 45 mM Mes and 1 mM $MgCl_2$ for at least 24 hours. Bent structures are indicated by * and straightened structures are indicated by **.

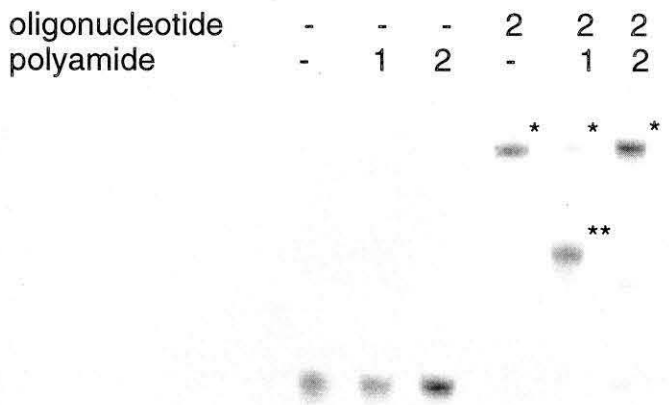


Figure 4a. PhosphorImage using a linear dynamic range of 50 of a 10% (75:1 acrylamide/bis ratio) polyacrylamide gel run in 45 mM Mes, 1 mM $MgCl_2$, pH 5.5 at 4° C showing a gel shift analysis of oligonucleotide **2** on a 3' ^{32}P end-labeled restriction fragment generated with EcoRI and HindIII. Incubations included 1 μM oligonucleotide and 100 pM **PA1** or **PA2** , as indicated at 22° C at pH 5.5 with 45 mM Mes and 1 mM $MgCl_2$ for at least 24 hours.

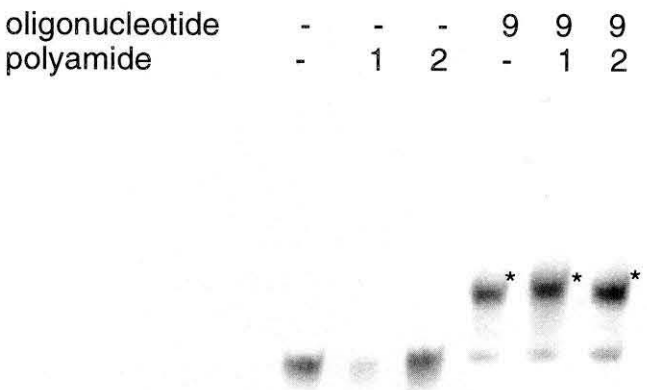


Figure 4b. This gel is the same as 4a, where oligonucleotide **9** was included in place of oligonucleotide 2. In both gels, bent structures are indicated by * and straightened structures are indicated by **.

oligonucleotides and the protein GCN4, which don't bend DNA to a significant degree (16, 17, 25).

These experiments were performed with simultaneous addition of polyamide and oligonucleotide. Next, we sought to determine if the order of addition was important, given that the reported half life of a bound third strand oligonucleotide on DNA is relatively long (approximately 12 hours under near physiological conditions) (26). As shown in Figure 5, simultaneous addition, preincubation with polyamide for at least 1 hour, or preincubation with third strand oligonucleotide for at least 1 hour made little difference in the ability of the polyamide to inhibit binding of the third strand oligonucleotide, where 1 hour is expected to be sufficient for significant association of both oligonucleotide and polyamide (26, 27). These results suggest that polyamides may be useful in targeting prebound transcription factors in cells.

It has been shown that the energy required to bend DNA with bidentate triple helical ligands is less than predicted by theoretical models of DNA as a smoothly bending wormlike chain with coulombic repulsion between phosphates placed at fixed distances (10, 24, 28, 29). However, the prebending of target DNA for TBP, a general transcription factor, significantly altered its binding affinity (30). We sought to examine the effects of prebending on the affinity of the polyamide by measuring K_I, the inhibition constant against triplex mediated bending. Figure 6 shows a sample gel shift titration, where the measured K_I for **PA1** is 6.2 ($\pm$0.3) x 10^{10}, whereas K_I for **PA2** is $<1 \times 10^9$ M^{-1}. K_I for **PA1** is slightly larger than K_{eq}, the equilibrium association constant, measured by DNase I footprinting and shows that the energy for straightening DNA bent by a ligand with a lower binding affinity is minimal (15). The higher value for K_I may be due to experimental differences in measurement from the use of different techniques and different low ionic strength solution conditions, or may represent the need for association with but not saturation of the target site to preclude binding by DNA bending third strand oligonucleotides. Bending effects with TBP may be much

