Spring 1-1-2008

Gene Regulation Through Artificially Induced DNA Conformational Changes

David Bednarski

GENE REGULATION THROUGH ARTIFICIALLY INDUCED DNA CONFORMATIONAL CHANGES

A Thesis

Submitted to the Graduate School of Pharmaceutical Sciences

Duquesne University

In partial fulfillment of the requirements for

the degree of Master of Science

Medicinal Chemistry

By

David Bednarski

May 2008

GENE REGULATION THROUGH ARTIFICIALLY INDUCED DNA

CONFORMATIONAL CHANGES

By

David Bednarski

In partial fulfillment of the requirements for the degree of Master of Science.

APPROVED: April 3, 2008

Steven Firestine, Ph. D.
Thesis Chairperson
Assistant Professor of Medicinal Chemistry
Eugene Applebaum College of Pharmacy
and Health Sciences
Wayne State University
Detroit, MI

Aleem Gangjee, Ph.D.
Professor of Medicinal Chemistry
Mylan School of Pharmacy Distinguished
Professor
Graduate School Pharmaceutical Sciences
Duquesne University
Pittsburgh, PA

Marc W. Harrold, Ph.D.
Professor of Medicinal Chemistry
Graduate School Pharmaceutical Sciences
Duquesne University
Pittsburgh, PA

David A. Johnson, Ph.D.
Associate Professor of Pharmacology-
Toxicology
Director of Graduate Studies
Graduate School Pharmaceutical Sciences
Duquesne University
Pittsburgh, PA

J. Douglas Bricker, Ph.D.
Dean, Mylan School of Pharmacy and the
Graduate School Pharmaceutical Sciences
Duquesne University
Pittsburgh, PA

ABSTRACT

GENE REGULATION THROUGH ARTIFICIALLY INDUCED DNA CONFORMATIONAL CHANGES

By

David Bednarski

May 2008

Thesis supervised by Steven Firestine, Ph. D.

The conformation of DNA has been shown to play an important role in the regulation of gene expression. A consequence of this finding is that agents that alter the conformation of DNA should also affect the regulation of gene expression. To explore this, we used tethered-triple helix oligonucleotides (TFO) to bend DNA. We show that the expression of a luciferase gene is regulated by the presence of an induced DNA bend. Bends occurring in the same orientation as RNA polymerase binding result in a 93% increase in expression. In contrast, bends induced in that opposite direction resulted in a 51% decrease in expression. These results prompted us to investigate the synthesis of three small molecules with the potential to induce a sequence selective bend in DNA. These studies revealed that the compounds were able to induce a DNA bend counter to an intrinsic bend present in a target DNA fragment.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF SCHEMES

LIST OF EQUATIONS

LIST OF ABBREVIATIONS

A – adenine

AT – adenine-thymine

bLTFO – broken long tethered-TFO

bp – base pair

bSTFO – broken short tethered-TFO

Boc – tert-butoxyxarbonylamino group

C – cytosine

CAP – catabolite activating protein

DNA – deoxyribonucleic acid

dsDNA – double stranded DNA

dsRNA – double stranded RNA

EndoV – endonuclease V

ET-743 – ecteinascidin-743

EtBr – ethidium bromide

FAM – fluorescein

FRET – fluorescence resonance energy transfer

G – guanine

GC – guanine-cytosine

HRP – horseradish peroxidase

hTGF- β_1 – human transforming growth factor

IHF – integration host factor

LEF-1 – lymphoid enhancer factor-1

LNA – locked nucleic acid

LTFO – long tethered-TFO

mRNA – messenger RNA

NF1 – nuclear factor I

PAGE – polyacrylamide gel electrophoresis

PMO – phosphorodiamidate morpholino oligonucleotide

PNA – peptide nucleic acid

RNA – ribonucleic acid

RNAi – RNA interference

siRNA – small interfering RNA

ssDNA – single stranded DNA

STFO – short tethered-TFO

T – thymine

TAMRA – tetramethylrhodamine

TFO – triple helix-forming oligonucleotide

UDG – uracil DNA glycosylase

UV – ultra-violet

VEGF – vascular endothelial growth factor

I. LITERATURE SURVEY

Gene expression is a fundamental biological process in which DNA sequences are converted to RNA and proteins crucial to the life of an organism. Cells regulate gene expression to control a multitude of functions from normal cell maintenance to apoptosis. Gene expression can be regulated at any step of the process, from transcription to protein modification. Aberrant gene expression is the result of a loss of normal regulation and is linked to a number of diseases ranging from cancer to Alzheimer's disease.[1-5] The development of new mechanisms of artificial gene regulation will lead to the creation of pharmacological agents capable of treating a variety of diseases.

A number of methods have been employed to establish artificial control over gene expression. These methods have utilized a wide range of ligands, from DNA to proteins and small molecules.[6-64] These methods seek to either inhibit or recruit the binding of proteins to DNA or mRNA to alter the expression of a target gene. An under utilized method to establish artificial gene regulation is to alter the conformation of DNA, particularly upstream of a target gene. The flexibility of DNA and the conformations that it can adopt, particularly DNA bends, play important roles in various processes including transcription. Though DNA can be intrinsically bent and can be bent by the binding of various proteins, these methods are not acceptable pharmacological treatment methods. Such treatment methods require the bending of DNA through artificial means, including nucleic acids and small molecules.

I.A. Artificial Gene Regulation

There are a number of mechanisms that have been used to establish artificial control over gene expression including antisense oligonucleotides, RNAi, zinc finger

proteins, triple helix oligonucleotides, peptide nucleic acids, and small molecules. These methods will be individually discussed below.

Antisense Oligonucleotides

Gene expression has been modified through targeting mRNA with the intent of either blocking its translation or degrading it through enzymatic action, silencing the expression of its parent gene. A number of oligonucleotides have been used to form duplexes with target mRNA to inhibit the translation of the mRNA and silence the expression of a gene. These complementary oligonucleotides have included DNA, the oligonucleotide analogue PNA, LNA and PMO.[6-10] DNA and PNA will be discussed in greater depth (see Triple Helix Oligonucleotides and Peptide Nucleic Acids sections).

LNA, locked nucleic acid, are oligonucleotide analogues of RNA. The bases of LNA contain methylene linkages between the 2' oxygen and the 4' carbon of the ribose ring. This linkage locks the 3'-carbon in the *endo* conformation, resulting in high-affinity binding and increased melting temperature.[6-7] LNA complementary to target mRNA is able to form duplexes that activate RNase H, resulting in the degradation of mRNA and the silencing of the expression of a target gene.

PMOs or phosphorodiamidate morpholino oligonucleotides are oligonucleotide analogues that display the standard nucleic acid bases. The deoxyribose rings of DNA and the ribose rings of RNA are replaced with six-membered morpholine rings while the phosphate linkers were replaced with phosphorodiamidate groups. The inclusion of the phosphorodiamidate groups eliminated the negative charge of the phosphate backbone under physiological conditions allowing PMOs to bind to their RNA targets without the

negative charge repulsion present in the formation of DNA/RNA and dsRNA duplexes. PMOs were targeted to 25-bp regions of mRNA to block the binding of the ribosomal intitiation complex, inhibiting the translation of mRNA.[8-9] In addition, PMOs can interfere with the splicing of introns from pre-processed mRNA to alter the expression of a gene.[8,10-11]

RNA Interference

Genes have been down regulated and completely silenced by interfering with the translation of mRNA through the formation of double-stranded RNA using a technique called RNA interference, RNAi.[12] RNAi methods are derived from the native degradation of RNA through the formation of double stranded RNA. These methods redirect this process to target the mRNA of a gene of interest.[12-13] Double-stranded RNA, homologous to a target sequence of mRNA, is delivered to the nucleus of a cell (Fig. 1). Riobnuclease III cleaves these dsRNA strands into 21- and 22-nucleotide fragments called small interfering RNA, siRNA. The generated siRNA bind to the target mRNA strands and create new RNA duplexes. Like the dsRNA parent of the siRNA fragments, the mRNA/siRNA duplexes are degraded by a protein complex called Dicer, silencing the expression of the originating gene. This method has been used to silence reporter genes in *Drosophila* and mammalian cell lines.[12] RNAi was also used to produce high-amylose potatoes by silencing two genes encoding for starch branching enzymes.[13]

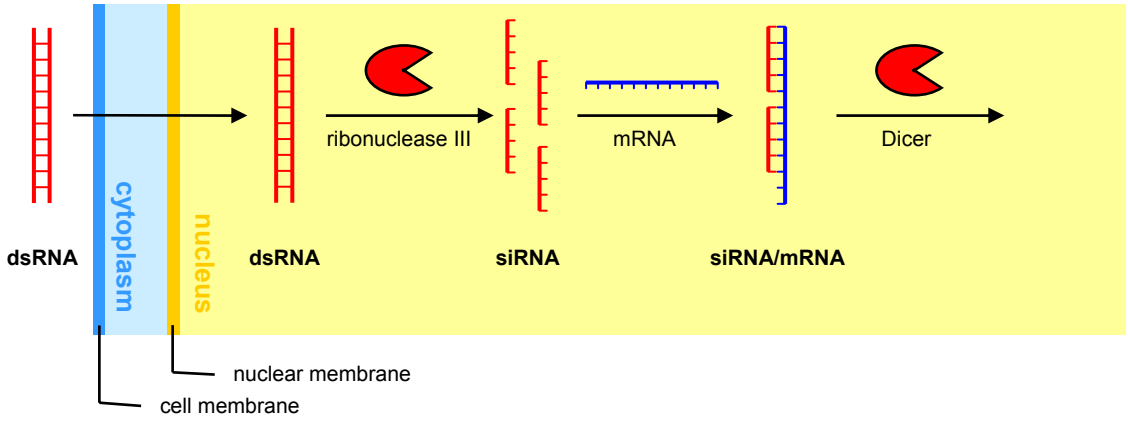


Figure 1. RNA interference occurs when dsRNA enters the nucleus of and is broken down into 21- and 22-nucleotide siRNA by ribonuclease III. siRNA then bind to target mRNA creating siRNA/mRNA duplexes that are cleaved into small fragments by ribonuclease III, preventing translation of the mRNA.

Zinc Finger Proteins

Protein-based systems have also been used to activate transcription. These activation systems were composed of two main domains, a DNA binding domain responsible for binding to a specific DNA sequence and an activation domain that recruites a transcriptional complex. These systems included both fusion proteins and independent proteins that activate transcription upon dimerization. A system composed of DNA-binding domain proteins and activation domain proteins was used to activate transcription, both *in vitro* and *in vivo*, upon dimerization by a dimerization ligand.[14] The proteins that have been used for targeting DNA sequences and for activating transcription are highly variable, though one of the most commonly used protein-based DNA-binding domains were zinc finger-proteins.[15-16]

Zinc finger-proteins are composed of C_2H_2ZF domains, an α-helix packed against two anti-parallel β-strands with stability provided by the coordination of a zinc ion by the side chains of two cysteine and two histidine residues (Fig. 2). Zinc finger proteins are able to bind to a variety of DNA sequences due to the variability of the DNA binding

region of the proteins. The N-terminus residues of the zinc finger domain α-helix make contact with the major groove of DNA at a recognition triplet, which is customizable based on the amino acids of the particular domain.[15-16] Unfortunately, zinc finger proteins cannot target all DNA sequences with equal efficiency; the zinc finger domain has difficulty targeting pyrimidines; cytosine and thymine.

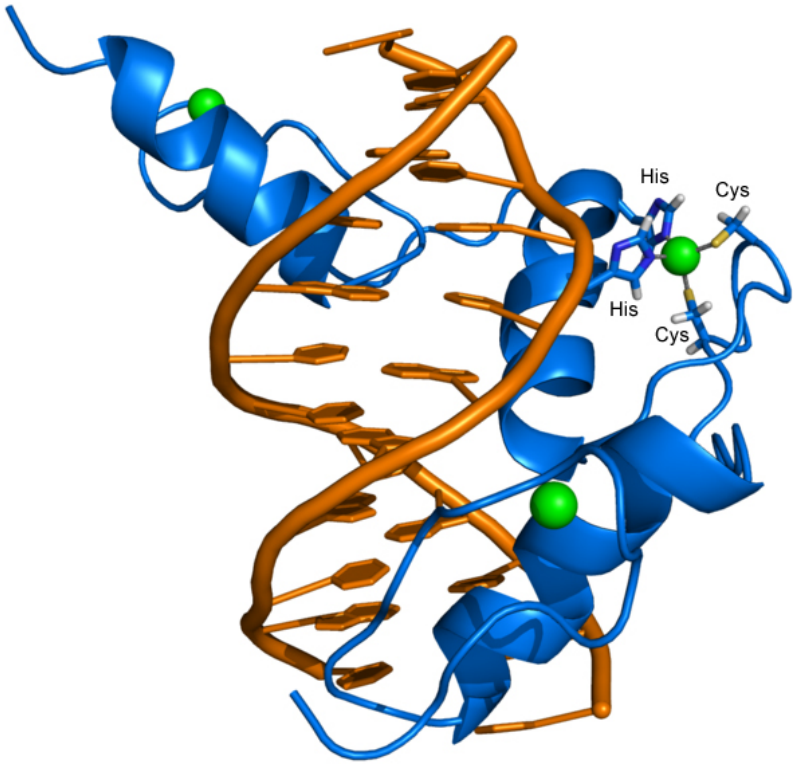


Figure 2. Polydactyl zinc fingers bound to DNA. The zinc fingers are blue, the DNA is orange and the zinc ions are green. Zinc ions are coordinated by two cysteine and two histidine residues. Figure from PDB 1A1L, ZIF268 zinc finger-DNA complex (GCAC SITE).[17]

Multiple zinc fingers domains have been linked to form polydactyl zinc finger proteins to target longer, more specific DNA targets (Fig. 2). Polydactyl zinc finger-proteins have been used as DNA-binding domains for a number of gene activation motifs with activation domain proteins such as TBP and variations of the herpes simplex virus VP16 activation domain.[18-19] These fusion proteins were used to activate the expression

of oncogenic genes ErbB-2 and ErbB-3, angiogenic genes VEGF-A and CD144, apoptotic gene BAX, and γ-globin as a potential treatment for sickle cell anemia.[15-16,20-21] When conjugated to KRAB (Krüppel associated box), ERD, ERF repressor domain, or SID (mSIN3 interaction domain), zinc finger-proteins served as binding partners for repressor domains.[19]

Triple Helix Oligonucleotides

TFOs are single strands of DNA that are able to bind in the major groove of DNA through hydrogen bonds with the purine members of DNA base pairs, though they can tolerate some pyrimidines in their binding region. There are two modes for this hydrogen bonding to occur, parallel and antiparallel, dependent on the orientation of a TFO in the major groove.[22-24] The purine-rich DNA strand serves as a reference when determining the binding mode of the TFO. Parallel binding occurs when the TFO and a purine-rich DNA strand are in parallel, aligned with their 5' and 3' ends in the same orientation. The hydrogen bonding that occurs in the parallel binding orientation is called Hoogsteen base pairing, or the pyrimidine motif (Fig. 3A). TFO thymines bind to duplex adenines, while protonated cytosines of TFOs bind to duplex guanines. Due to TFO cytosines requiring protonation to bind, triplex formation in parallel systems is pH dependent, favoring acidic environments.

The other TFO binding motif that a TFO can bind into duplex DNA is one with an antiparallel orientation. A TFO and a purine-rich DNA strand of a DNA duplex run opposite of one another in terms of 5'-3' directionality. The hydrogen bonds formed between TFOs and antiparallel DNA are known as Reverse Hoogsteen or the purine

motif (Fig. 3B). Reverse Hoogsteen base pairing is comprised of purines hydrogen bonding to purines; adenine binds to adenine and guanine binds to guanine. The antiparallel orientation can also contain thymines of a TFO strand bound to the adenines of the purine-rich duplex DNA strand.[25] Unlike regular Hoogsteen base pairing, Reverse Hoogsteen bonds do not have a pH requirement.

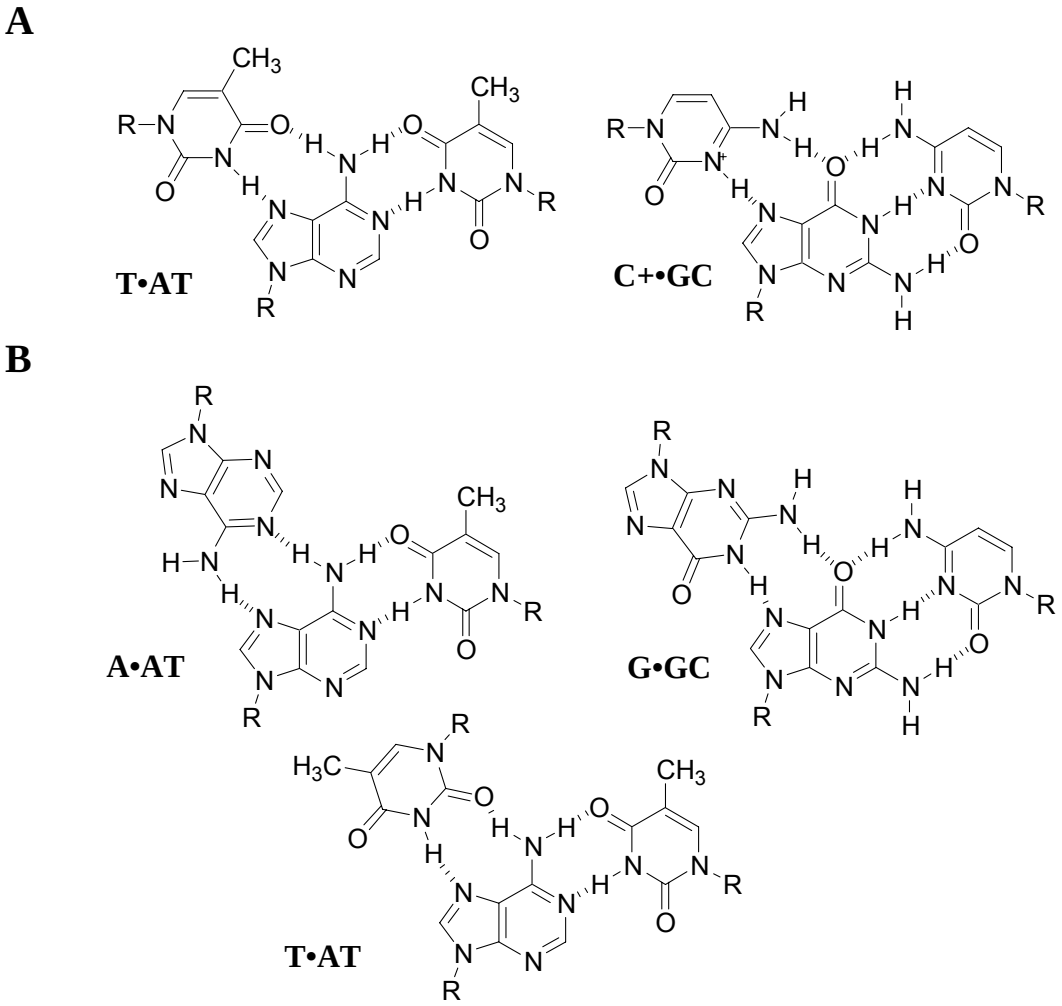


Figure 3. (**A**) Hoogsteen base pairs used in the formation of DNA triplexes; thymine to a Watson-Crick base paired adenine and thymine and a protonated cytosine to a Watson-Crick base paired guanine and cytosine. (**B**) The Reverse Hoogsteen hydrogen bonding motif found in triplex DNA.

Single TFOs have been used to both inhibit and activate transcription. TFOs have been used as antigene inhibitors that bind to target DNA sequences to inhibit gene

expression, most often blocking transcriptional initiation.[26] The expression of c-*myc* genes, rat α1(I) genes and oncogenic Ets2 genes were inhibited by binding TFOs to their respective promoters.[27-30] TFOs have been used to block specific DNA-protein interactions integral to transcriptional activation. TFOs inhibited gene expression when bound to the binding sites of Sp1 and SPy transcriptional factors in the promoter regions of Ha-*ras* and c-Src genes.[31-32] In addition to preventing transcriptional activation, TFOs were used to inhibit transcriptional elongation by arresting the progression of bacterial RNA polymerase in *in vitro* experiments through the binding of TFOs to sites within a gene sequence.[33]

Transcription was also inhibited by TFOs that mediated genomic modifications. DNA-damaging agents, such as psoralen, were coupled to TFOs to achieve site-directed DNA damage.[34] Psoralens intercalated cellular DNA to form monoadducts and double strand crosslinks upon activation by UV irradiation.[35-36] TFO/psoralen conjugates have been used to selectively inactivate transcription in skin cells as a treatment for psorosis.[35] Duplex DNA was distorted upon formation of either monoadducts or crosslinks by psoralen. These distortions are recognized and repaired by nuclear excision repair mechanisms that leads to strand breaks and recombination that eliminated and activated gene expression.[36-37]

TFOs also activated transcription by acting as the DNA-binding domain for protein and oligonucleotide activation domains. TFOs were conjugated to hairpin DNA that contained transcription factor binding sites.[38] Transcriptional machinery will be recruited to the TFO/hairpin-DNA conjugate, when bound to DNA, initiating transcription.

As guanine-rich oligonucleotides, TFO were prone to self-association into quartets and tetrads in physiological concentrations of potassium.[34] TFOs were also limited by negative charge repulsion between the phosphate backbones of the duplex DNA and the TFO. Replacement of three or four adjacent 2'-*O*-methyl sugars of the TFO backbone with 2'-*O*-aminoethyl riboses alleviates this repulsion and stabilizes the triplex.[39]

Peptide Nucleic Acids

PNAs are oligonucleotide analogues in which the sugar phosphate group is replaced with a uncharged *N*-(2-aminoethyl)glycine unit (Fig. 4A).[40-42] Peptide nucleic acids have also been used to artificially regulate gene expression. Unlike TFOs, PNAs were not limited by negative charge repulsion with the backbones of target DNA. PNAs are uncharged under physiological conditions and are resistant to nucleases and proteases, making them a more stable alternative to TFOs for targeting transcription to alter transcription.

PNA binds to DNA by forming both Watson-Crick and Hoogsteen base pairs. PNAs can bind to dsDNA by displacing the second strand of DNA (Fig. 4B). A second strand of PNA can bind to a PNA/DNA duplex to form a triplex. The displaced DNA strand forms a single-stranded displacement loop (D-loop) recognized by RNA polymerase as an initiation site comparable to a strong promoter.[43] PNAs have been targeted to sites where transcriptional activation was desired, such as upstream of γ-globin genes to treat sickle cell anemia.[40]

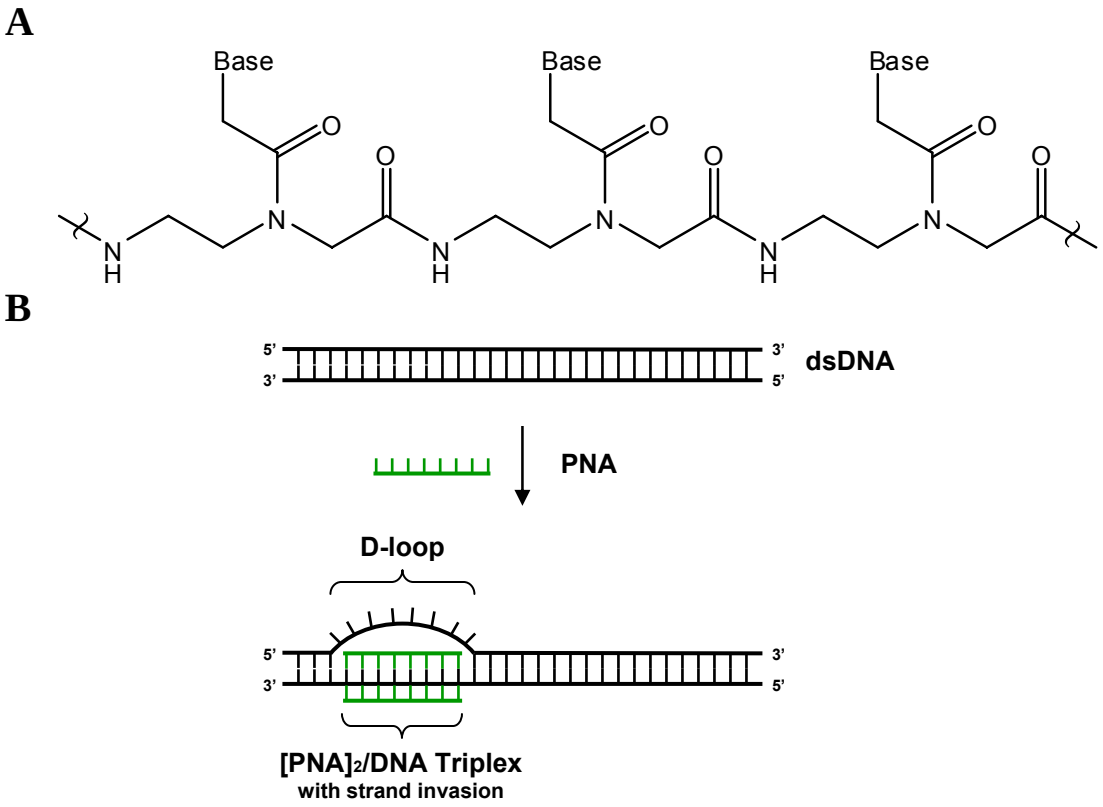


Figure 4. (**A**) PNA backbone consisting of uncharged *N*-(2-aminoethyl)glycine units linked by amide bonds. (**B**) PNA binds to DNA forming triplexes that displace one strand of DNA forming a D-loop.

Additionally, PNAs have been used to inhibit transcription. PNAs were used in an antigene mechanism by binding to the ssDNA of open transcription complexes in chromosomal DNA, arresting transcription by inhibiting elongation.[44] Antigene PNA was used to down regulate the mutant *KRAS* gene, with minimal effect to wild-type *KRAS*.[41] Like TFOs, PNA has been used to block the binding sites of proteins required for gene activation.[42] PNAs were also used to deliver DNA damaging agents, such as psoralen, to a particular sequence to induce DNA repair and silence a gene.[36]

While lacking the negative charge repulsion limitation of TFOs and other oligonucleotides, PNAs are limited in their ability to penetrate cell membranes. However, conjugations of PNA with cell penetrating peptides and nuclear localization

peptides have been shown to ease the passage of PNA through cell and nuclear membranes.[41,45]

Small Molecules

There are a number of small molecules that have been used to control gene expression by targeting transcription by targeting transcriptional factors and DNA to inhibit or activate transcription. These will be discussed below.

Small molecules have been used to inhibit transcription by targeting transcription factors. An example of small molecules that functioned in this capacity was β-peptide peptidomimetics. These peptide backbone analogues inhibited the interaction between p53 and hMD2.[46-47] The hMD2 protein negatively regulates p53, a tumor suppressor transcription factor. β-peptide mimics of p53 bound to hMD2, allowing p53 to activate the transcription of genes that suppressed tumor growth and genes that promoted cell cycle arrest and apoptosis. Similarly, small molecules were used to inhibit the interaction of p53 with MDM2, which is overexpressed in tumor cells and inhibited the activity of p53 to suppress tumor growth.[47]

Small molecules can also target DNA to affect transcription and artificially regulate gene expression. Most compounds that target DNA bind to the minor groove and display a preference for AT-rich DNA sequences. This selectivity is largely due to the steric hindrance presented by the exocyclic amine of the guanine nucleotide in the minor groove. Polyamides netropsin and distamycin (Fig. 5) are naturally occurring antibiotics that served as the parent compounds of a series of sequence selective minor groove binding agents. Netropsin and distamycin displayed a preference for binding to

AT-rich regions of DNA. The binding energy of these compounds is derived from hydrogen bonds formed between the amide-NH groups of the compounds and the hydrogen bond acceptors in the minor groove. Additional binding energy was provided by electrostatic interactions between the positively-charged amindine tails of the compounds and the negatively-charged phosphate oxygens of the DNA backbone. Binding affinity is also enhanced by the hydrophobic interactions between the aromatic rings of the compound and the carbons of the ribose ring of DNA.

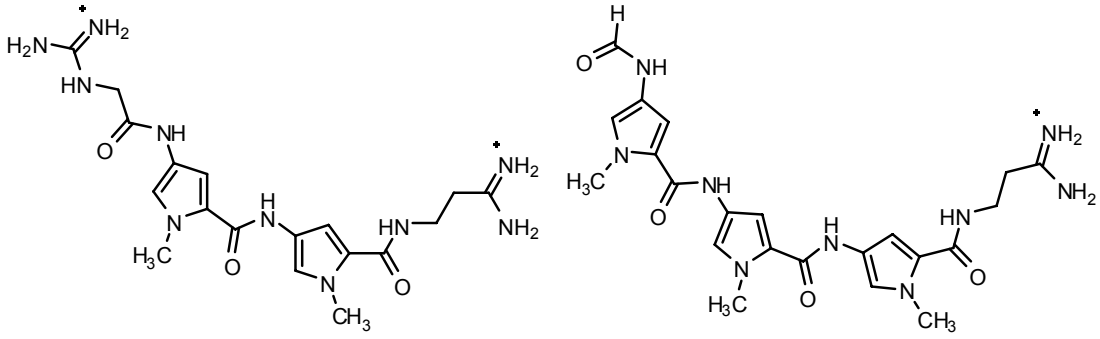


Figure 5. Polyamides netropsin and distamycin.

Polyamides, represented by distamycin, were able to bind to the DNA minor groove in two different binding modes based upon the stoichiometry of binding either 1:1 or 2:1. The first binding mode, 1:1, consisted of a single distamycin molecule bound in the minor groove of an AT-rich region of DNA (Fig. 6A). The 1:1 binding motif is the predominant binding mode at lower concentrations of the antibiotic. Distamycin can bind in either orientation (Fig. 6A).[48]

At higher concentrations of distamycin the second binding mode, 2:1, becomes the prevalent binding motif.[49] The 2:1 binding mode involved two distamycin molecules binding to the same AT-rich DNA minor groove site (Fig. 6B). The minor groove is widened to accommodate the binding of two molecules. The binding of the first

distamycin lowers the binding energy for the second distamycin and increases the second distamycin's rate of binding.[50] The 2:1 binding mode is stabilized by dipole-dipole interactions between the pyrrole rings and the amide bonds of the molecules. In this binding mode each distamycin is bound to a single strand of DNA, preventing bifurcated hydrogen bonds, which in turn increased the selectivity of the duplex toward AT-rich sequences in the minor groove.

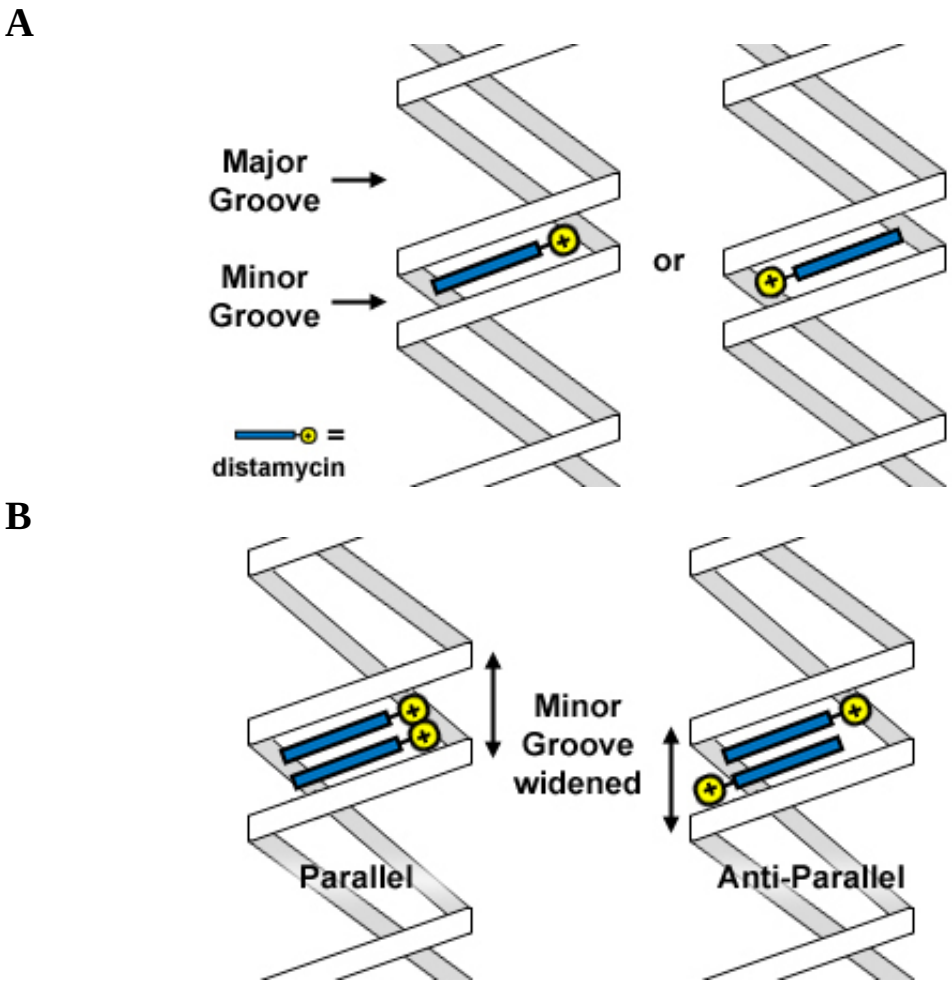


Figure 6. Two binding modes seen in NMR studies of distamycin-DNA binding. The (**A**) 1:1 binding mode and (**B**) 2:1 binding mode, distamycin to DNA. There are two orientations possible for the 2:1 binding mode, parallel and anti-parallel, both of which result in a widening of the minor groove. Parallel binding has not been experimentally observed.

The 2:1 binding mode could occur in two different orientations; parallel and anti-parallel (Fig. 6B). The parallel orientation involved the stacking of two distamycin molecules such that their positively-charged amidine tails lay in the same direction. Due to charge-charge repulsion of the amidine tails, parallel 2:1 binding does not occur. The anti-parallel 2:1 binding mode, however, has been experimentally observed and is the most common higher order structure formed. In the antiparallel binding mode, the distamycin molecules were stacked with their positively-charged tails at opposite ends of the binding motif, such that there is no charge-charge interaction between them.

The 2:1 binding mode of distamycin allows both strands of a DNA sequence to be targeted in the minor groove in a sequence selective manner. Dervan used distamycin as a lead compound in the creation of sequence selective minor groove binding

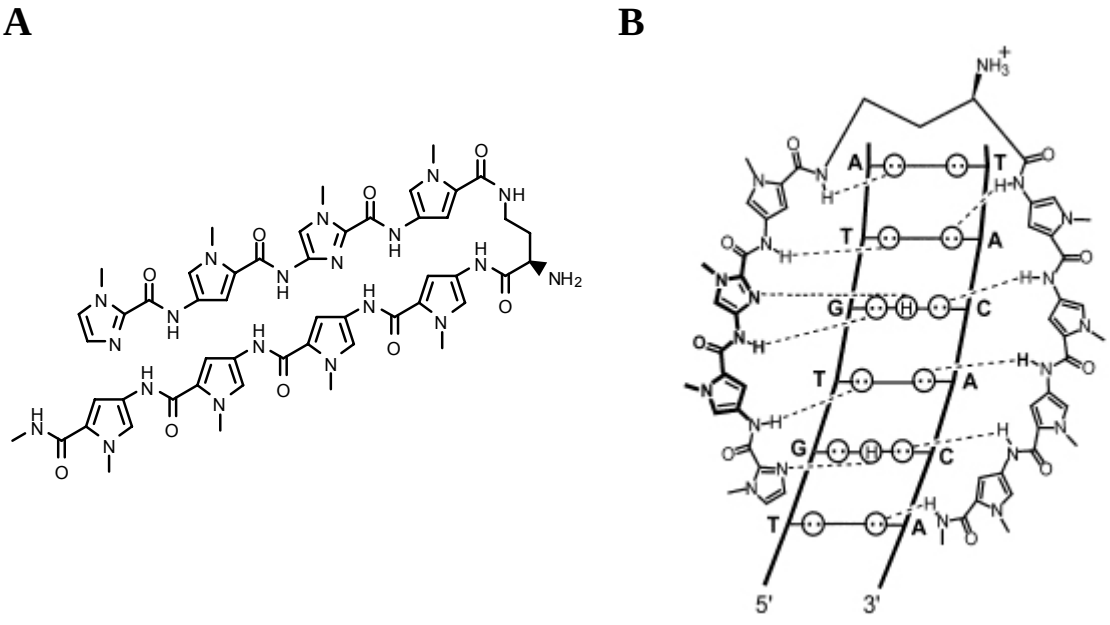


Figure 7. (**A**) Example of Dervan's hairpin polyamide.[52] (**B**) Hairpin polyamide from **A** bound to the minor groove of the DNA. DNA is shown from a perspective of looking at the base pairs from their edge, with circled H's representing hydrogen bond donors while circles with two dots, lone pair electrons, represent hydrogen bond acceptors.[52]

polyamides.[51] Dervan's polyamide minor groove binders utilized various heterocycles linked by amide bonds. These polyamides, called lexitropsins, bound to DNA in the anti-parallel 2:1 binding mode. Two lexitropsin molecules were then linked at one end with an aliphatic amino acid to create a hairpin polyamide (Fig. 7A). The side-by-side binding of these polyamide chains, with each chain targeting one strand of DNA, allowed this system to target all four base pair arrangements in any sequence (Fig. 7B).

The hairpin polyamide system utilized three different heterocyclic systems to recognize DNA; pyrrole (Py), imidazole (Im) and hydroxypyrrole (Hp) (Fig. 8). The *N*-methylpyrroles of the lead compound, distamycin, could target the N3 of adenine and the O2 of thymine in a bifurcated fashion with the connecting amide-NH groups. In the 2:1 binding motif, where only a single strand of DNA would be targeted by the moiety, either adenine or thymine could be targeted by an *N*-methylpyrrole (Fig. 8). *N*-methylpyrroles could also accommodate the O2 of cytosine, normally inaccessible in the 1:1 binding motif due to steric hindrance of the exocyclic amine of guanine in the minor groove. Hairpin polyamides could therefore use *N*-methylpyrroles to target adenine, thymine or cytosine nucleotides.

Recognition of guanine presented a problem due to the protrusion of the exocyclic amine group into the minor groove. Guanine targeting was accomplished by the incorporation of imidazoles (Im) into the molecule. Imidazoles act as hydrogen bond acceptors to the exocyclic (Fig. 8B). The ability to target guanine allowed for the creation of hairpin polyamides capable of distinguishing between GC and CG base pairs with a combination of pyrrole and imidazole; Im/Py could target GC, while Py/Im could target CG base pairs.

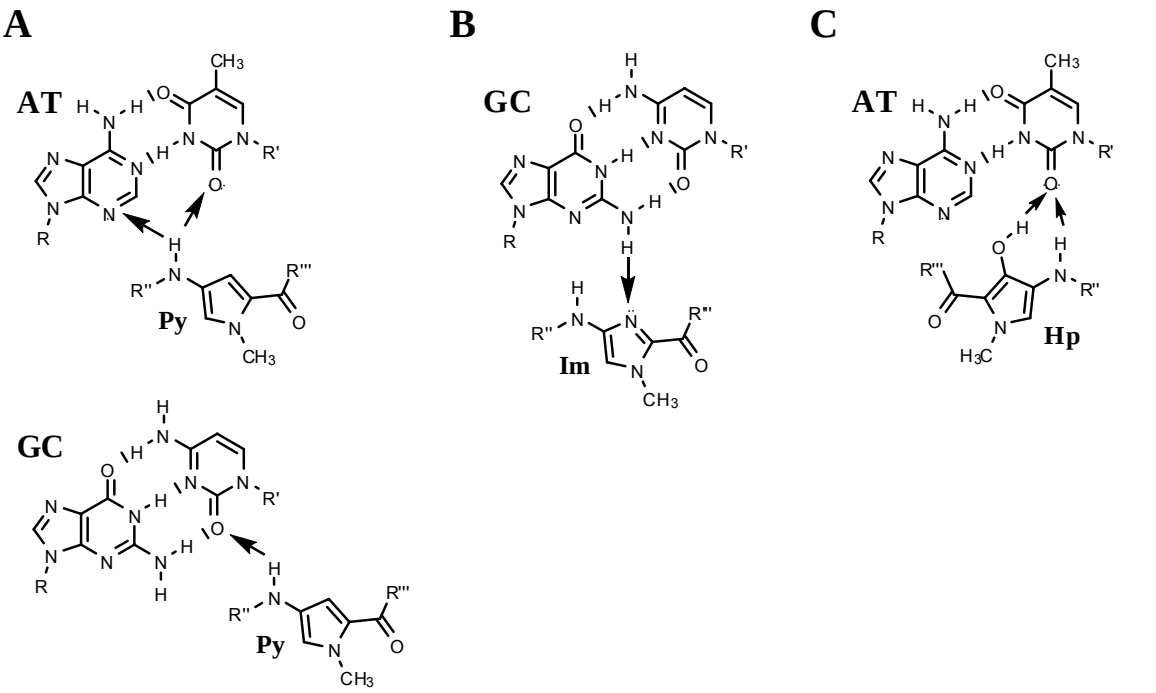


Figure 8. Minor groove binding schemes of heterocycles used in Dervan's polyamides in AT and GC base pairs; (**A**) pyrrole, Py, (**B**) imidazole, Im, and (**C**) 3-hydroxypyrrole, Hp.

Dervan's hairpin polyamides also possessed the ability to distinguish between AT and TA base pairs. While paired pyrroles could target both adenine and thymine, there was no distinction between AT and TA. By pairing an *N*-methylpyrrole with a 3-hydroxypyrrole (Hp) in a hairpin polyamide AT and TA could be differentiated and independently targeted. *N*-methylpyrroles cause a shift in the hydrogen bonding scheme of a polyamide depending on whether the N3 of an adenine or the O2 of a thymine is targeted. 3-Hydroxypyrrole alleviated this shifting by displaying a preference for thymine due to the presence of the hydroxy group.[53-54] This hydroxy group also may hydrogen bond with the O2 of thymine (Fig. 8C).[54] Combinations of *N*-methylpyrrole and 3-hydroxypyrrole in hairpin polyamides allowed for the independent targeting of AT and TA base pairs, Py/Hp and Hp/Py, respectively.

Hairpin polyamides are able to target AT, TA, GC and CG base pairs independently (Table I). This allowed molecules to be created to specifically target any sequence in DNA. In addition to *N*-methylpyrrole, 3-hydroxpyrrole and imidazole a number of other heterocycles have been placed into polyamide chains including thiazoles, furans, oxazoles and benzimidazoles, among many others.[54-55]

Table I: The sequence recognition of hairpin polyamides by heterocylic pair, where "+" indicates recognition and "–" indicates no recognition.

Pair	AT	TA	GC	CG
Py/Hp	+	-	-	-
Hp/Py	-	+	-	-
Im/Py	-	-	+	-
Py/Im	-	-	-	+

Polyamides are the only class of small molecules capable of selectively targeting any predetermined DNA sequence and have been used to both inhibit and activate transcription in a sequence-specific manner.[56] Polyamides were used to block the binding sites of transcription factors to inhibit the expression of adjacent genes by preventing crucial protein-DNA interactions.[56-60] Polyamides were used in this fashion to block the binding of TFIIIA to inhibit the expression of the 5S ribosomal gene *in vitro*.[57-58] Gene expression was inhibited in other instances where polyamides were used to block the binding sites of transcription factors such as LEF-1, Ets-1.[56] The transcription of other genes, bHLH and hTGF- β_1, were inhibited by binding polyamides in the respective promoter regions.[56,59] Polyamides were also bound to the hypoxia response element to prevent the binding of hypoxia-induced factor to reduce the expression of VEGF.[60]

The customizability of polyamides to target any sequence of DNA permitted polyamides to serve as the DNA-binding domains of dual-domain transcriptional

activators.[56,61-62] The activation domains of these activator complexes included both peptide ligands and other small molecules. Hairpin polyamides tethered to activating peptides derived from the herpes simplex virus VP16 activator, were used to stimulate promoter specific transcription.[62]

Peptide-based activation domains were replaced with small molecules to create completely synthetic small molecule transcriptional activators.[63-64] Hairpin polyamides fused to isoxazolidines mimics of activation domains, were used as artificial transcription factors (Fig. 9).[63] Fusions of hairpin polyamides to a molecule called wrenchnolol initiated transcription in a sequence-specific manner.[64]

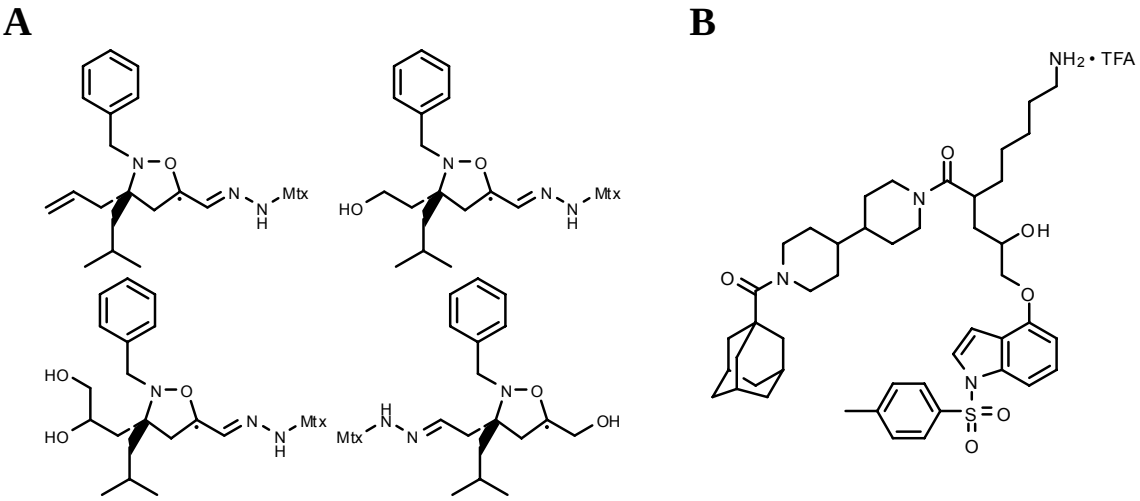


Figure 9. The structures of (**A**) isoxazolidines and (**B**) wrenchnolol.

I.B. DNA Bending

One method which has not received much attention in the artificial regulation of gene expression is the role that DNA conformation plays in gene regulation. The common structural perception of DNA is that it exists as a rigid, rod-like, double helix. DNA, however, is not a rigid molecule. DNA possesses a level of flexibility critical to numerous cellular activities. The flexibility of DNA is derived from the potential

deformability of its structure, specifically the structures and interactions of the nucleotide bases (Fig. 10).[65-67] This flexibility is best described in terms of "inverse stiffness" and is expressed as a persistence length, the distance over which DNA is effectively linear.[68-70] The persistence length is dependent on both the rigidity of DNA and sequence-derived, intrinsic DNA curvature. DNA fragments shorter than 150 base pairs are considered to have rigid, linear character, while DNA sequences longer than 150 base pairs possessed more flexibility and have slightly curved structures.[69] The flexibility of DNA allow for protein interactions necessary for the various activities of DNA, ranging from packaging in nucleosomes to transcription.