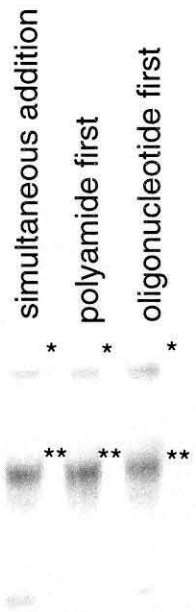


Figure 5. PhosphorImage using a linear dynamic range of 50 of a 10% (75:1 acrylamide/bis ratio) polyacrylamide gel run in 45 mM Mes, 1 mM $MgCl_2$, pH 5.5 at 4° C showing a gel shift analysis of oligonucleotide **2** on a 3' ^{32}P end-labeled restriction fragment generated with EcoRI and HindIII. Incubations included 1 μM oligonucleotide and 100 pM **PA1** at 22° C at pH 5.5 with 45 mM Mes and 1 mM $MgCl_2$ for at least 24 hours. Here, the first lane shows simultaneous addition of polyamide and oligonucleotide, the second, polyamide added one hour before oligonucleotide, and the third with oligonucleotide added one hour before polyamide. Bent structures are indicated by * and straightened structures are indicated by **.

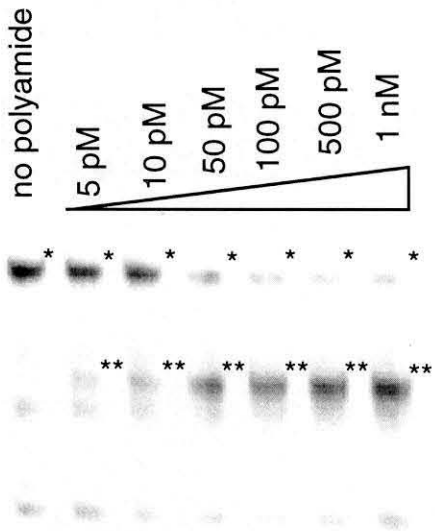


Figure 6. PhosphorImage using a linear dynamic range of 50 of a 10% (75:1 acrylamide/bis ratio) polyacrylamide gel run in 45 mM Mes, 1 mM $MgCl_2$, pH 5.5 at 4° C showing a gel shift analysis of oligonucleotide **2** on a 3' ^{32}P end-labeled restriction fragment generated with EcoRI and HindIII. Incubations included 1 μM oligonucleotide and a titration of **PA1** from 5 pM to 1 nM at 22° C at pH 5.5 with 45 mM Mes and 1 mM $MgCl_2$, equilibrated for at least 24 hours. Bent structures are indicated by * and straightened structures are indicated by **.

larger, given the large number of protein-DNA contacts supporting the bent structure (31, 32). Additionally, TBP-bound DNA is bent towards the major groove, not the minor groove. Despite this, TBP can in fact be displaced by distamycin bound at the same site with high bulk solution concentrations(18, 19).

To confirm that binding of the polyamide is indeed in the intervening duplex between triplex sites, DNase I footprinting was performed, as shown in Figure 7. While no binding in the intervening duplex and no displacement of the third strand oligonucleotide are seen for **PA2**, **PA1** protects the intervening duplex. The footprint remains at the triple helix target sites, where oligonucleotides remain bound in a nonbent conformation. This footprinting pattern is consistent with the binding model presented.

The mechanism of action of polyamides is assumed to be directly through rigidification of the double helix. Alternative modes of action to be considered are simple steric blockage, or disruption of the solvation shell or counterion shell to alter third strand oligonucleotide affinity. Modeling of the triple helix-mediated bend shows a linker that is displaced from the intervening duplex to accommodate the bend angle, where shorter linkers are displaced further from the duplex than longer linkers. The inability of **PA1** to displace bending oligonucleotides with longer linkers and bend angles less than 60^0 argues against the steric blockage explanation. Disruption of the solvation shell in the minor groove is likely to be steric and therefore unlikely to extend into the triple helical region where no polyamide is bound. Furthermore, this minor groove effect is unlikely to affect binding of a third strand in the major groove, where simultaneous binding has previously been demonstrated (16). Similarly, the high charge density of double helical and triple helical DNA results in a large counterion shell around the molecule and in a very short Debye screening length where charge effects are unlikely to extend for multiple base pairs from the polyamide binding site to the third strand oligonucleotide binding site (33, 34). In fact, it may be partially