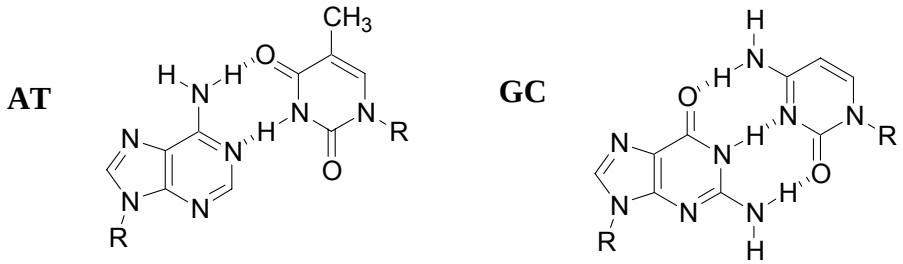


Figure 10. Watson-Crick base pairs; adenine, A, to thymine, T, and guanine, G, to cytosine, C.

The innate flexibility of DNA allows it to adopt various conformations, distorting it from the aforementioned rod-like structure. These conformations range from simple bends to full loops.[71] These distortions are the result of the widening of one of the grooves of a DNA region and compressing the complementary groove on the opposite face of the double helix. These alterations in groove width are introduced into a DNA sequence by one of two general mechanisms, either nucleotide sequence dependent intrinsic bends or bends facilitated by an external force, such as a protein, cation or drug.[69,72-75]

Intrinsic conformation changes are the result of DNA sequences that possess a higher level of flexibility than other sequences. The most common examples of these sequences are AT-rich DNA sequences.[73] These sequences are found throughout the genome and interact with a variety of proteins including transcriptional proteins.[76-77] Static bends of DNA are derived from steric interactions between base pairs of a particular sequence. The best known example of an intrinsically bent DNA sequence is the A-tract. A-tracts consist of a sequence of four to six adjacent adenine-thymine base pairs that naturally bends the DNA toward the minor groove. The magnitude of the bend is about 18° from linear DNA (as measured from the center of the helical axis) and is caused by widening of the major groove as a result of a combination of tilt and roll helical deformations between the adenine-thymine base pairs (Fig. 11).[72,74,78-80] These A-tracts are found within the genome and have roles in various DNA processes, including replication and transcription, either by the curve that they possess or by serving as a binding site for proteins that bind to pre-bent DNA.[67,72,81-85]

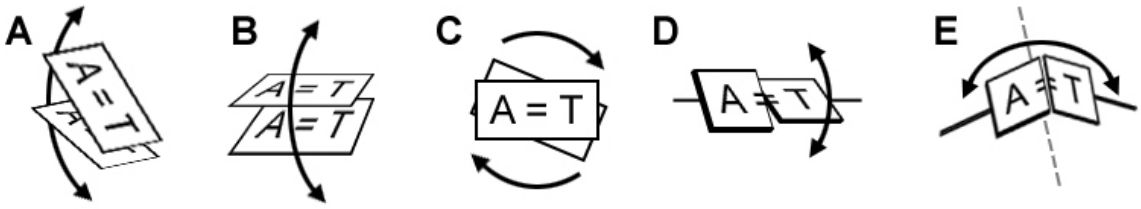


Figure 11. The helical parameters of DNA are influenced by the sequence of a region. These are the basic helical deformations that can be caused by a base pair or between base pairs as a result of steric interaction. Tilt (**A**) results in a change in the short axis between base pairs, while a roll (**B**) results in a change along the long axis between two base pairs. A twist (**C**) deformation results in adjacent base pairs twisting about the axis of a DNA strand in relation to one another. A propeller twist (**D**) involves a twist within a base pair, resulting in a less planar base pair and widening of the major groove. A buckle (**E**) is a base pair deformation that also results in a less planar base pair; however this change is due to a bend in the short axis between the nucleotides.

While DNA conformation is affected by the sequence of nucleotides, DNA conformation is also naturally altered by external forces, namely interactions with DNA-bending proteins. These proteins included those that are architectural in nature serving only to bend the DNA, and those which possess other major functions, in addition to bending DNA. The methods by which these proteins bend DNA are varied, but can be divided into two basic classes (Fig. 12); proteins that contact bent DNA on the convex side of the bend, bending DNA away from their binding position, and proteins that contact the concave side of DNA bends, bending the DNA around itself.[69,75]

The first class of DNA bending proteins, known as convex benders, bent DNA through hydrophobic interactions between amino acids of the protein and base pairs in the minor groove of the DNA. The presence of the amino acids in the minor groove, and the resulting interactions, widened the minor groove, compressed the major groove and bent the DNA away from the protein. A number of proteins functioned by this method including transcriptional factors such as the eukaryotic TATA binding protein (TBP) (Fig. 12).

The second class of DNA bending proteins, concave benders, functioned by bending DNA toward or around themselves. Many of these concave benders electrostatically interact with the DNA.[86] Examples of proteins that bend DNA from the concave side of the bend included the eukaryotic histone octamer, the prokaryotic integration host factor (IHF) and the prokaryotic catabolite-activating protein (CAP) (Fig. 12).

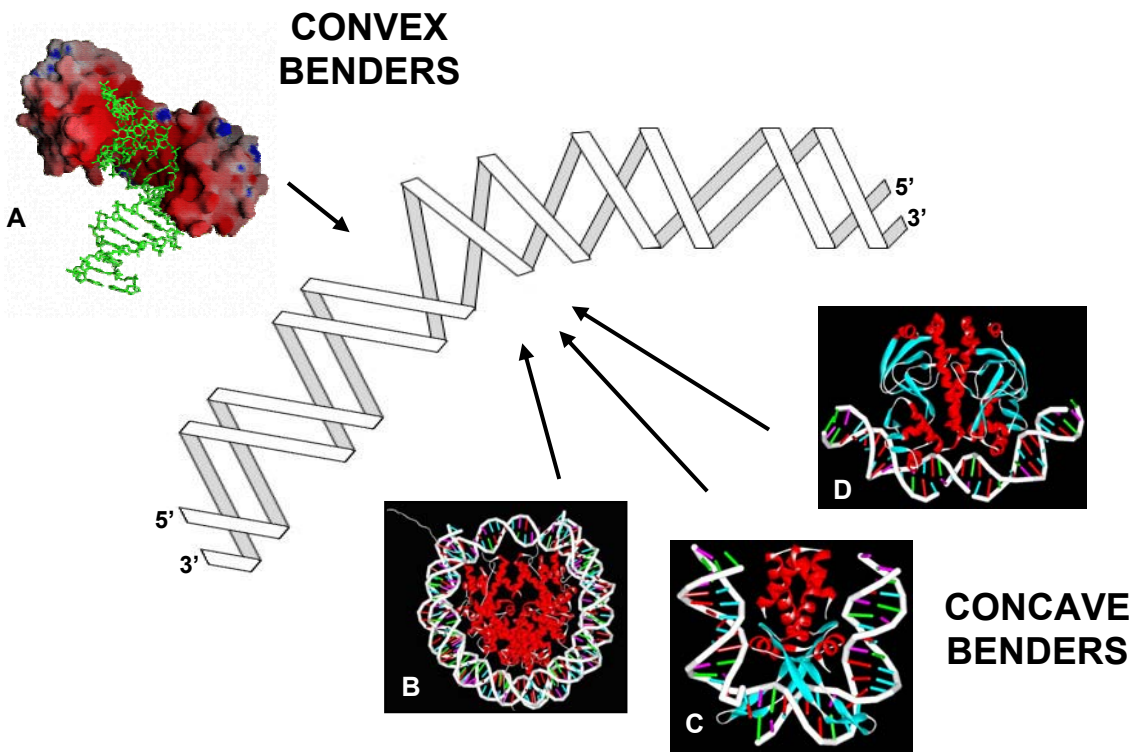


Figure 12. Examples of DNA bending proteins. Protein that bends from the convex side of the induced bend TBP, **A**. Proteins that bend DNA from the concave side of the induced bend include nucleosome formation about a histone octamer, **B**; IHF, **C**; and CAP, **D**.

I.C. DNA Bending and Transcription

I.C.1. *Potential mechanisms for the role of DNA bending in transcription*

The focus of this thesis is the crucial role DNA bending plays in transcriptional regulation. Various studies have shown that DNA bending participates in transcription in two main capacities. First, RNA polymerase causes DNA bending upon binding to the promoter sequence of a gene. Second, the binding of transcription factors in the promoter region of a gene also results in bent DNA, although it is possible that the binding sites of some transcription factors are already pre-bent. Pre-bent DNA sequences may function as a mechanism to attract the transcription factors to bind to their sites.[67,85]

While it is generally agreed that DNA bending upstream in the promoter region of a gene is important to the regulation of genetic expression, the mechanism of how bending regulates gene expression remains in question.[77,85] There are two basic theories regarding the role played by DNA bends on gene regulation. The first theory suggests that DNA bends function in an architectural capacity, bringing distally bound proteins into proximity with the RNA polymerase (Fig. 13A), while the second theory suggests that DNA bending facilitates DNA wrapping about RNA polymerase (Fig. 13B).

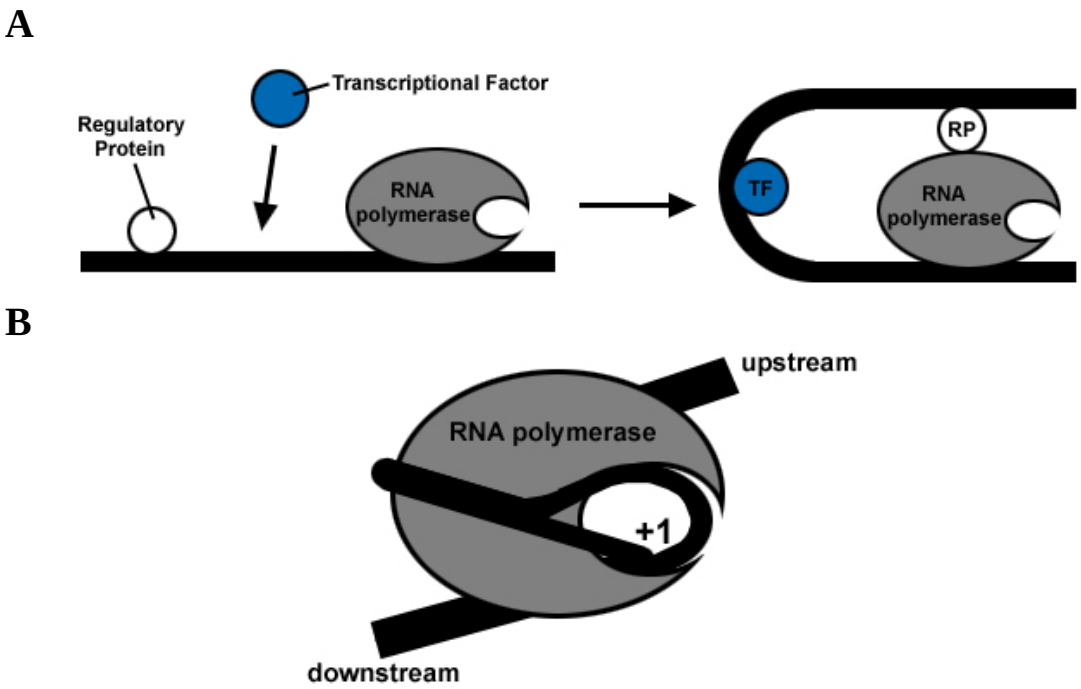


Figure 13. The two models of DNA bending and transcriptional regulation. (**A**) Interactions between RNA polymerase and distally bound transcription factors brought into proximity by DNA bend and (**B**) DNA wrapping facilitated by DNA bending.

The initiation of transcription requires the assembly of multiple protein complexes. The proteins involved in these complexes may be bound at nonadjacent DNA sites and must be brought into contact for the formation of the transcriptal initiation complex.[23,82-83,87-93] For example, transcription factors TFIIA, TFIIB, TFIIC, TFIID

(also known as TBP), TFIIE, TFIIF and TFIIH and RNA polymerase in eukaryotic systems all most interact to form the transcriptional complex. The control of expression also involves interactions of the transcriptional complex and RNA polymerase with proteins bound upstream such as activators, enhancers and repressors. Interactions between upstream and downstream proteins are facilitated through the distortion of the intervening DNA helix, bringing the proteins of interest into proximity (Fig. 13A). In the case of longer intervening DNA sequences, greater than 500 base pairs, the natural flexibility of the DNA will allow bending to occur and the proteins will be brought into the required proximity. However, for shorter distances, DNA is more rigid and requires the binding of DNA-bending proteins to facilitate the formation of protein-protein interactions necessary for the completion of the transcription initiation complex. Thus, DNA bending regulates transcription by facilitating the interaction of RNA polymerase and distally bound proteins.

The second theory regarding the role of upstream DNA bending and gene regulation asserts that bending is required for wrapping DNA about RNA polymerase in both prokaryotic and eukaryotic transcriptional complexes.[83,94-96] DNA wrapping has been visualized by footprinting, DNA supercoiling experiments, microscopy and protein-DNA crosslinking analysis.[96] DNA becomes wrapped around RNA polymerase upon its binding to the promoter region of a gene. RNA polymerase comes into contact with DNA both upstream and downstream of the transcriptional start site (Fig. 13B). DNA wrapping facilitates strand separation, required for transcription, by introducing torsional strain in the double helix.

Wrapping of DNA around RNA polymerase requires about 60 kJ/mol of energy to alter the conformation of DNA.[96] The presence of other DNA bending proteins, such as TATA-binding protein (TBP) in eukaryotic transcription complexes, lessens the cost of wrapping DNA around RNA polymerase.[95,97] Upon binding, TBP induces an 80° bend toward the major groove of DNA prior to RNA polymerase DNA wrapping. Pre-bent DNA lowering the cost of DNA wrapping is also seen in prokaryotic systems, where RNA polymerase binds to promoter regions that contain intrinsically bent DNA.[83]

While the exact role that DNA bends play in transcriptional regulation is unknown, their position, subsequent directionality and intensity are all important factors in the transcriptional regulation of a gene. The orientation of a DNA bend, relative to the promoter and RNA polymerase binding position, has an enormous effect on the transcription of a gene. By changing the position of a facilitated bend, its location and orientation relative to a gene promoter is also changed. The orientation of a facilitated bends, such as those induced by Sox2, may be the result of single nucleotide changes in the binding site consensus sequences of DNA bending transcription factors.[93] Changing the orientation of a facilitated bend relative to a promoter will prevent optimal interactions between distally bound proteins and the transcription machinery resulting in decreased transcription levels.[93,98] This phenomenon is explored later, in the experimental validation of the role of DNA bending in transcription section.

In addition to the position and orientation of a DNA bend, the degree of a DNA bend plays an important part in determining gene transcription levels. TBP binds to and induces DNA bends in promoters that contain the TATA box consensus sequence, while the SRY-related Sox2 protein binds to HMG box sequences. The bend induced by TBP

at these binding sites varies in intensity, ranging from 30° to 106°.[99-100] Sox2 induces an 83° bend in its binding site, while a mutation in the protein resulted in a bend angle of 46°.[93] These studies indicate that there is a direct correlation between bend angle and transcription. In addition, these studies indicated that the presence of an optimal bend angle for transcription while bends that were too small or too large resulted in decreased transcription levels.[93,99-100] The induction of a bend by TBP at the TATA box was described by a two state model in which transcription was not activated by TBP while a DNA bend was only slightly induced, but upon the induction of an 80° DNA bend, transcription became activated, demonstrating that transcriptional activation requires the induction of a DNA bend of a certain magnitude.[100]

I.C.2. *Experimental validation of the role of DNA bending in transcription*

To analyze the role of protein-induced DNA bends in transcriptional regulation, a number of replacement studies were preformed. In these studies, the binding sites of various DNA-bending proteins were replaced with either intrinsically bent DNA or binding sites for heterologous DNA-bending proteins. Gartenberg and Crothers replaced the binding site for the DNA-bending protein CAP, located upstream of a *lac* promoter, with phased A-tracts *in vitro*.[101] Transcription was up-regulated by the presence of the A-tract. It was hypothesized that this up-regulation was due to the A-tract curve mimicking the bend induced by normal CAP binding. When the curvature of the A-tract was directed towards the opposite face of the helix, in opposition to a normal CAP-induced DNA bend, transcription was inhibited indicating that the orientation of a DNA bend upstream of a gene could influence the expression levels of that gene. Similar studies had been performed on the *lac* operon *in vivo* by replacing the binding site for

DNA-bending CAP with various A-tract sequence combinations either in-phase (bend occurs in the same direction as CAP) or out-of-phase (opposite direction compared to CAP bending) with the start of the target gene.[102] The results were the same as those seen by Gartenberg and Crothers in their *in vitro* experiments; the A-tracts phased as the natural CAP binding site up regulated the expression of the encoded gene, while the A-tracts bending the DNA out-of-phase with the normal CAP bend displayed down regulated gene expression. These studies also demonstrated the role of DNA bending and its importance in the regulation of genetic transcription through the effects of changing the phase of the bend. They illustrate the importance of the three-dimensional orientation of DNA bends its role in gene regulation.

The down-regulation and total inhibition of expression demonstrated by out-of-phase DNA bends can be explained by both theories on the roles of DNA bending in transcription. In terms of DNA wrapping, bending the DNA in the opposite direction of that which facilitates transcription prevents the DNA from interacting with the RNA polymerase in a manner to induce DNA wrapping. Bending the DNA opposite that of transcriptional activation would also prevent distally bound proteins from being brought into proximity for interaction at the promoter.

Later work revealed that A-tracts, while intrinsically bent, were also a binding site for the sigma subunit of RNA polymerase, bringing into question the actual effects of the intrinsic bend of an A-tract on transcription in both *in vitro* and *in vivo* environments.[103] The authors suggested that the effects on transcription were due to the placement of a new RNA polymerase binding site into the promoter. However, the fact that the activity of promoter was dependent upon the relative orientation of the A-tract bend suggests that

other factors may be involved besides the inclusion of a new RNA polymerase binding site.

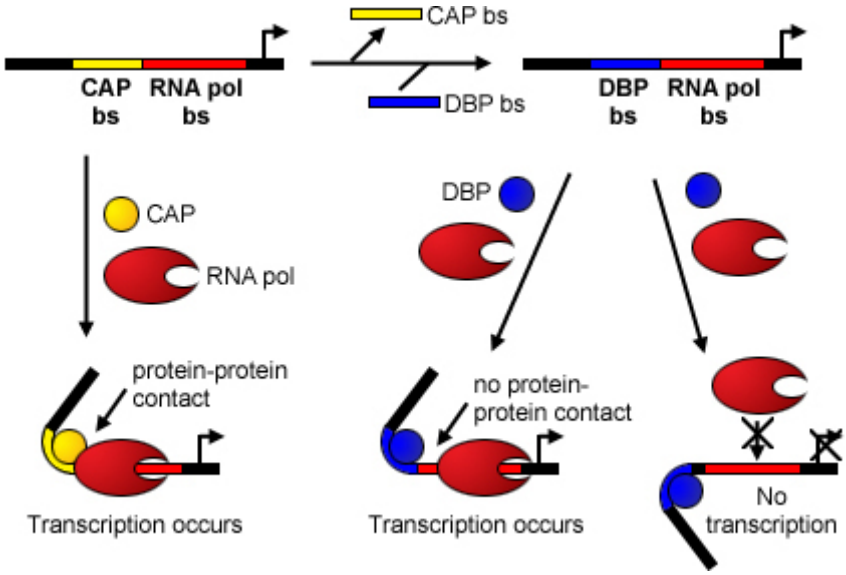


Figure 14. Protein replacement assay. The naturally occurring CAP binding site (CAP bs) was replaced with a DNA bending protein binding site (DBP bs). CAP both bends DNA and interacts with RNA polymerase (RNA pol) resulting in transcription. Orienting the DBP bs as the CAP bs resulted in transcription through the bending action of the DNA bending protein (DBP) without RNA pol interactions. Orienting the DNA bs such that the DBP induced DNA bend was opposite that of CAP resulted in no transcription.

The ability of an A-tract to function as an RNA polymerase binding site eliminated its use as an accurate tool to evaluate the role of DNA bending in transcription. Other approaches were required to determine the function of DNA bends in transcriptional regulation. A simple approach involved replacing the binding site of a native DNA-bending protein with the binding site of a DNA-bending protein foreign to the system (Fig. 14). Perez-Martin and Espinosa used this method to determine the importance of a DNA bend induced by CAP in transcription *in vivo*.[104] They replaced the native CAP binding site in the *fur* operon with the binding site of an unrelated repressor protein, RepA. RepA is a known DNA-bending protein and heterologous to CAP,

preventing CAP from interacting with both the RepA binding site and RNA polymerase. Like the A-tract replacement experiments, placing the RepA site in the same phase as a naturally occurring CAP binding site increased the interaction of RNA polymerase with the promoter and transcription was initiated. Conversely, when the RepA site was phased opposite to the normal CAP binding site orientation transcription was hindered.

Déthiollaz *et al* carried out a similar experiment, replacing the native CAP binding site in a *malT* promoter with a binding site for an IHF protein in *in vivo* experiments.[105] Unlike CAP, the heterologous IHF protein did not interact with RNA polymerase, limiting any effects on gene expression solely to the IHF-induced DNA bend. Transcription was activated in the altered *malT* promoter when IHF was bound to its binding site oriented as the original CAP binding site. These protein replacement studies demonstrated the importance of the DNA conformation in transcription, specifically the role of DNA bend orientation in relation to the start of a gene.

The role of DNA bending in transcription has also been analyzed through the insertion of a DNA-bending protein binding site into a gene promoter, between the binding sites of transcriptional factors whose interaction is required for transcriptional activation (Fig. 15).[23] The binding sites of YY1, LEF-1 and Sp1 proteins were evaluated for their abilities to modulate transcription. The binding sites for these proteins were inserted between eukaryotic transcriptional complex binding regions, where the RNA polymerase bound, and activator/enhancer regions, which increase the rate of transcription. All of the proteins were able to induce DNA bends at their inserted binding sites. The bends induced by YY1, LEF-1 and Sp1 facilitated protein-protein interactions between the transcriptional binding complex and proteins at the activator or enhancer

regions, allowing transcription to proceed. Transcription only occurred when the induced bend was properly oriented, allowing for interaction between proteins bound to the transcriptional complex binding site and to the activator/enhancer binding site. Improper phasing of the induced bend inhibited transcription by preventing the required interactions of the distally located proteins in the promoter region.

The importance of the orientation of a DNA bend relative to the transcriptional start site of a gene and its role in gene expression was also demonstrated by altering the flexibility of sequence of DNA upstream of a gene. Scaffadi's group created various mutants of the DNA bending protein Sox2 and its respective binding site, located upstream of an *Fgf4* gene.[93] These mutations showed a correlation between the flexibility of DNA and gene expression *in vivo*. When a mutation granted DNA greater levels of flexibility the level of transcription of the *Fgf4* gene increased. Conversely, decreased flexibility resulted in a loss of transcription and gene expression.

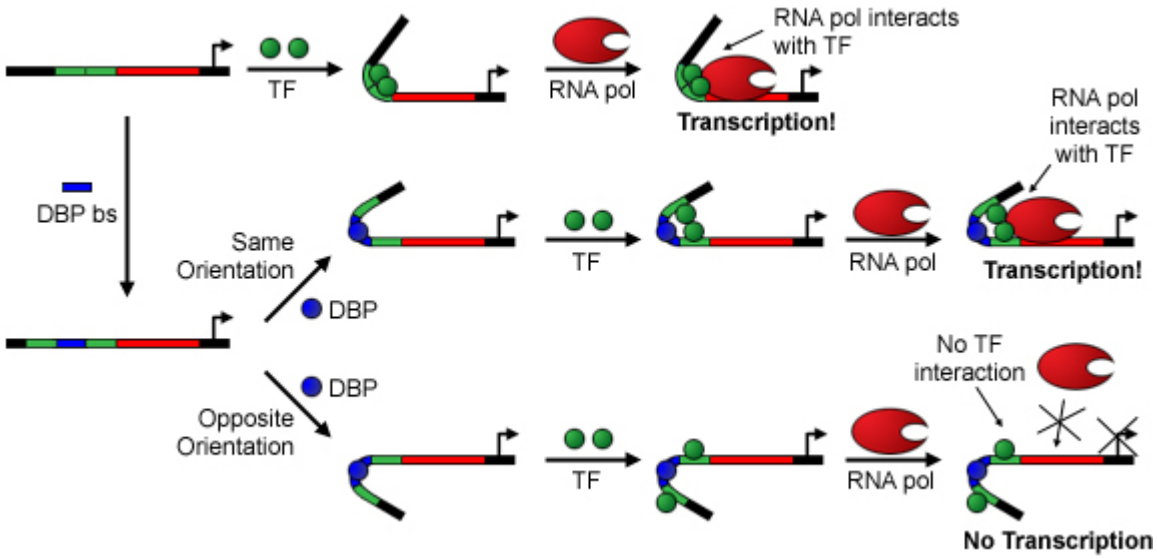


Figure 15. DNA bending protein binding site insertion into RNA polymerase transcription factor binding region.

I.D. Artificial DNA Bending

The ability of DNA bending to either up regulate or down regulate transcription indicated that it could serve as a potential mechanism to artificially regulate genetic expression to treat transcription related diseases. Targeting DNA as a means to regulate transcription has advantages over other methods of artificial gene modulation.[106-107] As opposed to targeting gene products such as mRNA or proteins, which exist in abundance in every cell, targeting DNA is attractive due to there being only a single target per cell. Targeting DNA bending has an additional advantage that it does not require competition with the potent binding of transcription factors. Only the overall shape of the transcriptional complex is important. On the other hand, targeting a specific sequence within genomic DNA becomes difficult with the potential existence of multiple binding sites, making sequence specificity a priority in designing DNA binding agents.

How does one induce a DNA bend? Within cells, DNA bending is present in the forms of intrinsically bent DNA sequences and the binding of DNA bending proteins. These methods would be impractical as mechanisms to artificially induce DNA bends. Sequence-dependent DNA bending would require alteration of the genomic DNA sequences, which would require the delivery and integration of a modified gene into the genome. Using DNA-bending proteins would be less invasive, but their use has two major drawbacks. First, the large size and multiple charges of many proteins make transport into the cell and nucleus difficult. Secondly, a protein would need to be engineered to target a specific sequence of DNA and retain its DNA bending ability.

DNA bends have been artificially induced with non-protein ligands. The mechanisms these moieties use to induce bends fall into one of two opposing categories; pull benders or push benders (Fig. 16).

Pull benders are composed of two DNA binding regions linked together with a molecular tether. The only way that the two DNA binding regions can simulatenously bind to DNA is if the DNA bends. Push benders bind to DNA and widened one of the grooves of DNA, resulting in the DNA being pushed away from the moiety and bending away from the bound ligand.

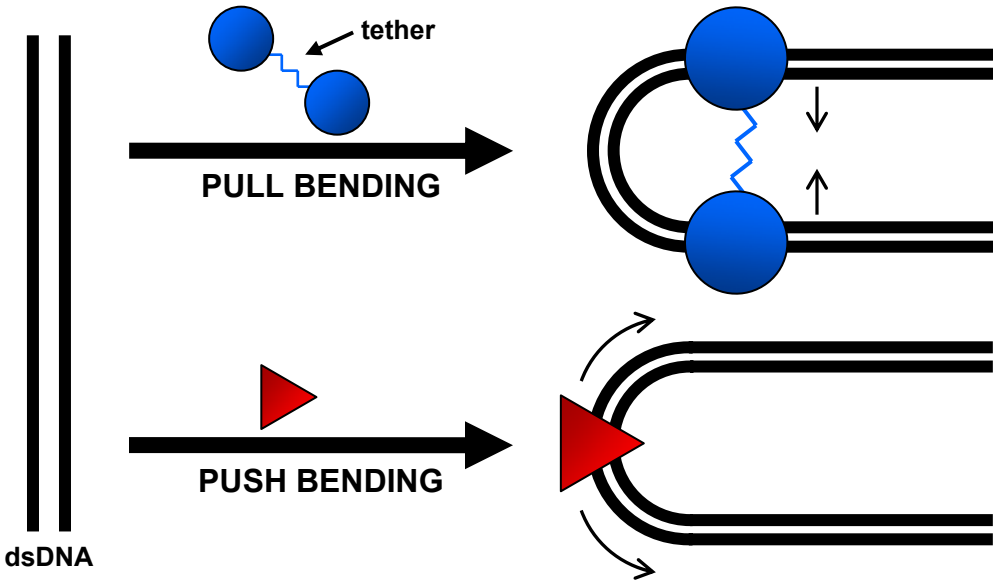


Figure 16. These are the general pull and push bending mechanisms. A pull bender is an agent with two, tethered DNA binding regions that, once bound, pull and hold DNA into a bent conformation. A push bender acts by binding into a DNA groove and widening it, pushing the walls of the groove away from the agent and bending the DNA away from the compound.

Tethered-triple helix-forming oligonucleotides, TFOs, are the quintessential example of pull-bending DNA ligands. Tethered-TFOs are comprised of two single-stranded, triple-helix forming DNA arms connected by a variable length molecular tether that does not interact with DNA.[75,108-109] The utilization of tethered-TFOs in DNA

bending strategies has been investigated using both parallel and antiparallel systems of triplex formation.[75,108-111] Both binding strategies used the same basic configuration; two TFOs, targeted to sites located one helical turn apart on duplex DNA, linked together by a linear molecular tether (Fig. 17). The tethered-TFOs bound to duplex DNA and restricted the movement of the duplex DNA, holding it in a bent conformation.

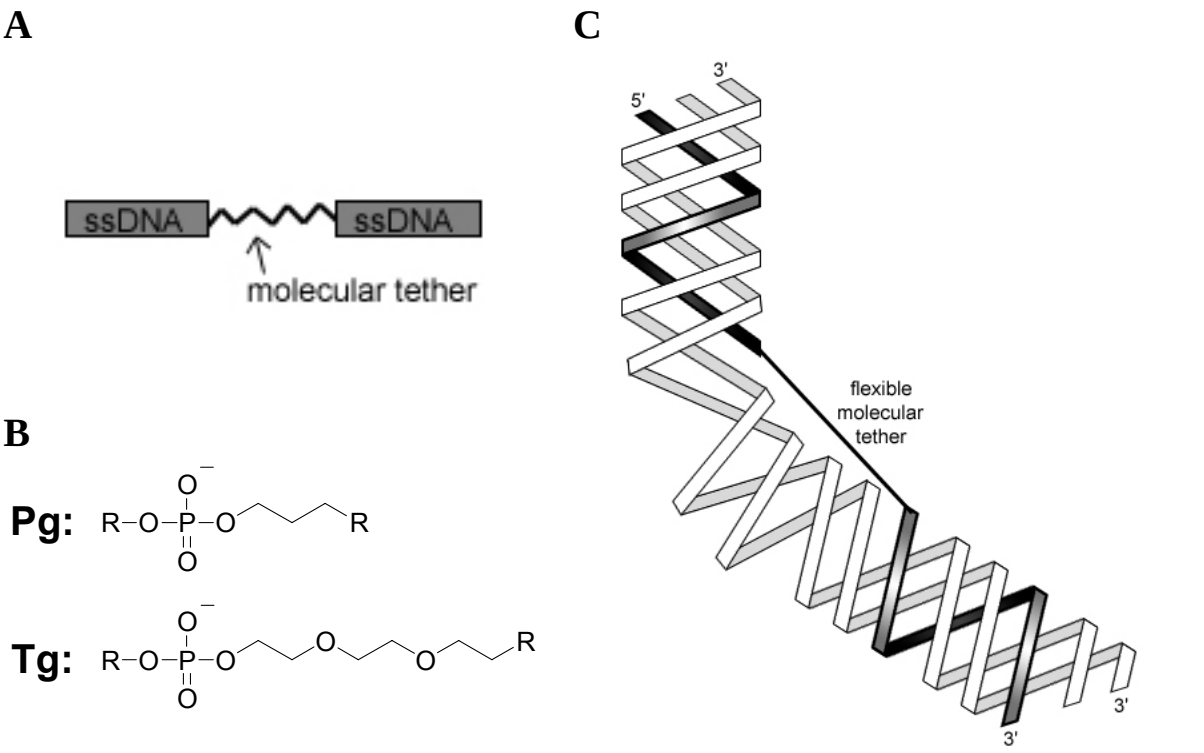


Figure 17. (**A**) The general structure of a tethered TFO, where the single stranded DNA, ssDNA, can run either 5' to 3' or 3' to 5'. (**B**) The linker subunits propylene glycol phosphodiester, **Pg**, and triethylene glycol phosphodiester, **Tg**, utilized by Akiyama and Hogan. (**C**) Bending of a DNA duplex by a TFO.

Akiyama and Hogan focused on tethered-TFO systems that utilized a variety of molecular linkers which varied in length, from 18 to 44 rotatable bonds, to connect a pair of antiparallel binding TFOs that were stabilized by T•AT and G•GC Reverse Hoogsteen base pairs.[107,109-110] The antiparallel orientation was selected due to its ability to form stable triplexes under neutral conditions. A variety of linkers were tested based on the

hypothesis that linkers with 25 or more rotatable bonds would not bend DNA, while a linker shorter than 25 rotatable bonds would induce a DNA bend once the tethered-TFO was bound.[111]

Akiyama and Hogan established a number of conclusions based on their work with tethered-TFOs.[108-110] They noted that the bends induced by the tethered-TFOs were toward the minor groove of the DNA and reasoned that this bend directionality explained why the free energy required to induce a bend in the DNA of around 60° with a tethered-TFO was less than they had projected. Akiyama and Hogan discerned that the sequence of the region of DNA spanning between the TFO binding sites affected the magnitude of the induced bend with AT-rich regions bending the most; though they claimed no relation to intrinsic bending. They also demonstrated that shortening the length of the tethered-TFO linker resulted in a direct effect on the magnitude of the induced DNA bend; the shorter the linker the greater the induced bend. A tethered-TFO with the shortest linker, three propylene glycol phosphodiester units with a total of 18 rotatable bonds, bound to duplex DNA with an AT-rich region between the TFO binding sites was able to induce a bend of about 60° in the duplex structure. They demonstrated that a slight DNA bend was detected when tethered-TFOs with linkers comprised of 30 and 33 rotatable bonds, while linkers longer of this length were computationally projected to not possess the ability to bend DNA.[111] Akiyama and Hogan reasoned that these linkers were not fully extended as expected, with probable gauche conformations in their linear structures, as opposed to the expected anti- conformations. Suprisingly, unlinked TFOs were also reported to induce slight bends in duplex DNA targets, though no explanation as to why this occurred was given.

The second category of DNA bending agents was the push benders. Push bending is the result of introducing steric bulk into the minor groove of DNA such that the groove widens and the DNA bends away from the molecule. Examples of such compounds include ecteinascidin-743 (ET-743) and calicheamicin (Fig. 18).

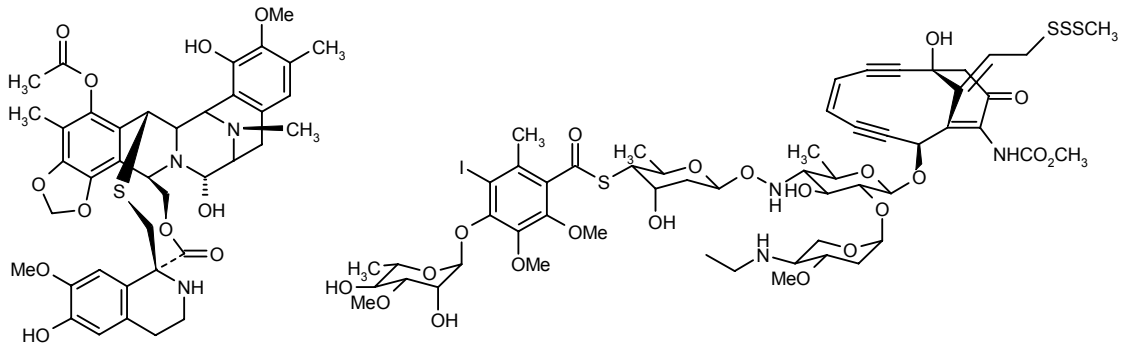


Figure 18. Ecteinascidin-743, ET-743, and calicheamicin-ζ_1^I.

ET-743 is a natural product possessing potent antitumor activity, specifically against soft tissue sarcomas and advanced breast cancer.[112-114] ET-743 covalently attaches to DNA via the exocyclic N2 position of guanine. ET-743 binds to DNA triplets with central guanine nucleotides, 5'- purine-G-C-3' and 5'-pyrridine-G-G-3'.[115-116] Once bound in the minor groove, ET-743 protrudes perpendicularly from its binding site (Fig. 19). The bulk of projecting molecule exerts steric force on the walls of the minor groove, widening it and causing the DNA to bend away from the molecule toward the major groove.

Calicheamicin (Fig. 18) is an antibiotic and tumoricidal agent belonging to the carbohydrate class of DNA binding compounds.[117-119] Calicheamicin possesses both an aryltetrasaccharide tail responsible for binding the compound to the minor groove of DNA and an enediyne region that is responsible for cleaving DNA upon binding. In order to bind to the minor groove of the DNA calicheamicin requires that the groove

widen to accommodate the bulk of the enediyne region of the molecule.[119] This widening of the minor groove corresponds to a compression of the major groove on the opposite face of the DNA.

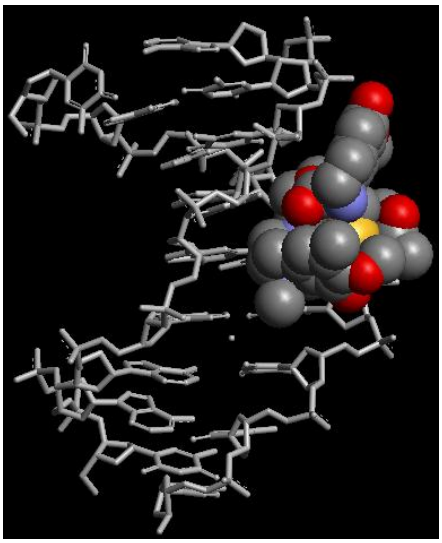


Figure 19. ET-743 bound in the DNA minor groove.

II. STATEMENT OF THE PROBLEM AND OVERVIEW OF EXPERIMENTAL APPROACH

A number of diseases have aberrant gene expression.[2-4,120-123] Thus, methods to artificially regulate gene expression would be a tremendous benefit. The long term goal of the projects presented here is to create sequence specific DNA bending agents to artificially regulate genetic expression (Fig. 20). These agents will induce a bend upstream of the promoter region of a gene resulting in either up-regulation or down-regulation of expression. The ability to artificially control the expression of a gene would allow for the potential treatment of a number of diseases including cancer,[1-3,120] cardiovascular diseases[121] and sickle cell anemia.[122-123]

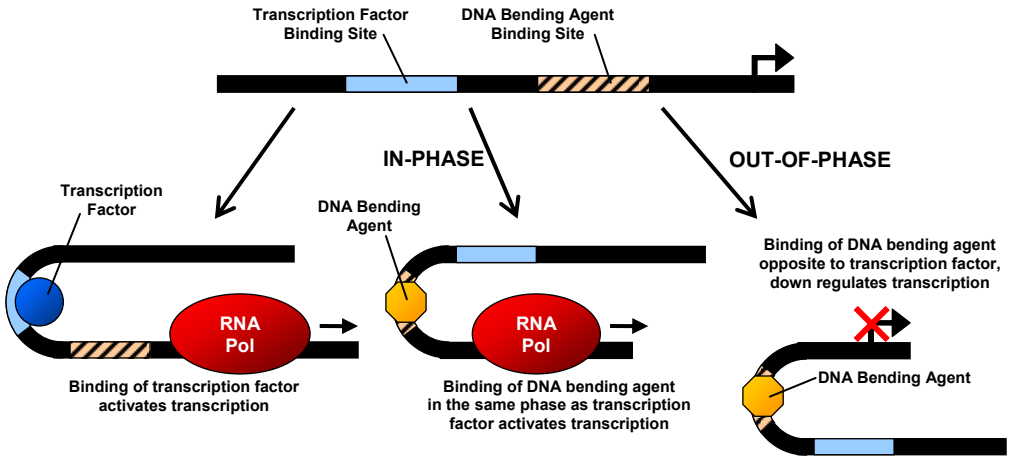


Figure 20. Regulation of transcription using an artificial DNA bending agent. Induction of a bend with an artificial DNA-bending agent similar to that of a DNA-bending protein will result in an "in-phase" DNA bend and an up regulation of transcription. Bending DNA contrary to that of DNA bending protein would result in an "out-of-phase" DNA bend and a down regulation in transcription.

Previous studies showed that both intrinsic DNA bends and bends induced by DNA binding proteins effected transcription levels with the orientation of the bend relative to the start of a gene being a critical determinate for gene regulation.[23,93,101-105] When the bend was oriented analogous to that caused by a naturally occurring transcription factor or RNA polymerase, an orientation known as an in-phase bend, a gene would be up regulated or activated.[23,101-105] Conversely, an out-of-phase DNA bend, oriented in the opposite direction, resulted in the down regulation or silencing of gene expression.[23,93,102,104]

How can one bend DNA? There are two general methods to artificially bend DNA, namely pulling DNA into a bent conformation or pushing it into bent conformation. Pull bending is accomplished using a molecule which contains two DNA-binding domains. The only way in which the two domains can bind to DNA is if the DNA bends. Push bending is accomplished through the widening of a DNA groove with

a DNA binding agent. This results in DNA bending away from the bound moiety. The work presented in this thesis is divided into two main projects, with each project examining one of the two methods to artificially bend DNA. The first project focuses on the artificial regulation of gene expression through DNA bends induced by tethered-TFOs. The second project involves the synthesis and investigation of a small, sequence selective, minor groove binding agent that function as a push bender through the introduction of bulk into the minor groove of a DNA target sequence.

II.A. Project 1: Tethered-TFOs and Gene Expression

The first project will address the hypothesis that a non-protein moiety could regulate gene expression through the induction of DNA bending. This project is essential in the development of artificial gene regulating agents as it will either validate or refute the concept that DNA bending alone could influence gene expression and that this influence could be achieved through artificial means. We will use Akiyama and Hogan's tethered-TFO system to artificially induce DNA bending in a synthetic promoter/gene system.[108-110] After bending of the target sequence is verified, we will examine the effects of bending on the expression of our target gene. The orientation of the bend will be varied and we anticipate that proper orientation of the induced DNA bends will result in either the activation or repression of the target gene (Fig. 21). The above hypothesis is supported by previous studies utilizing intrinsically bent DNA and DNA-bending proteins to study the relationship between DNA bending and gene expression.[23,93,101,104-105]

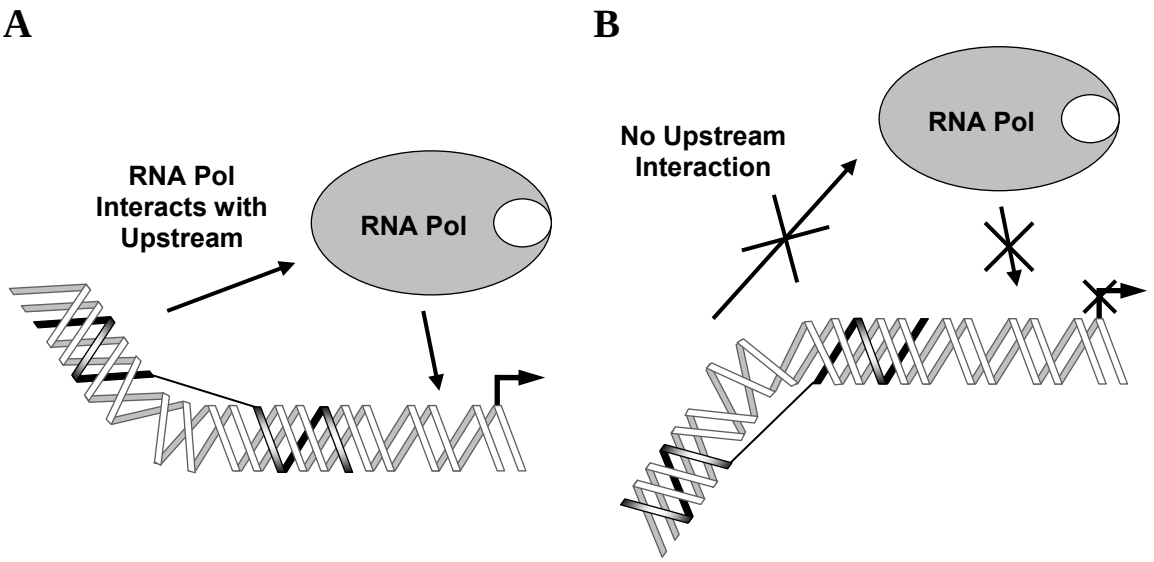


Figure 21. (**A**) Tethered-TFO bending DNA in-phase with a gene, causing the activation of transcription. (**B**) Tethered-TFO bending DNA out-of-phase with a gene resulting in the repression of transcription.

The sequence of the oligonucleotides and tethered-TFOs used in this study are shown in figure 22. The tethered-TFO that induced the greatest DNA bend in Akiyama and Hogan's experiments, labeled in our studies as the short tethered-TFO (STFO), will be used as the positive control.[108-110] The negative control in our research will be the tethered-TFO that was previously shown to display no DNA bending ability; we have dubbed this moiety the long tethered-TFO (LTFO). The negative control will be used to verify that any effects seen in our experiments will be due to DNA bending and not the binding of the tethered-TFO to the DNA targets. We will also utilize "broken" versions of these tethered-TFOs (bSTFO, bLTFO) where the molecular tether and one TFO region will be separated from the other TFO (Fig. 22). These broken variants of the tethered-TFOs will be used to verify that any affects on gene expression will be due to DNA bending and not the formation of DNA triplexes.

STFO: 5′-GGTGGGGTGGGTGGG-**PgPgPg**-GGGTGGGTGGGGTGG-3′
LTFO: 5′-GGTGGGGTGGGTGGG-**TgTgTgTg**-GGGTGGGTGGGGTGG-3′
bSTFO: 5′-GGTGGGGTGGGTGGG-3′ and 5′-**PgPgPg**-GGGTGGGTGGGGTGG-3′
bLTFO: 5′-GGTGGGGTGGGTGGG-3′ and 5′-**TgTgTgTg**-GGGTGGGTGGGGTGG-3′

Figure 22. The tethered-TFOs that we will use in this project. The short tethered-TFO, STFO, that will bend DNA and the long tethered-TFO, LTFO, which will serve as the negative control tethered-TFO com from Akiyama and Hogan. The broken variants will be called bSTFO and bLTFO. **Pg** represents a propylene glycol phosphodiester unit and **Tg** represents a triethylene glycol phosphodiester unit; see **Fig. 6** for structures.

Before testing the ability of tethered-TFOs to regulate gene expression, we will first verify that the target tethered-TFO binding sequence can be bent by STFO and not any of the other TFOs; LTFO, bSTFO and bLTFO. Gel mobility shift assays will be used for tethered-TFO induced bend analysis. DNA fragments containing the most flexible tethered-TFO target sequence used by Akiyama and Hogan will be used in these assays.[108-109] After establishing the bending ability of our tethered-TFOs, we will then investigate the effects of the bends will have on gene expression when the bends are induced in varied orientations upstream of a luciferase gene.