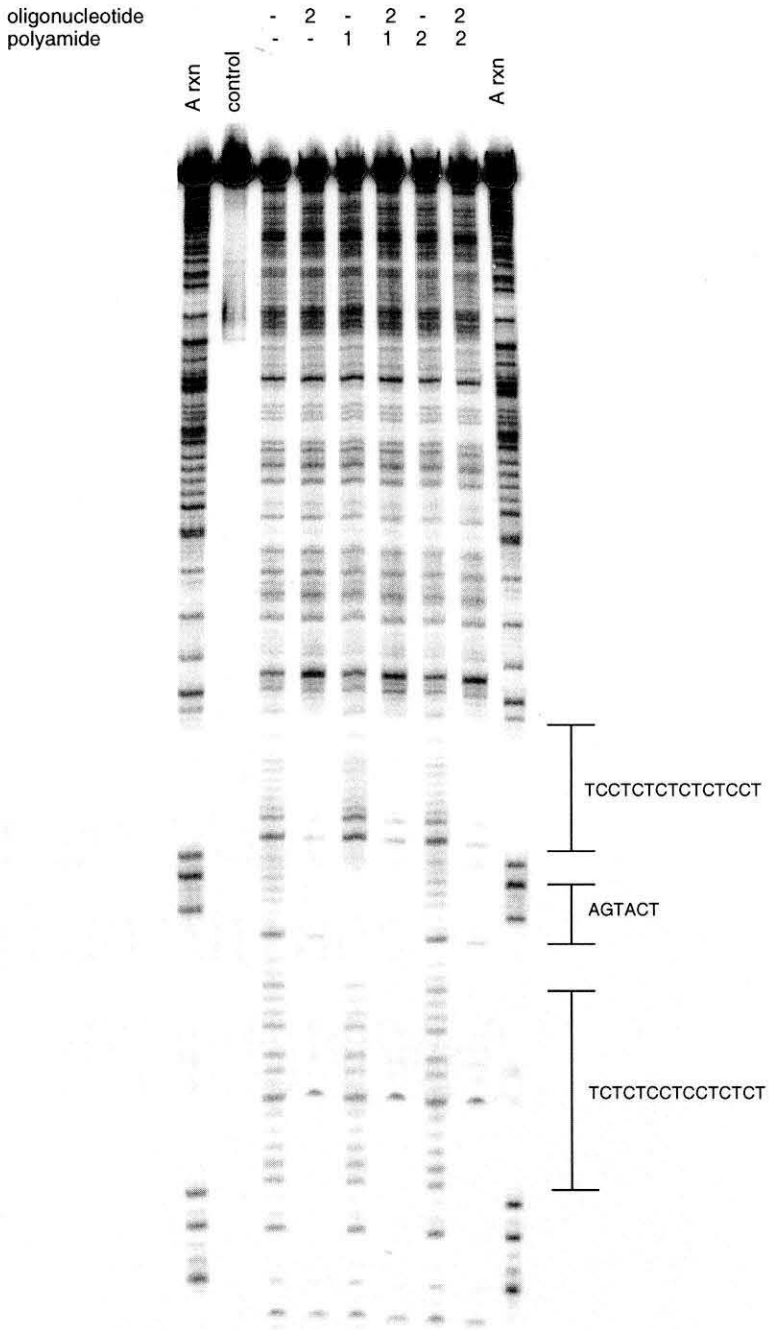


Figure 7. PhosphorImage using a linear dynamic range of 50 of an 8% denaturing polyacrylamide gel run in TBE showing DNase I footprinting analysis of oligonucleotide **2** on a 3' ^{32}P end-labeled restriction fragment generated with EcoRI and PvuII. Incubations included 1 μM oligonucleotide and 100 pM **PA1** and **PA2**, as indicated at 22° C at pH 5.5 with 45 mM Mes and 1 mM $MgCl_2$ for at least 24 hours.

because of this large counterion shell that DNA bending towards the minor groove is so easily accommodated (10, 24, 35).

By modifying DNA structure, generalizable sequence specific polyamides have been designed to displace a DNA bending ligand at an adjacent but nonoverlapping binding site. This ability to displace DNA bending ligands through rigidification may be useful in the design of polyamides as artificial regulators of gene expression, providing a potentially valuable tool in molecular biology and human medicine.