A series of plasmids containing the tethered-TFO target sequence used by Akiyama and Hogan will be positioned upstream of a promoter region and a luciferase gene (Fig. 23).[108-109] By varying the distance between the tethered-TFO target sequence and the transcriptional start site of the luciferase gene, the orientation, or phase, of the induced bend relative to the gene will change as well. TFO:DNA complexes will be subjected to an *in vitro* transcription/translation system. The expressed luciferase protein will then be quantified through Western blot analysis. We will determine whether the induced DNA bends have an effect on gene expression levels and, if so, determine the role that the phase of a bend has on artificial regulation.

A

5′-GGAGGGGAGGGAGGGCGCCGGCGCGGGAGGGAGGGGAGG-3′
3′-CCTCCCCTCCCTCCCGCGGCCGCGCCCTCCCTCCCCTCC-5′

B

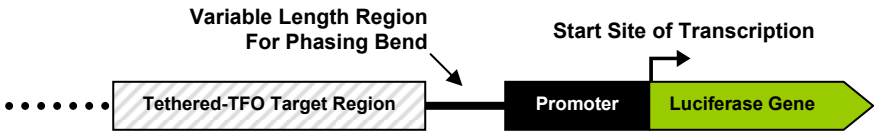


Figure 23. (**A**) The target sequence for these tethered-TFOs. The triplex target regions are highlighted in gray, while the center of the induced bend is underlined. (**B**) Region of interest in plasmid series containing the tethered-TFO target, the variable length region to phase the induced DNA bend, the promoter region and the start of the luciferase gene.

We expect to learn that artificially induced DNA bends can effect gene expression and that the phase of the bend plays and important role in gene expression. This project will serve to validate or refute the use of artificially induced DNA bends to regulate genetic expression.

II.B. Project 2: A Sequence Selective Push Bender

A series of small molecules that will function as sequence selective DNA push benders will be created in the second project. While tethered-TFOs pull DNA into bent conformations that can artificially regulate the expression of a gene, proving the concept of artificial gene regulation through induced DNA bends, they are poor pharmaceutical candidates due to their size and the difficulties in the formation of DNA triplexes *in vivo*.[124] Artificial gene regulation accomplished by a small molecule capable of passing through membrane barriers, selectively binding to a specific DNA target sequence and inducing a bend at that site upon binding would be ideal.

Our proposed push bending compounds were designed to widen the minor groove of a DNA target through steric interactions between the walls of the minor groove and a

bulky moiety introduced by the binding of our compounds. This bending mechanism will be coupled to a sequence specific, minor groove binding region to create sequence selective push bending molecules.

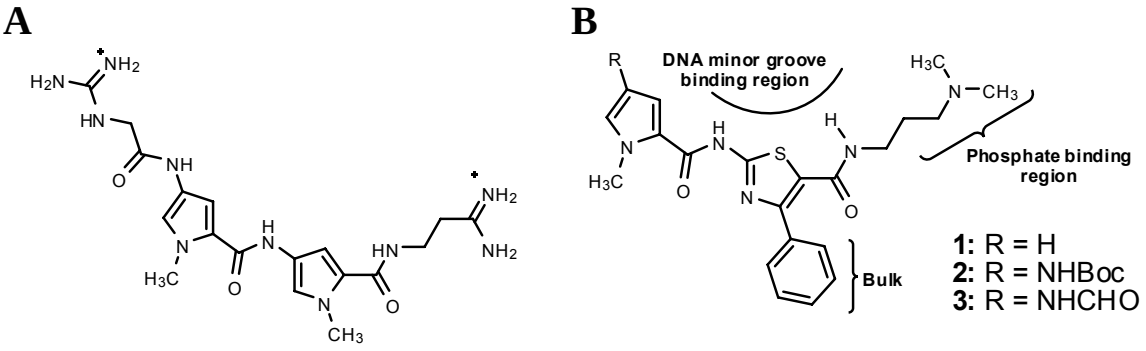


Figure 24. (**A**) The lead compound, netropsin. (**B**) The structures of the netropsin analogues, **1**, **2** and **3**. The DNA minor groove binding region and phenyl bulk are labeled.

We will synthesize and evaluate a series of compounds that will act as sequence selective DNA push bending agents (Fig. 24). Our compounds (**1**, **2** and **3**) were modeled after the mechanisms of push benders such as ET-743 and calicheamicin, which introduce a bulky functional group in the minor groove of DNA forcing the groove to widen and bend the DNA. Our compounds will couple a DNA binding region modeled after netropsin to a bulky region that will widen minor grooves upon binding. These analogues consisted of three main regions; a DNA minor groove binding region, a phosphate binding region and a bulky moiety.

The minor groove binding region of our compounds was modeled after the amide-linked, dual *N*-methylpyrrole binding region of netropsin. This region should direct binding to AT-rich sequences based upon its similarity to netropsin. A phenyl-bearing thiazole ring was substituted for the C-terminal *N*-methylpyrrole, and was oriented with the thiazole sulfur directed toward the minor groove-binding edge of the structure. This

substitution was done to ease the synthesis of our target compounds and to maintain the AT-rich DNA sequence binding preference of netropsin.[55]

The second region common to all of our compounds was the phosphate binding region located at the C-terminus of the structures. Although netropsin possessed a C-terminal amidine, we opted to use a single dimethylaminopropyl arm. The dimethylaminopropyl arm was one methylene unit longer than the amidine arms of netropsin, granting the chain greater flexibility and greater freedom of movement to interact with DNA backbone phosphates once our compounds were bound to the minor groove. The dimethylaminopropyl arm used in our compounds had been previously used by Dervan in the construction of various polyamide minor groove binders.[125] The substitution of a dimethylpropylamide tail for the amidine of the model compound will result in a decreased basicity which will enhance transport into the cell should these agents bend DNA as hoped.[126-127]

The bulky moiety that our compounds will introduce into the minor groove of their DNA binding sites was a phenyl ring attached to the 4-position of the thiazole ring of the DNA minor groove binding region of the structures. The phenyl ring will sterically interact with the walls of the minor groove, widening it, pushing the DNA into a bend away from the bound compounds. Molecular modeling, by Dr. Steven Firestine with Molecular Operating Environment (MOE), shows that this phenyl ring will be projected perpendicularly from the minor groove binding region, maximizing steric interactions within the groove (Fig. 25).

A **B**

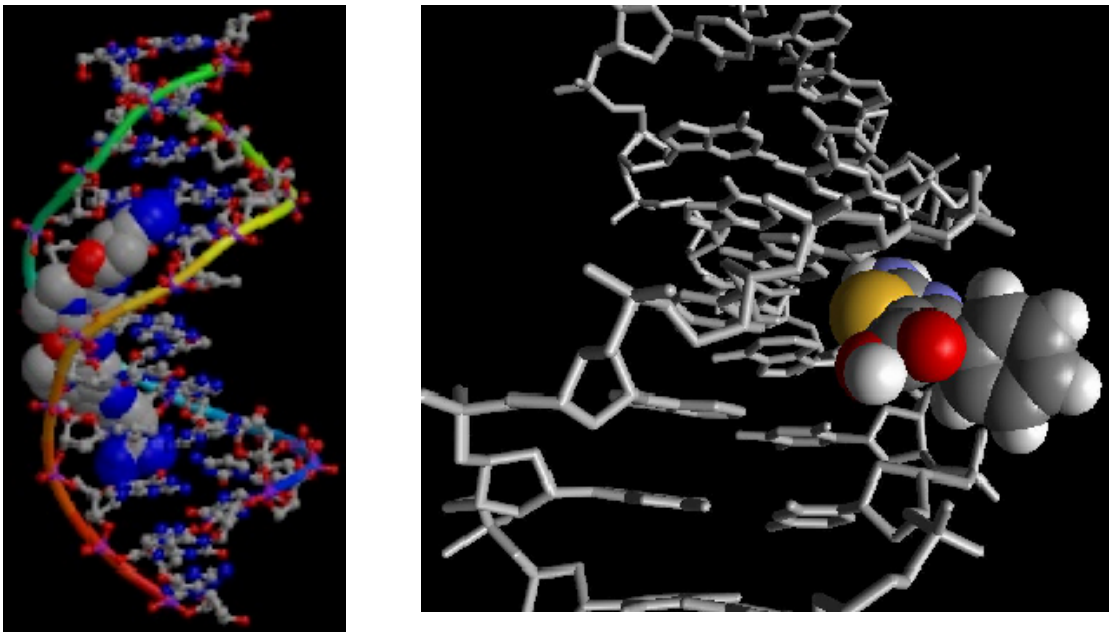


Figure 25. (**A**) Netropsin bound in the minor groove of DNA (1DNE from PDB). (**B**) Compound **1** modeled in the minor groove of DNA using MOE.

The N-terminal guanidine arm bound to N-methylpyrrole of the lead compound netropsin was replaced by three different functional groups including a proton (compound **1**), a tert-butoxycarbonylamino group (compound **2**), and a formamide group (compound **3**). The single proton at this position in compound **1** should provide insight into the role that this position plays in the binding of our compounds to the minor groove of the DNA by removing the hydrogen binding and phosphate backbone interactions from this location. Conversely, compound **2** possessed the largest functional group at the 4-position of the *N*-methylpyrrole; a tert-butoxycarbonylamino group. Previous, unpublished work from our laboratory had shown that the *t*-butyl group could alter the conformation of DNA. Thus, we were interested in examining whether the inclusion of two potential DNA bending groups would enhance DNA bending. Finally, a formamide was placed at the 4-position of the *N*-methylpyrrole in compound **3** in an effort to

enhance binding affinity by including an additional hydrogen bond donor from the –NH of the amide bond of the formamide.[48,51]

After synthesizing our compounds, they will be evaluated for their DNA binding ability, sequence binding preference and their ability to bend DNA. An ethidium displacement assay will be used to determine binding ability and sequence preference. We expect our compounds to maintain the AT-rich sequence preference displayed by the lead compound, netropsin. We also expect that the various analogues will possess different binding affinities, with the formyl-capped compound **3** having the greatest binding affinity. FRET analysis will be used to determine what effect, if any, our compounds have on the conformation of a DNA target. We expect to see a level of conformational distortion in DNA targets with the preferred DNA target sequence. These studies will provide important information to the future development of small, sequence selective DNA bending agents to artificially control the expression levels of target genes.

III. RESULTS

III.A. Project 1: Tethered-TFOs and Gene Expression

III.A.1. *Construction of Required Materials*

To investigate the role DNA bending played in gene expression, we utilized the tethered-TFO system described by Akiyama and Hogan.[108-110] This system required two major components, namely a series of plasmids (pBLP) containing the target sequence for the tethered-TFOs and the tethered-TFOs with varied linker regions.

III.A.2. *Construction of pBLP Plasmids*

The plasmid series pBLP was constructed through the insertion of oligonucleotides into a series of vectors derived from pBR322. The four plasmids of the pBLP series included two main sequence regions; a binding site for our tethered-TFOs spaced at various distances from a promoter sequence and a luciferase reporter gene. With the necessary fragments, we constructed four plasmids; pBLP77, pBLP80, pBLP83, and pBLP86. The nomenclature used for our plasmid series was derived from the location of the bend relative to the start of the luciferase gene. We constructed the pBLP plasmid series in a stepwise fashion (Fig. 26).

The purpose of the pBLP plasmid series was to position a luciferase gene at varying distances from the center of a target site for a tethered-TFO at which a DNA bend would be induced. The helical nature of DNA means that varying the distance between two points on a strand of DNA results in a different orientation of the bend relative to the start of the gene. Assuming that our plasmids adopt the most common form of DNA, B-DNA, each turn of the helix contains 10.5 base pairs. The pBLP plasmid series varied the distance between the center of the DNA bend and the transcriptional start of a luciferase gene by 3 base pair increments, ranging from 77 to 86 base pairs. This created four different phases of the bend in relation to the luciferase gene (Fig. 27A). At 10.5 base pairs per turn difference between the different phases can be calculated in terms of degrees (Fig. 27B).

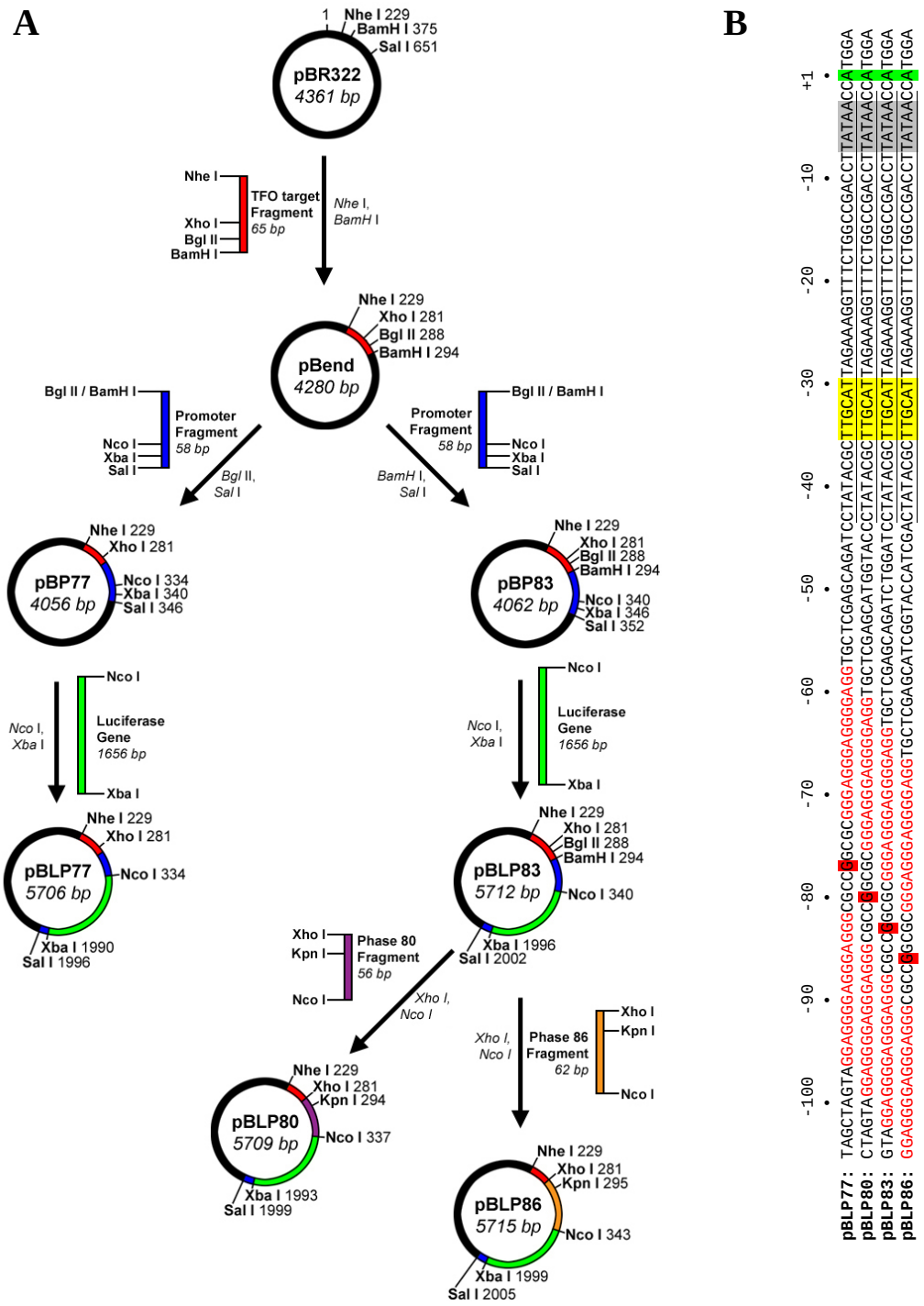
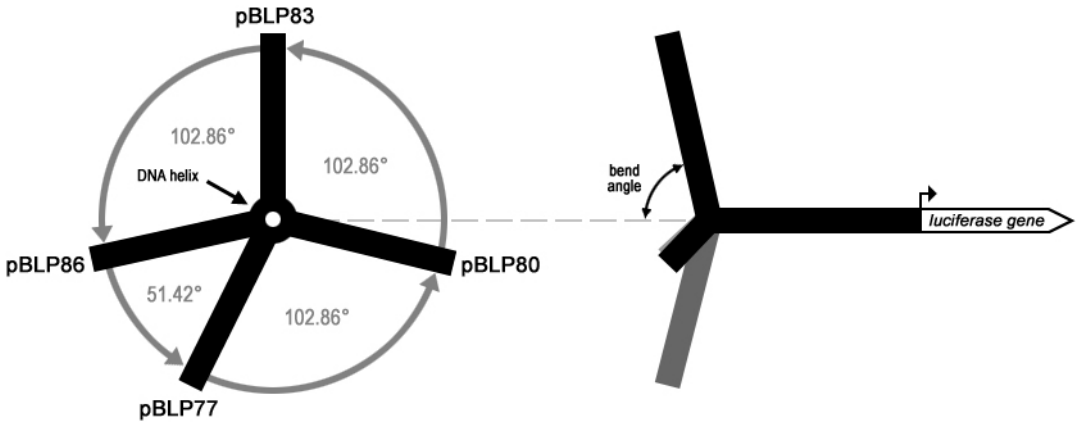


Figure 26. (**A**) Stepwise diagram of the creation of the pBLP plasmid series. (**B**) The sequences of plasmids pBLP77, pBLP80, pBLP83 and pBLP86. The tethered-TFO binding sites are red, with the center of the bend is highlighted in red. The *malT* promoter is under lined with the -10 region highlighted in gray and the -35 region is highlighted in yellow. The luciferase gene starts at the +1 position, highlighted in green.

A



B

Difference in Position in Degrees

Position	77	80	83	86
77	-	102.86°	154.28°	51.42°
80	102.86°	-	102.86°	154.28°
83	154.28°	102.86°	-	102.86°
86	51.42°	154.28°	102.86°	-

Figure 27. (**A**) Phases of the pBLP series of plasmid, relative to the start of the luciferase gene. Rotational degrees based on average of 10.5 bp/turn of the double helix. (**B**) Table of degrees of difference between the various phases of the pBLP series.

The pBend plasmid, named due to the incorporation of the tethered-TFO target sequence, served as a precursor to the rest of the pBLP series plasmids (Fig. 26). We inserted the TFO target fragment into the pBR322 plasmid at the *Nhe* I and *BamH* I restriction sites (Fig. 46). The successful creation of pBend was verified through restriction analysis at the unique *Xho* I restriction site, introduced into the plasmid with the TFO target fragment.

The pBLP plasmid series was created through the systematic insertion of various DNA fragments into the pBend plasmid. The first pBLP plasmid that we created was pBLP77, through the intermediate pBP77 plasmid. The pBP77 plasmid was built by the insertion of the promoter fragment into pBend at the *Bgl* II and *Sal* I restriction sites (Fig. 26). The overhang at the 5'-end of the promoter fragment was capable of ligation to both

digested *BamH* I and *Bgl* II sites. In placing the promoter fragment at the *Bgl* II position of pBend, the center of the tethered-TFO target sequence was located 77 base pairs upstream from the *Nco* I restriction site, where the start site of a luciferase gene would eventually be located. The promoter fragment not only positioned the luciferase insertion site relative to the center of the tethered-TFO bend, but also incorporated the *malT* promoter upstream of the *Nco* I insertion site for the luciferase gene. The unique *Nco* I and *Xba* I restriction sites of pBP77, introduced with the promoter fragment, served as a restriction analysis sites for verification of the pBP77 plasmid. *Bgl* II was also used in verification of the successful creation of pBP77, where the lack of digestion indicated a loss of the *Bgl* II restriction site, which in turn indicated the insertion of the promoter fragment.

The pBLP77 plasmid, the first complete pBLP plasmid, was created through the insertion of the luciferase gene into pBP77 (Fig. 26). The 1656-bp luciferase gene was excised from the pGL3 plasmid at the *Nco* I and *Xba* I restriction sites and subsequently inserted into the pBP77 plasmid at these same sites. The *Nco* I site of the pBP77 plasmid, the location of the transcriptional start site of the luciferase gene, was positioned 77 base pairs from the center of the tethered-TFO target sequence, where the bend in the DNA would occur. The creation of pBLP77 was verified by restrictive digestion with *Xho* I and comparison of plasmid length to a similarly digested pBP77 plasmid.

The pBLP83 plasmid, like the pBLP77 plasmid, was built in two steps: insertion of the promoter fragment into pBend to create the intermediary pBP83 plasmid followed by the insertion of the luciferase gene into pBP83 to create pBLP83. The pBP83 plasmid was constructed by inserting the promoter fragment into the pBend plasmid, as was done

with pBP77 (Fig. 26). Unlike pBP77, the promoter fragment was inserted into the *Sal* I and *BamH* I restriction sites, as opposed to the *Bgl* II site. Contrasting promoter fragment insertion at the *Bgl* II site, insertion at the *BamH* I site did not destroy the restriction site sequence. The promoter fragment included the *malT* promoter as well as positioned the *Nco* I site, for luciferase gene incorporation, 83 base pairs from the center of the tethered-TFO target sequence. We verified pBP83 through restriction analysis with *Nco* I and *Xba* I enzymes. The integrity of the *BamH* I restriction site was also tested to verify proper fragment insertion.

The pBLP83 plasmid was obtained through the addition of the luciferase gene from the pGL3 plasmid into the pBP83 plasmid at the *Nco* I and *Xba* I restriction sites (Fig. 26). The luciferase gene was positioned downstream of the *malT* promoter with its transcriptional start site located 83 base pairs from the center of the tethered-TFO target site, 6 base pairs further or about a half turn around the DNA double helix different than in the pBLP77 plasmid. The luciferase gene in pBLP83 was phased opposite that of the gene in the pBLP77 plasmid. Successful incorporation of the luciferase gene was confirmed through *Xho* I digestion and length comparison of the pBLP83 and pBP83 plasmids via gel electrophoresis.

The center of the tethered-TFO induced DNA bend was phased in two opposite orientations, relative to the transcriptional start of the luciferase gene, in pBLP77 and pBLP83 (Fig. 27). Phasing the gene in the intermediary positions was accomplished through modification of the pBLP83 plasmid. The pBLP80 plasmid was created through the insertion of the Phase 80 fragment, with a *malT* promoter, into the pBLP83 plasmid at the *Xho* I and *Nco* I restriction sites (Fig. 26). The Phase 80 fragment altered the distance

between the transcriptional start site of the luciferase gene and the center of the tethered-TFO target site, lessening the distance by 3 base pairs. The center of the tethered-TFO induced bend was positioned 80 base pairs from the start of the luciferase gene in a phase between those of the pBLP77 and pBLP83 plasmids (Fig. 27). The pBLP80 plasmid was verified by restriction analysis with *Kpn* I and *BamH* I, where *Kpn* I cut the pBLP80 plasmid and *BamH* I did not. The *Kpn* I restriction site was introduced to pBLP80 with the Phase 80 fragment, while the *BamH* I restriction site was lost when the Phase 80 fragment was incorporated.

The pBLP86 plasmid was constructed in the same manner as pBLP80. The Phase 86 fragment, with a *malT* promoter, was inserted intp the pBLP83 plasmid at the *Xho* I and *Nco* I sites (Fig. 26 and 46). The Phase 86 fragment changed the position of the transcriptional start site of luciferase to 86 base pairs from the center of the tethered-TFO target site. This repositioning phased the center of the tethered-TFO bend opposite that of the pBLP80 plasmid and intermediary to the pBLP77 and pBLP83 plasmids (Fig. 27). The pBLP86 plasmid was verified with *Kpn* I and *BamH* I restriction analysis in the same manner as pBLP80.

III.A.3. *Construction of Tethered-TFOs*

We used the tethered-TFOs of Akiyama and Hogan as bending moieties for our experiments.[108-109] We chose to use the tethered-TFO that had previously been shown to achieve the maximum bend angle as our positive control and we have dubbed this the short TFO or STFO (Fig. 28). The molecular tether used in this ligand was composed of three propylene glycol phosphodiester units and contained 18 rotatable bonds between its two triple helix-forming regions.

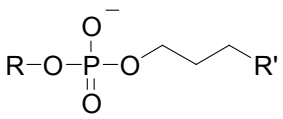
Short TFO (STFO)

5’-GGTGGGGTGGGTGGG-**PgPgPg**-GGGTGGGTGGGGTGG-3’

Broken Short TFO (bSTFO)

5’-GGTGGGGTGGGTGGG-3’
5’-**PgPgPg**-GGGTGGGTGGGGTGG-3’

Propylene glycol phosphodiester unit (Pg)



Long TFO (LTFO)

5’-GGTGGGGTGGGTGGG-**TgTgTgTg**-GGGTGGGTGGGGTGG-3’

Broken Short TFO (bSTFO)

5’-GGTGGGGTGGGTGGG-3’
5’-**TgTgTgTg**-GGGTGGGTGGGGTGG-3’

Triethylene glycol phosphodiester unit (Pg)

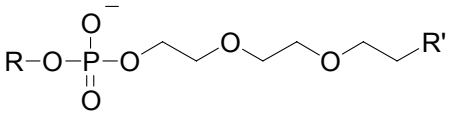


Figure 28. Sequences and molecular tether composition phosphodiester units. The molecular tethers of the STFO and the bSTFO contained 18 rotatable bonds, while the molecular tethers of LTFO and bLTFO each contained 44 rotatable bonds.

In addition to the DNA-bending positive control, STFO, we needed a negative control. We selected the longest tethered-TFO that Akiyama and Hogan investigated as our negative control.[108-109] This long tethered-TFO, LTFO, possessed a molecular tether composed of four triethylene glycol phosphodiester units with 44 rotatable bonds between the triplex-forming oligonucleotide arms (Fig. 28). LTFO displayed no DNA bending properties due to the length and rotational freedom of its molecular tether. The triethylene tether of LTFO was long enough to span a helical turn of DNA and allow the

formation of two triplexes without altering the conformation of the DNA. We used LTFO as a negative control to separate any gene expression effects caused by simple tethered-TFO binding from any alterations in gene expression related to DNA bending by STFO.

We created an additional set of negative controls in the form of broken versions of the two tethered-TFOs, bSTFO and bLTFO (Fig. 28). These consisted of two single-stranded DNA oligonucleotides. These broken oligonucleotides were created to determine whether or not bending was dependent upon tethering of two TFOs and whether gene expression was altered simply by the formation of DNA triplexes.

III.A.4. *DNA Bending by Tethered-TFOs*

Prior to exploring the effects of tethered-TFO-induced DNA bends on gene expression, we first examined the ability of our selected tethered-TFOs to bend or not. We used gel mobility analysis to accomplish this. In this method DNA fragments of identical length migrate through a gel to the same extent as long as the fragments have the same shape. Different shapes (circles, linear and bent) migrate at rates inversely proportional to the end-to-end distance of a DNA fragment; the shorter this distance the slower the fragment will migrate through a gel matrix.[128] Thus, DNA containing a bend will migrate slower than linear DNA due to a shorter end-to-end distance.

We were unable to use the pBLP plasmids created above for gel mobility analysis because the plasmid was greater than the persistence length of DNA. Thus, we PCR amplified a region of the pBend plasmid to generate two target DNA bending fragments each containing the tethered-TFO binding site at a different location relative to the ends

of the fragment. Two sets of primers and subsequent PCR were used to produce two 129-base pair DNA target fragments (Fig. 29).

A

C1:5′-CTAGCTAGCTAGTAGGAGGG-3′
C2:5′-CGCGGATCCAGATCTGCTCG-3′
E1:5′-CGTGCTGCTAGCTAGT-3′
E2:5′-TGTAGGAGCTATAGGC-3′

B

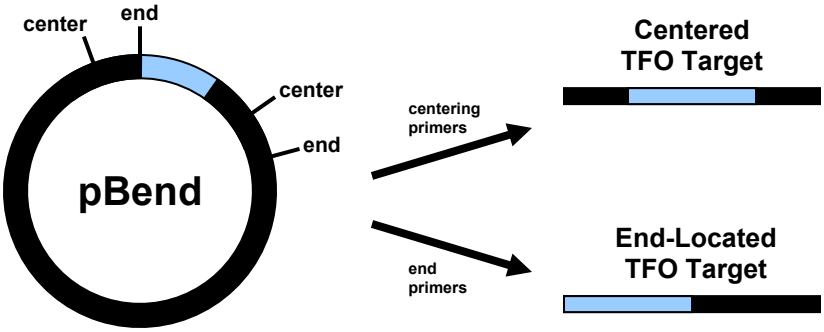


Figure 29. (**A**) Sequences of PCR primers used to create the centered TFO target DNA fragement, C1 and C2, and to create the end-located TFO target fragment, E1 and E2. (**B**) PCR amplification of 149 bp centered and end-located TFO target fragments.

The C1 and C2 primers were used to amplify the TFO target DNA fragment with a centrally located tethered-TFO target site. The presence of the bend directly in the center of the fragment would create the shortest possible end-to-end distance and thus correspond to the slowest possible mobility through a polyacrylamide gel. The E1 and E2 fragments were used to create a DNA target with a tethered-TFO target sequence located at one end of the fragment. A bend induced at this location would only affect the end-to-end distance to a small degree and, likewise, slow fragment gel migration only marginally.

The two target DNA fragments, centered and end-located, were complexed with various tethered-TFOs and analyzed by non-dentauring PAGE. Analyzing both the

centered and end-located fragments allowed us to determine whether shifts in the gel mobility of a complexed fragment was due to either DNA bending or another phenomena, such as an increase in the molecular weight. We expected to see the mobility of the centered TFO target fragment complexed to STFO, slowed to a greater extent than any other fragment complex.

We compared the migration of the uncomplexed tethered-TFOs, the uncomplexed DNA fragments and the TFO:DNA complexes in a gel mobility assay (Fig. 30). The tethered-TFOs were too small to be seen in our polyacrylamide gels and migrated off the bottom of the gel.

The uncomplexed, centered TFO-target DNA fragment had mobility consistent with its expected size. When coupled to STFO, the centered TFO-target DNA fragment displayed the greatest loss of mobility of any DNA fragment:tethered-TFO complex, running similar to the 225-bp band of the DNA ladder (Fig. 30, lane 5). In contrast, complexes of the centered TFO-target DNA fragment and LTFO or either of the broken tethered-TFOs, bSTFO or bLTFO, resulted in only a minor change in mobility when compared to the uncomplexed DNA fragment.

The uncomplexed end-located TFO target DNA fragment migrated similarly to the 125-bp band of the DNA ladder, close to its 129 bp length. The end-located TFO target fragment migrated slowest when complexed to STFO with a migration similar to 175-bp fragment, not as slow as the centered TFO target:STFO complex. When complexed to LTFO and the broken tethered-TFOs, the end-located target DNA fragment migrated as a 150-bp fragment, as did the complexes of these three tethered-TFOs and the end-located TFO target fragment.

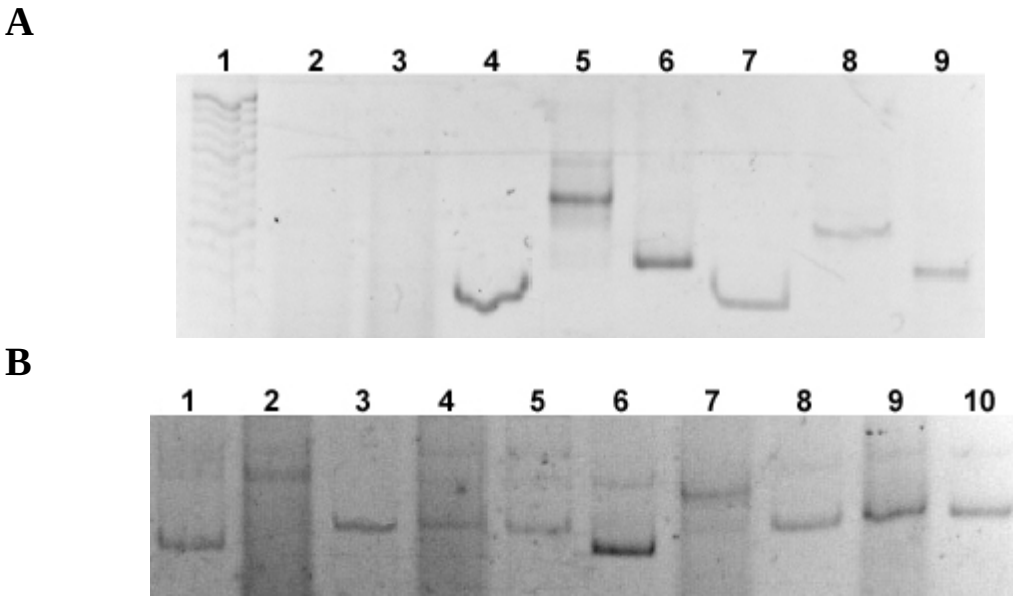


Figure 30. The gel mobility assay gel picture. In the lane assignments C represents the centered TFO target DNA fragment and E represents the end-located TFO target DNA fragment. (**A**) Lane 1, 25-bp DNA ladder molecular marker; Lane 2, STFO alone; Lane 3, LTFO alone; Lane 4, C alone; Lane 5, C + STFO; Lane 6, C + LTFO; Lane 7, E alone; Lane 8, E + STFO, Lane 9, E + LTFO. (**B**) Lane 1, C alone; Lane 2, C + STFO; Lane 3, C + bSTFO; Lane 4, C + LTFO; Lane 5, C + bLTFO; Lane 6, E alone; Lane 7, E + STFO; Lane 8, E + bSTFO; Lane 9, E + LTFO; Lane 10, E + bLTFO.

The gel mobility studies showed that STFO had a large effect on the migration of the DNA target fragments, particularly the centered target (Fig. 30). LTFO, with a longer molecular tether than STFO behaved in the same manner as observed by Akiyama and Hogan, thus validating its use as a negative control (Fig. 30).[108-109] As expected, the broken variants of our tethered-TFOs, bSTFO and bLTFO, were unable to bend DNA also confirming them as acceptable control ligands (Fig. 30).

Are the observed changes due to changes in the shape of the TFO:DNA complexes or due to the increased molecular weight of the complex? Comparison of the TFO complexes with the centered and end TFO target DNA provide an answer. Coupling the two TFO-target DNA fragments with STFO resulted in significantly slower migration rates compared to their respective, unencumbered states. However, there is a

significant difference between the STFO:C complex and the STFO:E complex. Since the molecular weight for both complexes is the same, the difference in migration is best explained by a difference in the conformation of the two complexes.

The angle of the STFO-induced DNA bend in the TFO-target DNA fragments was calculated using equation 1 in which θ is the bend angle, μ_m is the relative mobility of a DNA fragment when the bend is at its center and μ_e is the relative mobility of a DNA fragment when the bend is at the end of the fragment.[109]

$$\theta = 2\arccos\left(\frac{\mu_m}{\mu_e}\right) \qquad \textbf{Equation 1}$$

We calculated the induced bend angles for both the STFO and LTFO tethered-TFOs from the gels pictured in figure 30A (Table II).

Table II: Bending Angle Calculations for Tethered-TFO Induced Bends[a]

Sample	μm	μe	θ
STFO	2.18	2.79	77.23
LTFO	3.39	3.39	0.00

[a] The relative mobilities, μ_m and μ_e, were measured in cm. The bend angles, θ, were calculated with equation 1.

STFO was able to induce a DNA bend of 77.23°, while LTFO was unable to induce any DNA bend. These results were similar to those reported by Akiyama and Hogan; the STFO was able to induce a bend while the LTFO was not.[108,110] There are, however, differences in the degree that STFO bent our DNA target and that reported by Akiyama and Hogan. We calculated a bend angle of 77°, while Akiyama and Hogan determined that STFO induced a bend of about 53°.[108,110] There are several possible reasons for this difference in bend angle. First, our target fragments differed from those used by Akiyama and Hogan. Our 129-bp DNA targets were amplified from our pBend

plasmids using PCR, while Akiyama and Hogan used restriction enzymes to remove their 171 bp targets from their own plasmids.[108-109] Our target fragment was below the persistence length of DNA, while Akiyama and Hogan's fragment was equal to or above the persistence length.

Second, the composition of the spacer region of our respective DNA target fragments is different from that previously used. We used a spacer region that contained only GC base pairs which is reported to be the most flexible spacer of those tried by Akiyama and Hogan.[110] Akiyama and Hogan used a different, less flexible sequence for their calculation of tethered-TFO-induced bend angle.

III.A.5. *Regulation of Gene Expression by Tethered-TFO-Induced DNA Bending*

After verifying that our tethered-TFOs could induce a bend in our DNA target, we next determined whether these artificially induced DNA bends could regulate expression of a reporter gene. We conducted this experiment using an in vitro transcription/translation system. This allowed us to avoid the problems of transporting our TFOs into cells and also avoid the problem of verifying the formation of a triplex in vivo.

We coupled the various tethered-TFOs with the four pBLP plasmids, and then subjected the complexes to *in vitro* transcription and translation. We quantified the luciferase produced by these complexes to determine the effect the tethered-TFOs, and their conformational alterations, on gene expression.

The tethered-TFOs were coupled to linearized pBLP plasmids in a manner analogous to the method used in the gel shift analysis. The pBLP plasmids (pBLP77, pBLP80, pBLP83 and pBLP86) were linearized by digestion with the *Nde* I restriction

enzyme to remove the conformational obstruction presented by supercoiling and thus ease the formation of triplexes between the plasmids and the tethered-TFOs (Fig. 31). The linearized plasmids and tethered-TFOs were combined in a one-to-one molar ratio and allowed to complex overnight at 37°C. Control plasmids were subjected to the coupling conditions in the absence of tethered-TFOs.

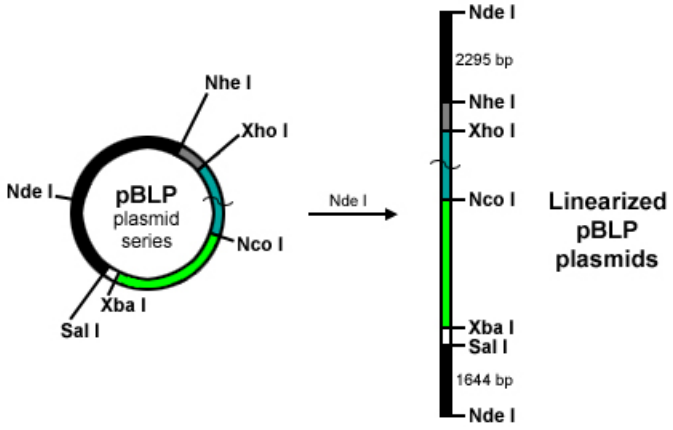


Figure 31. *Nde* I linearization of the pBLP plasmid series. The "~" on the circular and linear plasmids represents the variable region that contains the *malT* promoter.

The luciferase gene of the linearized pBLP plasmid:tethered-TFO complexes was expressed using the EcoPro T7 *in vitro* transcription/translation system system from Novagen. This system utilized T7 RNA polymerase that was compatible with the *malT* promoters present in our pBLP plasmid series. After incubation of our plasmid/tethered-TFO complexes with the EcoPro system for one hour at 37°C, β-galactosidase was added to each reaction as an internal control. Each reaction was then heat denatured and luciferase expression was analyzed by SDS-Page followed by Western analysis using both anti-luciferase HRP-conjugated and anti-β-galactodidase HRP-labeled antibodies. The Western blot was examined by chemiluminence using an ECL Western blotting kit and photographed using the Kodak Digital Science Image Station.

The amount of luciferase produced was determined by measuring the intensities of both the luciferase and β-galactosidase protein bands using the Kodak ID Image software. From these intensities we were able to calculate the amount of luciferase expressed using equation 2; where A_{luc} is the amount of luciferase expressed (ng) in the sample lane, $I_{b\text{-}gal}$ and I_{luc} are the intensities of the β-galactosidase and luciferase bands, respectively, and $A_{b\text{-}gal}$ is the amount of β-galactosidase (ng) added to the sample lane prior to running SDS-PAGE.

$$A_{luc} = \left(\frac{I_{b-gal}}{I_{luc}} \right) A_{b-gal} \qquad \textbf{Equation 2}$$

We compared the amount of luciferase expressed by each plasmid in the presence or absence of the various tethered-TFOs. We calculated the percent change for all of the pBLP plasmids complexed with each of the tethered-TFOs according to equation 3 (Fig. 32).

$$\%\text{Change} = \left(\left[\frac{A_{luc+TFO}}{A_{luc}} \right] 100 \right) - 100 \qquad \textbf{Equation 3}$$

Complexes of pBLP77 and STFO resulted in a 51% reduction of luciferase expression, yet no statistically significant change was observed over basal levels for complexes of pBLP77 with the other three tethered-TFOs (Appendix A). A statistically significant loss in luciferase expression was also seen in pBLP86/STFO complexes. When compared to the luciferase expression of the pBLP77/STFO complex, there was no statistical difference between the expression loss in pBLP86 and pBLP77.

A

Plasmid	STFO	LTFO	bSTFO	bLTFO
pBLP77	-51.3 ± 19.3	-5.9 ± 5.5	-4.1 ± 5.3	-4.0 ± 9.2
pBLP80	31.1 ± 13.3	-1.4 ± 5.9	-1.2 ± 7.5	-0.2 ± 3.4
pBLP83	93.3 ± 23.5	9.8 ± 6.4	10.8 ± 6.6	2.6 ± 5.6
pBLP86	-35.2 ± 4.6	-0.6 ± 3.1	-0.1 ± 2.7	-2.0 ± 4.1

B

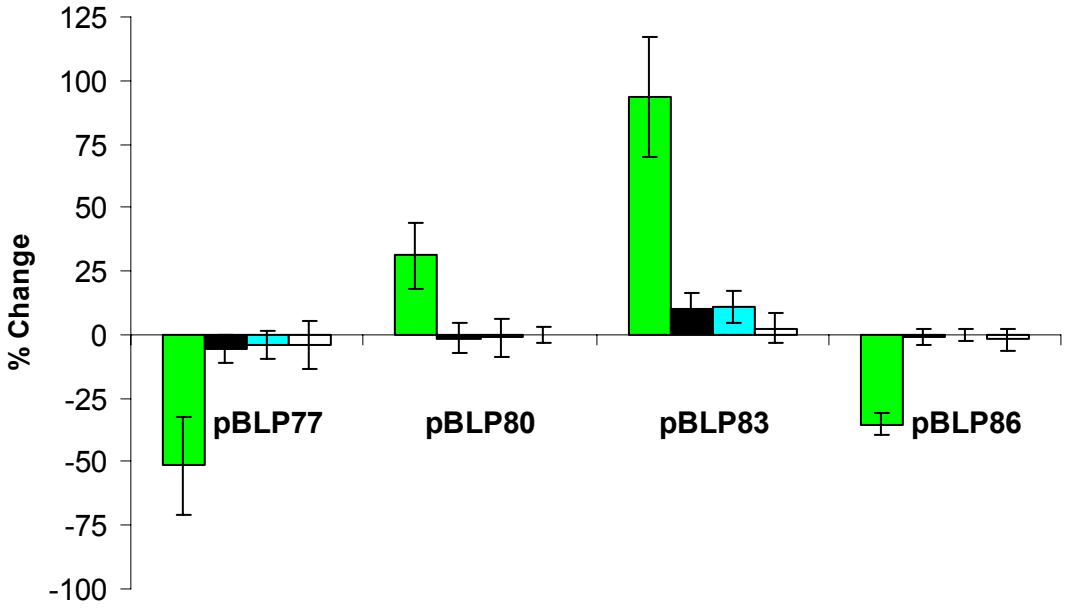


Figure 32. Averaged percent change in luciferase expression from uncomplexed pBLP plasmids. (**A**) Percent change in luciferase expression for each plasmid complexed to a tethered-TFO as compared to the same, uncomplexed plasmid; all values are percentages. (**B**) The green bars represent STFO complexes, the black bars represent LTFO complexes, the blue bars represent bSTFO complexes and the white bars represent bLTFO complexes. These are the averaged results of six assays for pBLP77 and pBLP83 and five assays for pBLP80 and pBLP86.

In contrast to the loss of luciferase expression seen in complexes of pBLP77 and pBLP86, a marked increase in luciferase expression was observed when STFO was complexed with either the pBLP80 or the pBLP83 plasmids. The increased expression levels were elevated to 31.1% and 93.3%, respectively, above the corresponding basal luciferase levels. These elevated expression levels were significantly different from the

decreased luciferase expression levels witnessed in complexes of STFO and either pBLP77 or pBLP86. In contrast, there was no statistical difference in luciferase expression for complexes of pBLP80 and pBLP83 with LTFO, bSTFO and bLTFO. Indeed, when all complexes of LTFO, bSTFO and bLTFO are examined, there is no statistically significant difference between the luciferase expression levels of any of the complexes.

Our data indicates that tethered-TFO-induced DNA bends are able to affect the levels of luciferase expression in the manner that we predicted. STFO, shown to induce a DNA bend in our gel mobility studies, demonstrated the greatest effect when complexed to the pBLP plasmids, whether increasing or decreasing the level of luciferase expression. On the other hand, all of the negative controls, which are unable to bend the DNA, do not significantly affect the expression of luciferase. The lack of a significant effect on luciferase expression by LTFO indicates that the bend induced by STFO, and not some other factor, is the most likely explanation for the observed change in luciferase expression.

The luciferase expression experiment verified the role of DNA bending in artificially regulating gene expression when the results were viewed in terms of bend phasing. The pBLP plasmids were designed with varying distances between the center of the tethered-TFO target sequence and the transcriptional start of the luciferase gene. The variation in this distance correlated to a change of phase in the induced bend of the DNA due to differences in the relationship of these two sites in terms of three-dimensional orientation (Fig. 27). We expected to see an increase in luciferase expression when a bend was introduced in one of the pBLP plasmids with a corresponding decrease in the

pBLP plasmid that was phased opposite it. In addition, intermediary effects were expected for the other two plasmids of the series which are phased in between the plasmids. In our experiment the pBLP83:STFO displayed nearly double the basal expression level of the uncoupled plasmid. On the other hand, pBLP77, with an induced bend opposite the pBLP83 bend, displayed a loss of about half of the luciferase expression seen in the basal expression of the gene in this plasmid. The intermediary pBLP plasmids displayed luciferase expression levels less extreme than those seen in pBLP83 and pBLP77 when bent by STFO.