REFERENCES

1. Perez-Martin, J., Rojo, F. & DeLorenzo, V. (1994) *Microbiological Reviews* **58,** 268-290.
2. Polaczek, P., Kwan, K., Liberles, D. A. & Campbell, J. L. (1997) , submitted.
3. Strauss, J. K. & Maher, L. J. (1994) *Science* **266,** 1829-1834.
4. Manning, G. S. (1983) *Biopolymers* **22,** 689-729.
5. Dlakic, M., Park, K., Griffith, J. D., Harvey, S. C. & Harrington, R. E. (1996) *J. Biological Chemistry* **271,** 17911-17919.
6. Bolshoy, A., McNamara, P., Harrington, R. E. & Trifonov, E. N. (1991) *Proc. Natl. Acad. Sci., USA* **88,** 2312-2316.
7. Rice, P. A., Yang, S., Mizuuchi, K. & Nash, H. A. (1996) *Cell* **87,** 1295-1306.
8. Liberles, D. A. & Dervan, P. B. (1996) *Proc. Natl. Acad. Sci., USA* **93,** 9510-4.
9. Akiyama, T. & Hogan, M. E. (1996) *Proc. Natl. Acad. Sci., USA* **93,** 12122-12127.
10. Akiyama, T. & Hogan, M. E. (1996) *J. Biological Chemistry* **271,** 29126-29135.
11. Akiyama, T. & Hogan, M. E. (1997) *Biochemistry* **36,** 2307-2315.
12. Larsson, A., Akerman, B. & Jonsson, M. (1996) *J. Physical Chemistry* **100,** 3252-63.
13. McCarthy, J. G. & Rich, A. (1991) *Nucleic Acids Research* **19,** 3421-9.
14. Barcelo, F., Muzard, G., Mendoza, R., Revet, B., Roques, B. P. & Pecq, J.-B. L. (1991) *Biochemistry* **30,** 4863-73.
15. Trauger, J. W., Baird, E. E. & Dervan, P. B. (1996) *Nature* **382,** 559-61.

16. Parks, M. E. & Dervan, P. B. (1996) *Bioorganic & Medicinal Chemistry* **4**, 1045-50.

17. Oakley, M. G., Mrksich, M. & Dervan, P. B. (1992) *Biochemistry* **31**, 10969-75.

18. Chaudhary, D. & Miller, D. M. (1995) *Biochemistry* **34,** 3438-3445.

19. Chiang, S.-Y., Welch, J. J., Rauscher, F. J. & Beerman, T. A. (1996) *J. Biological Chemistry* **271,** 23999-24004.

20. Dorn, A., Affolter, M., Muller, M., Gehring, W. J. & Leupin, W. (1992) *EMBO J.* **11,** 279-86.

21. Baird, E. E. & Dervan, P. B. (1996) *J. Am. Chem. Soc.* **118,** 6141-6.

22. Brenowitz, M., Senear, D. F., Shea, M. A. & Ackers, G. K. (1986) *Proc. Natl. Acad. Sci., USA* **83,** 8462-8466.

23. Ferber, M. J. & Maher, L. J. (1997) *Analytical Biochemistry* **244,** 312-320.

24. Liberles, D. A. & Dervan, P. B. , unpublished data.

25. Keller, W., Konig, P. & Richmond, T. J. (1995) *J. Molecular Biology* **254,** 657-667.

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

27. Albert, J. S., Baird, E. E. & Dervan, P. B. (1997) , submitted.

28. Bloomfield, V. A., Crothers, D. M. & Tinoco, I. (1974) in *Physical Chemistry of Nucleic Acids* (Harper and Row, New York), pp. 159-166.

29. Fenley, M. O., Manning, G. S. & Olson, W. K. (1992) *J. Physical Chemistry* **96,** 3963-3969.

30. Parvin, J. D., McCormick, R. J., Sharp, P. A. & Fisher, D. E. (1995) *Nature* **373,** 724-727.

31. Kim, J. L., Nikolov, D. B. & Burley, S. K. (1993) *Nature* **365,** 520-527.

32. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. (1993) *Nature* **365,** 512-520.

33. Bond, J. P., Anderson, C. F. & Record, M. T. (1994) *Biophysical Journal* **67,** 825-836.

34. Philpott, M. R. & Glosli, J. N. (1995) *J. Electrochem. Soc.* **142,** L25-L28.

35. Ray, J. & Manning, G. S. (1994) *Langmuir* **10,** 2450-2461.