The results of our gene expression assay were consistent with those of the replacement studies that explored the role of DNA bending in gene transcription.[102-105] As demonstrated in the replacement studies, gene expression was enhanced when artificial DNA bends were introduced in one orientation, in the pBLP83 complex, in a direction presumably analogous to those introduced by native DNA bending proteins.[23,93,102-105] Altering the orientation of the artificially induced DNA bend such that its directionality was opposite that of the enhancing, or in-phase, orientation, resulted in a loss of gene expression. This decrease or loss of gene expression due to an out-of-phase bend was also observed in the replacement studies.[23,93,102,104]

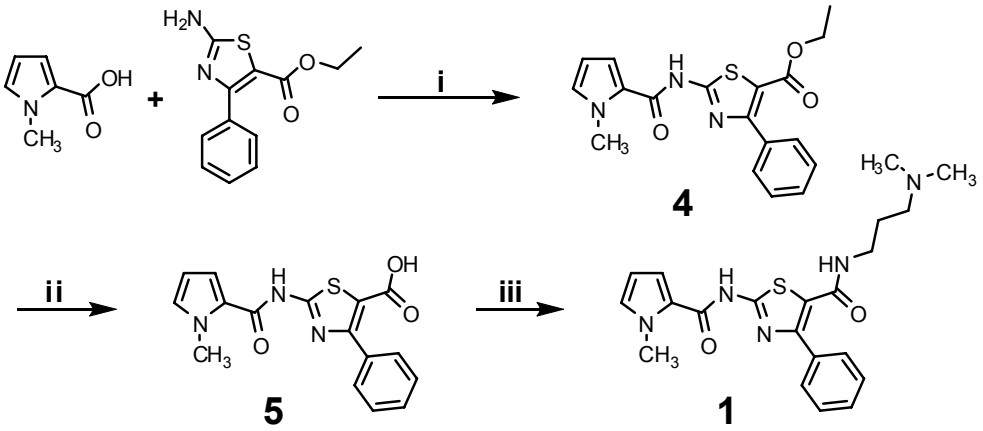
III.B. Project 2: A Sequence Selective Push Bender

We were able to demonstrate that the expression of a target gene could be influenced, both up- and down-regulated, through phased DNA bends induced upstream of the gene by tethered-TFOs. Unfortunately, tethered-TFOs are poor pharmaceutical candidates. Ideally, artificial gene regulation through DNA bending would be

accomplished through small molecules capable of passing through membrane barriers, selectively binding to a specific DNA sequence and inducing a DNA bend upon binding. These small molecules would function as push bending agents; widening the minor groove of a DNA target upon binding. In the second project of this thesis we synthesized three DNA push bending agents (Fig. 24), tested them for DNA binding preference and the ability to induce a bend in a DNA target.

III.B.1. *Synthesis of 2-[(1-Methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (3-dimethylamino-propyl)-amide, 1*

Scheme I[a]



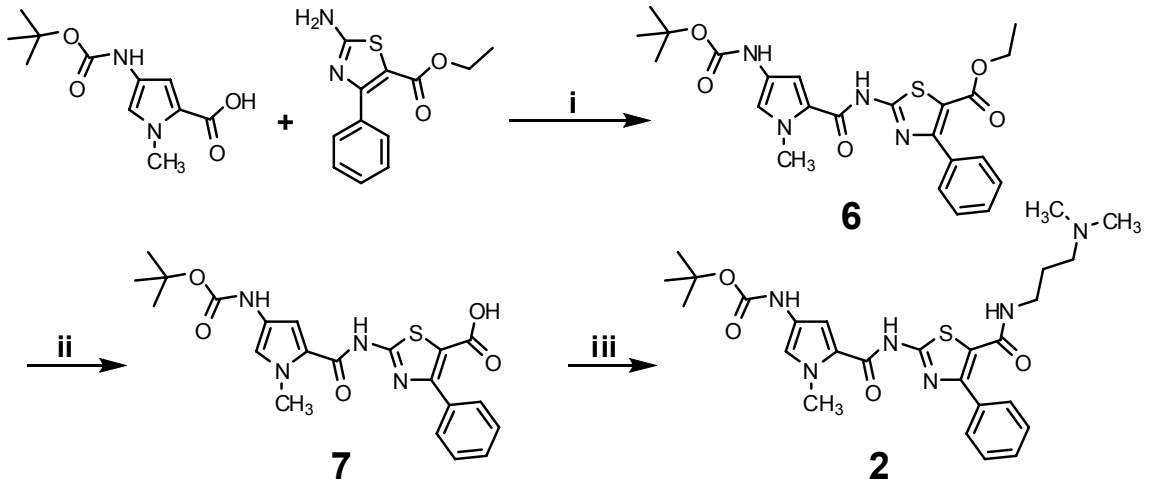
[a]Reagents: (i) EDCI, DMAP, DMF, room temperature, 62% yield; (ii) LiOH/methanol, 60°C, 75% yield; (iii) 3-(dimethylamino)propylamine, HBTU, NMM, DMF, 51% yield.

The construction of **1** (Scheme I) was accomplished by condensation of *N*-methylpyrrole-2-carboxylic acid and ethyl-2-amino-4-phenyl-5-thiazole carboxylate using standard peptide bond coupling conditions (EDCI, DMAP) as described by Boger *et al*.[129] The resulting product, **4**, was hydrolyzed with 1N LiOH to generate the free carboxylic acid, **5**, in good yield. The addition of the necessary cationic tail was again performed using standard peptide coupling conditions. We found that is was necessary to use the more reactive HBTU and NMM as coupling conditions for the formation of the

final peptide bond.[130] Purification of the final product, **1**, using standard silica gel chromatography provided the desired material in 51% yield.

III.B.2. *Synthesis of {5-[5-(3-Dimethylamino-propylcarbamoyl)-4-phenyl-thiazol-2-ylcarbamoyl]-1-methyl-1H-pyrrol-3-yl}-carbamic acid tert-butyl ester, 2*

Scheme II[a]

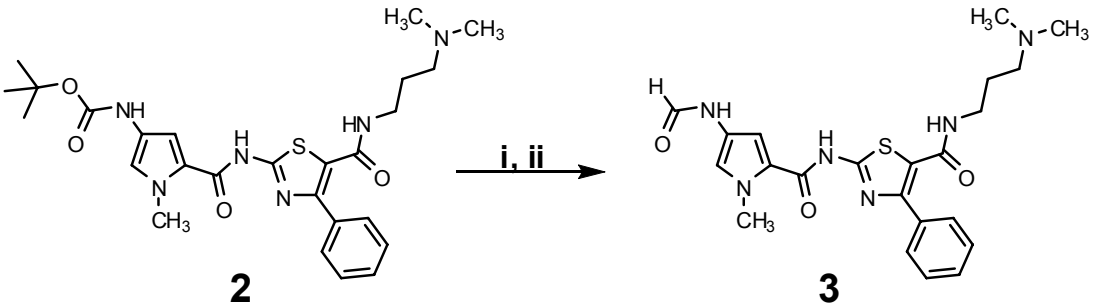


[a]Reagents: (i) EDCI, DMAP, DMF, room temperature, 52% yield; (ii) LiOH/methanol, 60°C, 72% yield; (iii) 3-(dimethylamino)propylamine, HBTU, NMM, DMF, 60% yield.

Condensation of 4-tert-butoxycarbonylamino-1-methyl-1H-pyrrole-2-carboxylic acid and ethyl-2-amino-4-phenyl-5-thiazole carboxylate using EDCI, DMAP generated the desired product, **6** in 52% yield (Scheme II). Hydrolysis of the ethyl ester with lithium hydroxide generated the acid, **7**, which was reacted with 3-(dimethylamino) propylamine under conditions similar to those for the preparation of **1** to yield the final product **2** in 60% yield.

III.B.3. ***Synthesis of 2-[(4-Formylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (3-dimethylamino-propyl)-amide, 3***

Scheme III[a]



[a]Reagents: (i) CH_2Cl_2, benzenethiol, TFA, room temperature; (ii) Ethanol, ethyl formate, reflux, 73% yield.

The formamide derivative of **2**, compound **3**, was obtained as outlined in scheme III. Deprotection of the Boc group with trifluoroacetic acid and benzenethiol yielded the free amine which was immediately reacted with ethyl formate to yield **3** in 73% overall yield.

III.B.4. ***DNA Binding of Compounds 1, 2 and 3***

We used an ethidium bromide displacement assay to determine the binding affinities and sequence preference of compounds **1**, **2** and **3**. Ethidium bromide (EtBr) fluoresces when bound to DNA, but possesses only weak fluorescence in the absence of DNA. Thus, displacement of bound EtBr from DNA by other DNA binding agents would result in a decrease in the fluorescence signal. By measuring changes in the fluorescence as a function of compound concentration, we were able to determine the binding constants for each compound for each DNA target.

In this assay, we investigated the ability of our compounds to bind to three different DNA targets; to an AT-rich sequence of DNA, a GC-rich sequence of DNA and

*Sca*I-linearized pUC19 plasmid. The AT-rich and GC-rich DNA target sequences were generated from an oligonucleotide that could adopt a hairpin structure (Fig. 33). The *Sca* I-linearized plasmid was created by digestion of pUC19 with *Sca*I restriction enzyme and was used as a representation of a random sequence of AT and GC base pairs.

A

5′-CG**AAAAA**C A A
3′-GC**TTTTT**G A A

B

5′-CG**GGCGG**C A A
3′-GC**CCGCC**G A A

Figure 33. The DNA hairpin targets investigated in the ethidium bromide displacement assay, (**A**) the AT target and the (**B**) GC target.

The binding constants for compounds **1**, **2** and **3** utilized a method published by Boger, et al.[129] The collected data was normalized to the background fluorescence of ethidium bromide in assay buffer, while the maximum fluorescence (100%) was set equal to the fluorescence of ethidium bromide bound to the various dsDNA targets prior to titration of the compounds. The data collected as a function of the concentration of the individual compounds were converted into a percentage of this maximum fluorescence, after being normalized to the background (Fig. 34).

The binding constants for compounds **1**, **2** and **3** were found using equation 4; where $K_{compound}$ is the binding affinity of the drug to the target DNA, K_{EB} is the binding coefficient for ethidium bromide for the DNA targets, [EB] is the concentration of ethidium bromide present, and $[compound]_{50}$ is the concentration of drug that reduced the fluorescence of the ethidium bromide/DNA complexes by 50%.

$$K_{compound} = \frac{K_{EB}[EB]}{[compound]_{50}} \qquad \textbf{Equation 4}$$

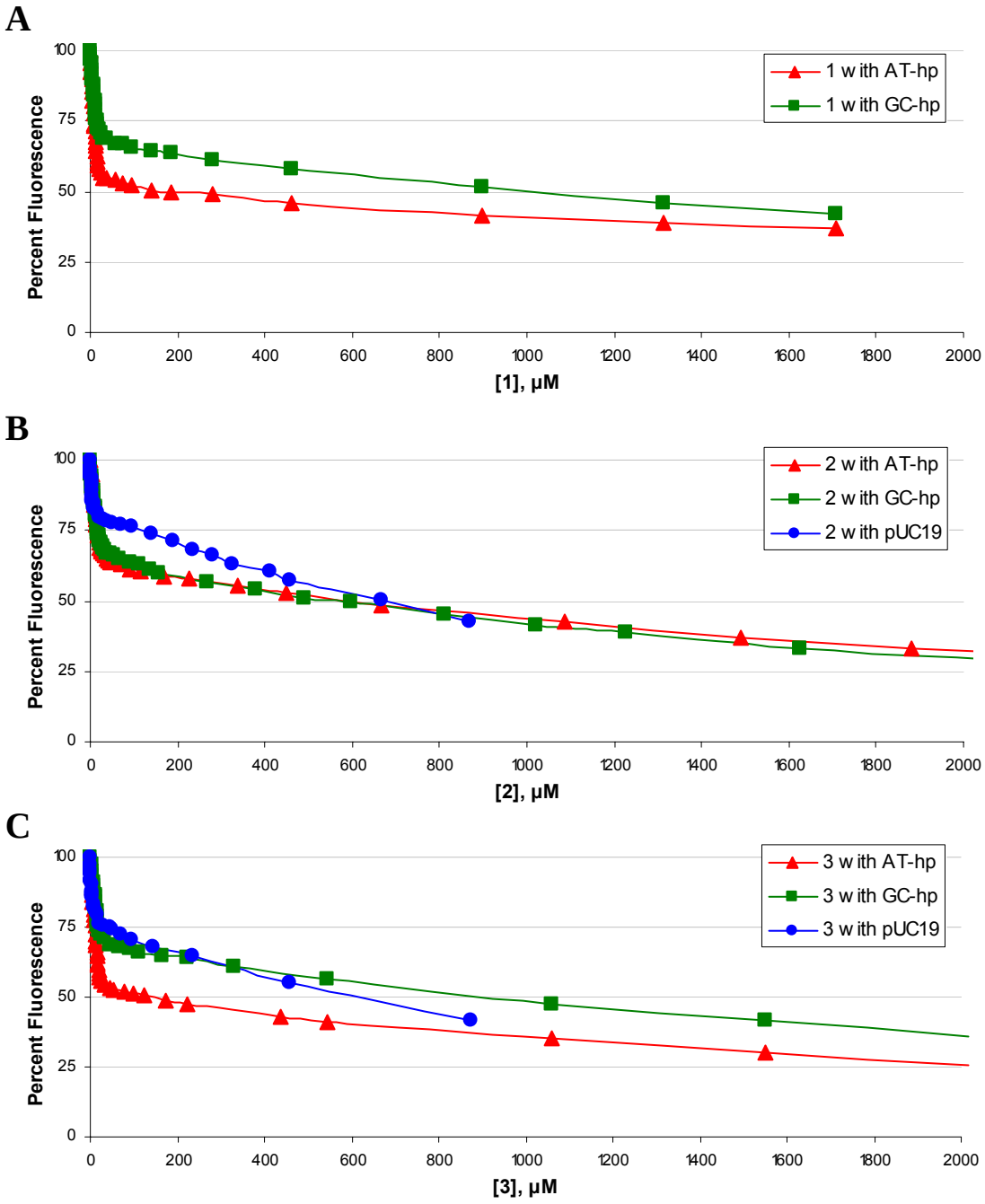


Figure 34. Plot of the change in fluorescence of ethidium bromide bound DNA targets in the presence of increasing concentrations of our compounds; AT-hairpin, ▲; GC-hairpin, ■; and the linear pUC19 plasmid, ●. (**A**) Compound **1**. (**B**) Compound **2**. (**C**) Compound **3**.

Table III: Binding Constants of Compounds **1**, **2** and **3**

Compound & Target [a]	$[Compound]_{50}$ (x 10^{-4} M)	$K_{compound}$ (x 10^{6} M^{-1})
Netropsin + 5'-AATT-3'	-	[b]25.0
Netropsin + 5'-ATAT-3'	-	[b]2.2
Distamycin + AT-hp	-	[c]6.5
1 + AT-hp	1.6720	0.026
1 + GC-hp	9.8650	0.004
2 + AT-hp	5.8335	0.008
2 + GC-hp	5.7950	0.008
2 + pUC19	6.7405	0.007
3 + AT-hp	1.3965	0.032
3 + GC-hp	9.0350	0.005
3 + pUC19	6.0900	0.007

[a]: hp represents hairpin DNA. [b]: Sidorova et al., 1995. [c]: Boger et al., 2000.

For our data, we used K_{EB} of 10 x 10^{6} M^{-1} for all calculations with [EB] at 4.4 x 10^{-6} M.[129] The calculated binding constants are presented in table III along with the binding constants of netropsin for two different AT target sequences and for distamycin for the AT-rich DNA hairpin.[48,129] Compounds **1** and **3** displayed binding preferences for the AT-rich hairpin DNA over the other DNA targets as demonstrated by their binding constants. Compound **2**, on the other hand, showed no preference for a single DNA target, binding to all three at comparable levels, all of which were less than that of compounds **1** and **3** toward the AT-rich hairpin. Compounds **1**, **2** and **3** all displayed binding constants two or three orders of magnitude less than a related compound, netropsin, towards the DNA targets.

Our compounds were designed with the goal of creating sequence specific DNA minor groove push bending molecules. The ethidium bromide displacement assay verified that compounds **1** and **3** bound to DNA with a noticeable level of sequence specificity. Compounds **1** and **3** displayed a five to six-fold greater preference for AT-rich DNA sequences over GC-rich and random DNA sequences. The sequence indiscriminate, low level DNA binding displayed by compound **2** was most likely due to

the tert-butoxycarbonylamino group at the head of the compound and the steric hindrance that it presented, preventing the compound from binding in the minor groove at any site. The lower binding constants, as compared to netropsin and distamycin, of compounds **1** and **3** could be related to the inclusion of the phenyl ring at the 4-position of the thiazole. The bulk of the ring may have dramatically effected the positioning of our compounds in the minor groove and the ability of the minor groove binding edge of the compounds to fully interact with the floor of the minor groove. Other thiazole-containing lexitropsins displayed an inability to bind to DNA, let alone AT-rich DNA sequences, though the majority of these compounds contained a substituted thiazole ring configured such that the nitrogen atom was on the minor groove binding edge of the compounds.[130]

III.B.5. *DNA Bending by Compounds 1, 2 and 3*

We determined the ability of compounds **1**, **2** and **3** to alter the conformation of DNA upon binding using fluorescence resonance energy transfer (FRET) analysis. FRET measures the distance between two fluorescent dyes, located at the 5'- and 3'-ends of a DNA target. Changes in the conformation of DNA, due to bending, should result in a change in the FRET signal between the two dyes. Thus, changes in the FRET signal upon drug binding would be indicative of DNA bending. We compared the fluorescence levels of two identical dsDNA sequences end-labeled with different dyes in the absence and presence of compounds **1**, **2** and **3** (Fig. 35).

The target FRET DNA sequence was 19-bp in length with an AT-rich central sequence. This target sequence was selected due to the AT sequence preference demonstrated by the compounds in the ethidium bromide displacement assay. Fluorescein and tetramethylrhodamine were selected as our FRET dyes as these dyes

A

```
5′-(6-FAM)GATGCCGAAATTCCGCTTC-3′
3′-       CTACGGCTTTAAGGCGAAG-5′
```

B

```
5′-(6-FAM)GATGCCGAAATTCCGCTTC       -3′
3′-       CTACGGCTTTAAGGCGAAG(TAMRA)-5′
```

Figure 35. The two fluorescently labeled DNA fragments for FRET analysis. (**A**) This fragment contains only the FAM fluorescence donor moiety, while (**B**) contains both the fluorescence donating FAM moiety but also the fluorescence accepting TAMRA moiety.

have been used by other research groups investigating conformational changes in DNA.[131] The first of our DNA target sequences, the donor fragment, possessed a single, 5'-terminal fluorescein moiety, FAM. Our other target DNA fragment, the donor-acceptor fragment, was labeled with a 5'-terminal FAM as a FRET donating group and a 3'-terminal tetramethylrhodamine, TAMRA, as a FRET acceptor.

Samples of each fragment were excited at 495 nm in the absence of the compounds and the emission spectra for each were recorded from 500 nm to 700 nm. The emission spectra for both the donor sequence and the donor-acceptor sequence can be seen in figure 36. While the emission spectrum for the donor sequence contained a FAM peak at about 520 nm, the spectrum of the donor-acceptor target contained two peaks; one at about 520 nm corresponding to FAM and a second, smaller peak at about 580 nm corresponding to energy transfer to TAMRA. The FAM peak of the donor-acceptor sequence spectrum was smaller than that in the donor sequence spectrum due to fluorescence resonance energy transfer. After analysis of the two DNA sequences alone, excess amounts of compounds **1**, **2** and **3** were added, incubated and then excited at 495 nm. The emission spectra from 500 nm to 700 nm were then recorded (Fig. 36).

Changes in the conformation of the DNA targets were calculated based on the differences in the spectra of the DNA target in the absence and presence of the compounds. The bending caused by the binding of our compounds was calculated by comparing the conformations of the FRET DNA target fragments with and without the compounds. To calculate these conformations, the efficiency of the energy transfer from the donor moiety to the acceptor moiety was determined using equation 5.[131] In this equation E was the efficiency of energy transfer; Φ_{em}^{D} was the integrated intensity of emission from 510 to 530 nm of DNA duplex with only the donor moiety; Φ_{em}^{DA} was the integrated intensity of emission from 510 to 530 nm of donor-acceptor labeled DNA duplex.

$$E = 1 - \left(\frac{\Phi_{em}^{D}}{\Phi_{em}^{DA}} \right) \qquad \textbf{Equation 5}$$

The calculated efficiencies of energy transfer are shown in table IV. The efficiency of energy transfer decreased when any of the compounds were bound to the DNA, though this loss in efficiency was greater for compounds **1** and **3** than compound **2**. The energy transfer between the FAM donor and the TAMRA acceptor was less efficient when the FRET DNA target was complexed with any of our compounds.

Table IV: FRET Derived Bending Data for Compounds **1**, **2** and **3**

Compound	E[a]	R[b] (Å)	θ[c] (°)	Δθ[d] (°)
No compound	0.249 ± 0.015	58.9 ± 0.7	55.3 ± 2.7	0.0
1	0.156 ± 0.004	65.0 ± 0.3	24.8 ± 2.8	-30.6
2	0.217 ± 0.008	60.7 ± 0.5	48.3 ± 2.1	-7.0
3	0.162 ± 0.017	64.5 ± 1.3	28.5 ± 8.7	-26.8

[a]: Efficiency of energy transfer; [b]: End-to-end distance; [c]: Bend angle from linear, where linear is 0°; [d]: Change in bend angle from FRET DNA target fragment without bound compound.

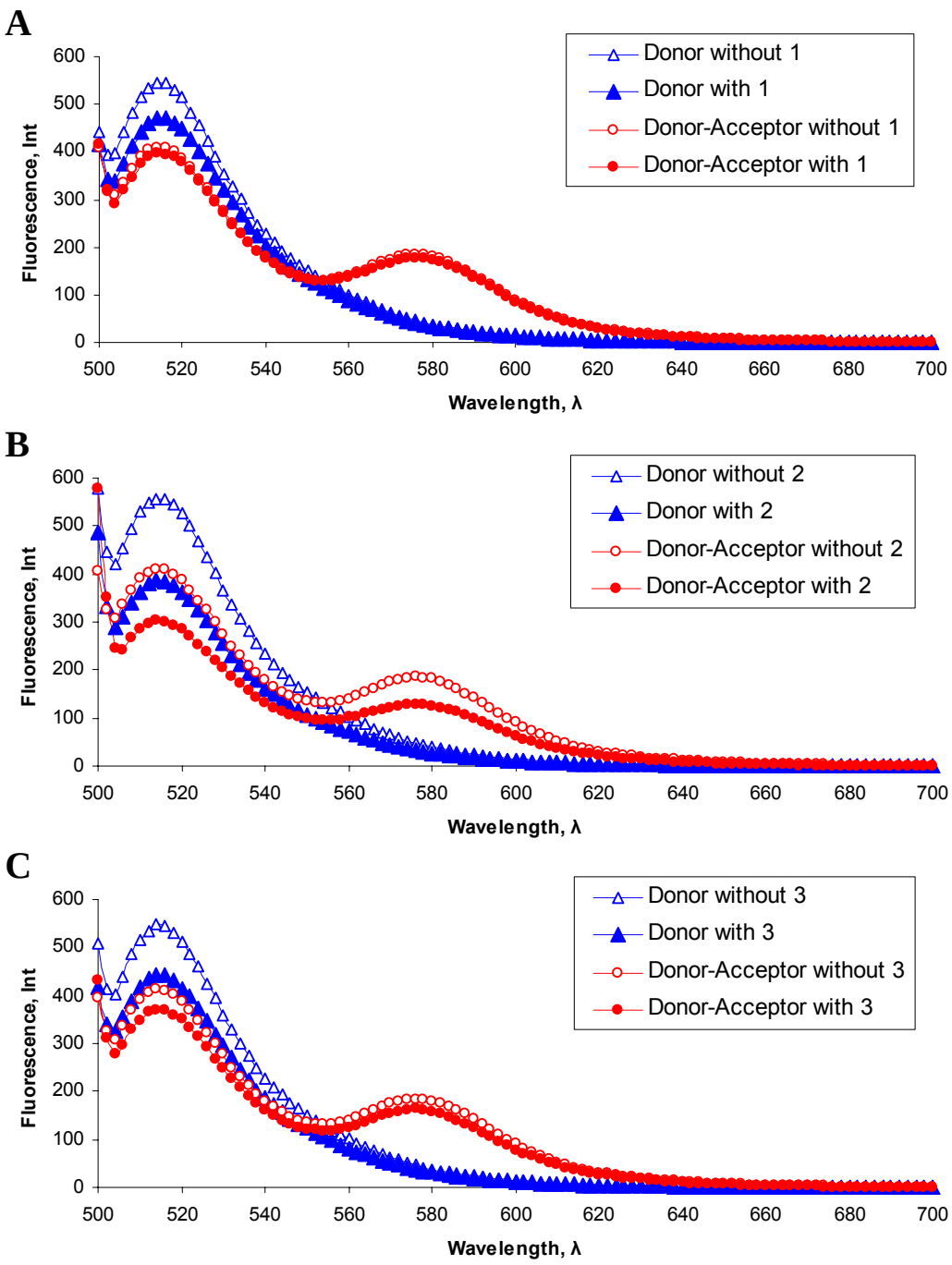


Figure 36. FRET spectra of donor fragment and donor-acceptor fragment in the absence and presence of the compounds; the donor fragment in the absence of the compounds (Δ) and in the presence of the compounds (▲), and the donor-acceptor fragment in the absence of the compounds (○) and in the presence of compounds (●). (**A**) Compound **1** data. (**B**) Compound **2** data. (**C**) Compound **3** data.

The end-to-end distance of the FRET DNA fragments is indirectly related to the efficiency of energy transfer. The more efficient the energy transfers within a DNA fragment, the closer together the ends of the fragment. The relationship between the efficiency of energy transfer and the distance between the two fluorescent dyes is given by equations 6 and 7.[131] In these equations, E represents the efficiency of energy transfer; R represented the end-to-end distance of interest; and R_0 was the "critical distance" at which energy transfer from donor to acceptor and spontaneous decay of donor is of equal probability. R_0 was a constant and is dependent upon the FRET pairs chosen. For our system, R_0 was 49 Å.

$$E = \frac{1}{\left[1 + \left(\frac{R}{R_0}\right)^6\right]} \qquad \textbf{Equation 6}$$

$$R = R_0\left(\sqrt[6]{\left(\frac{1}{E}\right) - 1}\right) \qquad \textbf{Equation 7}$$

The end-to-end distances calculated by equation 7 are shown in table IV. The maximum end-to-end distance for the two DNA sequences used in this FRET analysis, representing a completely straight strand of DNA, was 66.5 Å, based on the accepted average distance of 3.5 Å/DNA bp. The calculated end-to-end distances showed an unexpected trend. Upon addition of compounds **1**, **2** and **3** the end-to-end distances of the target DNA sequences did not decrease, as expected, but rather increased approaching the maximum length of these DNA sequences. The increased end-to-end distances and decreased efficiencies of energy transfer suggested that the FRET DNA targets were

becoming more linear and, therefore, less bent. This appears to be opposite of our original goal.

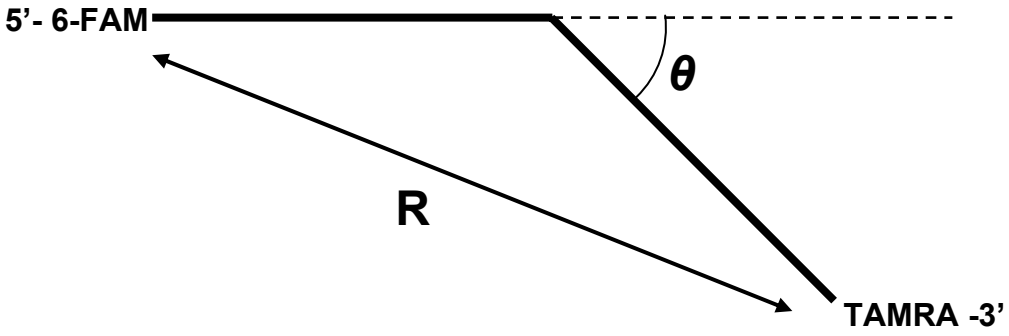


Figure 37. Relationship of end-to-end distance, R, of the FRET DNA fragments and angle from linear, θ. Decreasing distance R results in an increase in the θ angle and, conversely, increasing distance R results in a decrease in angle θ. The maximum distance for R was the linear length of the DNA fragment, 66.5 Å, with a θ angle of 0°.

The end-to-end distance of the FRET DNA in the presence and absence of our compounds allowed us to calculate the bend angles that were present in the various complexes (Fig. 37). As the end-to-end distance of the DNA fragments decreased, the fragment becomes more bent in relation to linear DNA. We assumed that our compounds were exerting their effects from the direct center of the FRET DNA target fragments, the center of the 5'-AAATT-3' binding site, and we set the linear length of the FRET DNA to be 66.5 Å(19-bp at 3.5 Å/bp). From these assumptions, we used equation 8 to calculate the bend angle relative to linear DNA. In equation 8, θ is the bend angle as seen in figure 37, R is the end-to-end distance of the DNA and R_0 is the maximum linear length of the DNA fragment (66.5 Å).

$$\theta = 2\cos^{-1}\left(\frac{R}{R_0}\right) \qquad \textbf{Equation 8}$$

The unliganded DNA displayed a bend greater than any of the complexes with our compounds. The bend in the DNA target fragment was calculated to be 55° from linear. This bend was sequence derived and, therefore, intrinsic in nature. The AT-rich target sequence located in the center of the DNA target was the most likely source of the intrinsic bend due to both the increased flexibility of AT-rich sequences and the steric interactions between the adjacent adenines causing a bend in the DNA as in an A-tract, widening the major groove and bending the DNA toward the minor groove.[72-73,78-80]

Our compounds demonstrated a lower efficiency of energy transfer, a greater end-to-end distance and a corresponding decrease in bend angle. The change in bend angle, $\Delta\theta$, was calculated as the difference between the bend angles in the presence versus the absence of our compounds (table IV). As seen in table IV, compounds **1** and **3** appeared to have straightened the bend of the DNA target sequence by about 30°. Compound **2**, which possess poor DNA binding ability, altered the bend of DNA by only about 7°.

Our compounds were designed as minor groove-binding push bending agents. However, complexes of DNA and our compounds resulted in more linear DNA fragments, as opposed to DNA fragments with DNA bending as expected. In order for the statically bent sequence of DNA to obtain a more linear conformation a force opposing the adenine-derived bend has to be applied to the DNA. It is known that the adenine-derived DNA bend is caused by a compression of the minor groove of DNA. Our compounds bind to the minor groove and widen it through interactions between the minor groove walls and the phenyl ring present in our compounds. Pushing the minor groove open in the intrinsically bent FRET DNA target, however, would result in a bend

in opposition to the naturally occurring bend at the target site and would ultimately result in forcing the DNA into a more linear conformation.

IV. DISSCUSSION

IV.A. Project 1: Tethered-TFOs and Gene Regulation

The objective of this project was to examine whether artificially-induced DNA bends could affect gene expression. Our studies showed that STFO was able to bend DNA and regulate gene expression to a far greater degree than the other tethered-TFOs tested (Fig. 32). The effect of this regulation was dependent upon the distance between the bend and the start of the luciferase gene. When STFO was coupled to pBLP77, the plasmid with the shortest distance between the center of induced bend and the start of the gene (77-bp), the greatest loss in gene expression was observed. However, when the center of the bend was moved 6-bp farther (pBLP83), the induced bend was moved to the opposite face of the DNA. At this position, the amount of luciferase expressed was nearly doubled. This dichotomy was expected, confirming our hypothesis that non-protein driven, DNA conformational changes alone could influence the expression of a gene.

The data from the gene expression assay indicated that changing the distance between the center of the tethered-TFO-induced DNA bend and the start of the luciferase gene would result in a change in level of gene expression. Plotting the percent change in luciferase expression as a function of the distance, in base pairs, between the center of the tethered-TFO-induced DNA bend and the start of the luciferase gene, it is apparent that

there is a cyclic function to the data (Fig. 38). It is interesting to note that the same cyclic function is observed for the relationship between distance and bend angle for intrinsically bent DNA and is reminiscent of plotted gel mobility data for DNA fragments with variably positioned DNA bends.[109] This relationship lends credence to the fact that DNA bending is directly related to gene expression.

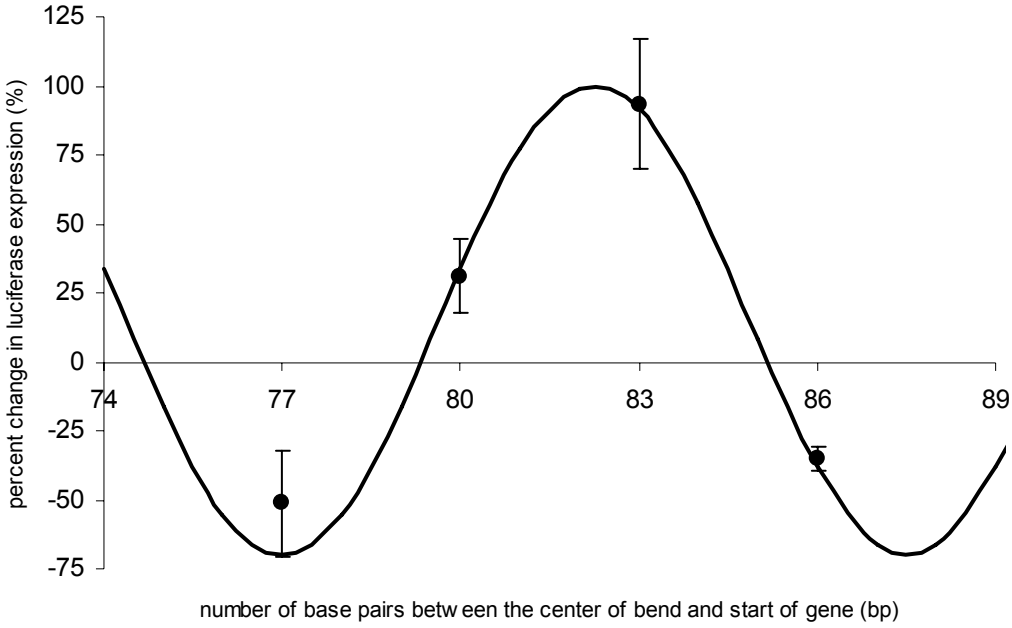


Figure 38. Graph depicting the experimental percent change in luciferase expression versus the distance, in base pairs, between the center of the tethered-TFO induced DNA bend and the start of the luciferase gene (black circles) and the curve from equation 10 showing changing expression levels as a function of bend phase as a result of distance (black line).

The data outlined in figure 38 is best fit with a sine function. The general formula for a sine curve is shown in equation 9. The result y is a function of x where A is the amplitude of the sine curve, w is the period of repetition (in radians), ϕ is the phase shift of the curve and B is the vertical shift of the sine curve.

$$y = A\sin(wx - \phi) + B \qquad \textbf{Equation 9}$$

From equation 9 and the gene expression assay data (Fig. 32), equation 10 was derived to describe the change in luciferase expression as a function of the distance between the induced DNA bend and the luciferase gene.

$$E = 85\sin[0.5984(b - 77) - 1.5708] + 15 \qquad \textbf{Equation 10}$$

In equation 10, *E* is the percent change in luciferase expression and *b* is the distance in base pairs from the center of the tethered-TFO target sequence to the start of the luciferase gene. The amplitude of the equation, *A*, was estimated to be about 85; the period of the equation, *w*, was based on 10.5-bp/turn, and was equal to $2\pi/10.5$ or about 0.5984; the vertical shift, *B*, was calculated to be 15; and the phase shift value, ϕ, was calculated to be 1.5708. The phase shift value, ϕ, was set such that the trough of the curve fell at 77-bp, the low point of the experimental data. Equation 10 best fit the experimental data and fell within the experimental error for each data point.

Based on equation 10, the greatest loss in luciferase expression, -70%, would occur when the center of the induced bend was positioned 77 base pairs upstream of the start of the luciferase gene; and every 10.5-bp increment from this point. On the other hand, the greatest increase in luciferase expression, +100%, would occur at 82.25-bp. Since no base pair can occur at this point, it is impossible to achieve a 100% increase in gene expression. The maximum achievable increase in luciferase expression would occur at a distance of 82 base pairs between the center of the induced bend and the luciferase gene with a +99.1% increase in gene expression.

We have demonstrated that artificially-induced, phased DNA bends were able to alter gene expression levels. How these induced DNA bends were influencing gene expression remains to be answered. There are two theories that can explain how DNA bending can change gene expression. These theories are DNA wrapping and RNA polymerase interactions with proteins or DNA upstream of the start of the gene. Both mechanisms have been described in the literature.

In both prokaryotic and eukaryotic systems, RNA polymerase has been shown to wrap DNA around itself after it binds to the promoter. Previous studies have shown that the DNA wraps about 300° around the polymerase. To accomplish this, about 90 consecutive base pairs are needed to come into contact with the RNA polymerase. Footprinting studies have shown that the base pairs range from the -70 to the +20 position.[83,94-96]

Induced DNA bends could have affected DNA wrapping either facilitating or hindering the wrapping around RNA polymerase. The major difference between the various TFO:DNA complexes was whether a DNA bend had formed and what was the orientation of that bend. Both of these factors could affect DNA wrapping. The out-of-phase, STFO-induced bends seen in complexes with pBLP77 and pBLP86 plasmids may have altered the conformation of DNA such that it was unable to wrap around RNA polymerase (Fig. 39A). The end result of this would be a weaker RNA polymerase:DNA complex, which would lead to a decrease in the efficiency of transcription. Conversely, the increased levels of luciferase expression witnessed in the pBLP80:STFO and pBLP83:STFO complexes could also be explained by facilitating DNA wrapping around the RNA polymerase (Fig. 39B). Facilitating DNA wrapping would decrease the energy

required, which would enhance binding of RNA polymerase and thus increase transcription. The other TFOs that have no effect on DNA conformation, LTFO, bSTFO and bLTFO, should have no effect on either RNA polymerase binding or on DNA wrapping (Fig. 39C).

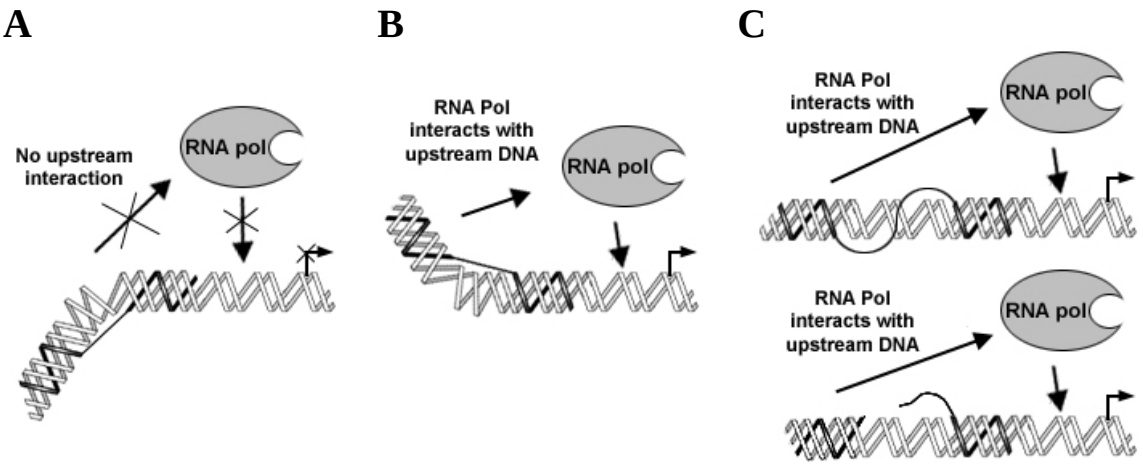


Figure 39. Role of tethered-TFO induced DNA bending in DNA wrapping. (**A**) Out-of-phase bends prevent DNA wrapping around RNA polymerase resulting in the down regulation of luciferase expression. (**B**) In-phase bends facilitate DNA wrapping that enhances luciferase expression. (**C**) Tethered-TFOs that did not bend DNA allow DNA wrapping to occur normally with baseline levels of luciferase expression.

Another mechanism by which tethered-TFO induced DNA bends could have affected the luciferase expression involved the inhibition or activation of protein-protein or protein-DNA interactions to RNA polymerase (Fig. 40). Again, the overall effect on gene expression would be dependent on the orientation of the bend relative to the start of the gene. An in-phase bend would facilitate interactions with upstream proteins or DNA, which in turn would increase gene expression. Conversely, an out-of-phase bend would hinder interactions and result in the down regulation of luciferase expression. Tethered-TFOs that did not affect the conformation of DNA would have no affect on the occurrence of these interactions and, consequently, luciferase expression would occur at

the baseline level. Although we cannot definitively rule out this mechanism, we believe that it is high unlikely to occur in our system for two reasons. First, our vectors were not designed with binding sites for transcription factors. Second, the EcoPro T7 transcription/translation system lacks these transcription factors due to the kit being designed to function with only a T7 or *E. coli* promoter.

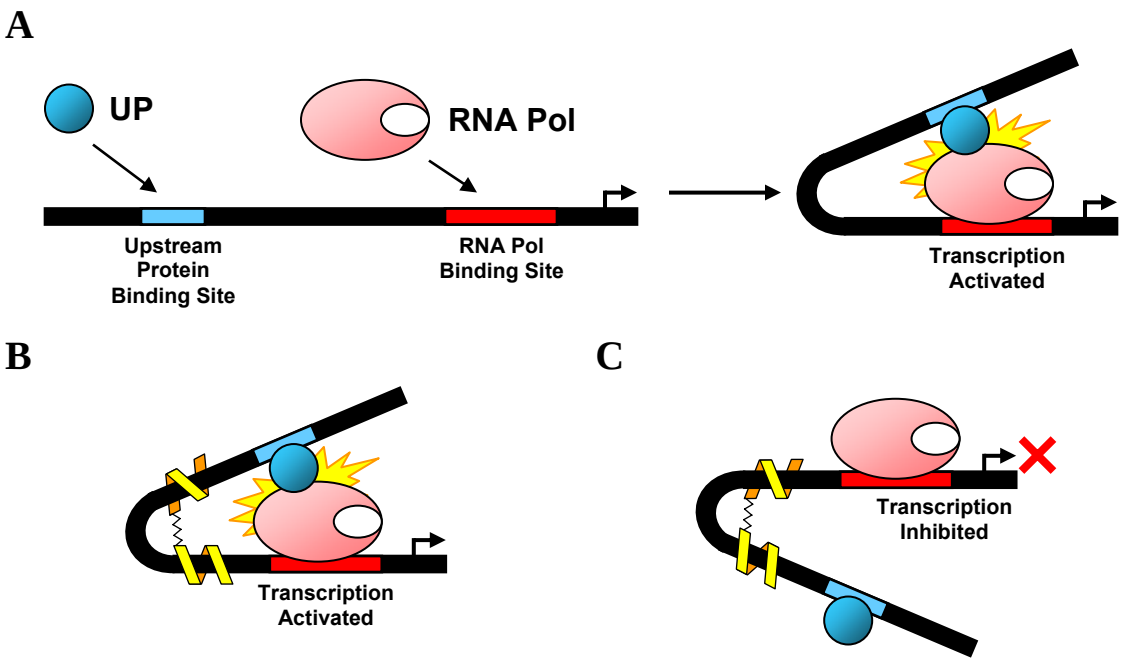


Figure 40. Role of phased DNA bends in facilitating interactions between RNA polymerase, RNA Pol, and upstream proteins, UP. (**A**) Linear DNA without tethered-TFOs or with non-bending tethered-TFOs display normal levels of transcription. (**B**) In-phase bends facilitate contact between RNA Pol and UP, up activating transcription. (**C**) Out-of-phase bends inhibit transcription by preventing interaction between the proteins.

Regardless of how tethered-TFO induced DNA bends were affecting gene expression, these results demonstrate that the induction of a DNA conformational change upstream of a gene can alter gene expression. This method of gene control presents a single mechanism by which a gene can be either activated or inhibited, dependent on the location and subsequent phase of an induce bend. This mechanism offers the potential for the development of a new class of gene expression controlling pharmaceutical agents.

Our study demonstrated the ability to alter gene expression with the induction of phased DNA bends upstream of RNA polymerase bind. The future development of pharmaceutical agents for the treatment of gene expression related diseases will require further exploration of this mechanism in eukaryotic systems. Nucleosome packaging plays an important role in eukaryotic transcription and will have to be dealt with for gene expression to be altered by artificially induced DNA bends.[132-133]

IV.B. Project 2: A Sequence Selective Push Bender

The tethered-TFO bending project verified that a non-protein moiety could induce a bend in DNA and affect the expression of the gene, validating the artificial induction of phased DNA bends as a mechanism to alter gene expression. It is unlikely, however, that tethered-TFOs themselves would be useful medicinal agents due to their size and complexity. The ideal agent would be a small molecule that could bind to a specific sequence of DNA and induce a bend to alter gene expression in the manner demonstrated by the tethered-TFOs. Thus, the objective of the second project was to explore the design of small molecule, sequence specific DNA bending agents.

Currently, there are a number of compounds that are capable of binding to DNA and some of these display sequence selectivity. Among the most selective DNA binders are the polyamide and lexitropsin classes of compounds. These compounds interacted with hydrogen bond donors and acceptors presented by nucleotides in the minor groove of the DNA duplex. Polyamides have served as useful molecules to delivery other functions to DNA and thus were logical choices for the construction of DNA bending agents.

In contrast to the minor groove binding agents, there are relatively few molecules that are capable of changing the conformation of DNA and most of them display poor sequence selectivity. However, we used some of the design principles found in existing DNA bending agents like ET-743 to construct our agents. ET-743 induces a conformational change in DNA through steric interactions between minor groove walls of the binding site and the perpendicularly displayed bulky aromatic ring system of the molecule. The bulk of the ring system widened the minor groove causing a corresponding compression of the major groove on the opposite face of the DNA, resulting in a DNA bend.

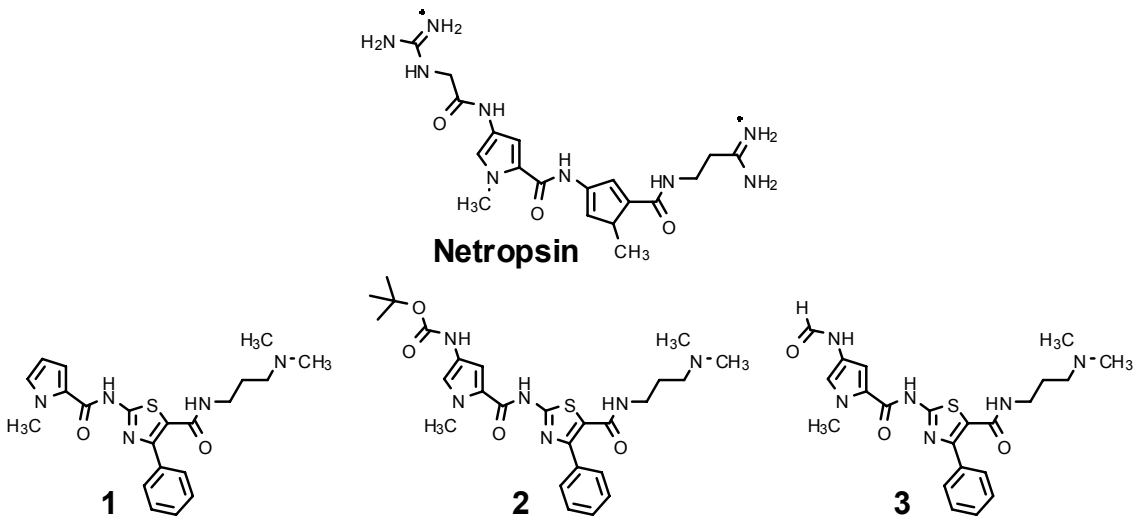


Figure 41. Netropsin and the three analogues; **1**, **2** and **3**.

We designed a series of compounds that incorporated the minor groove binding specificity of polyamides with the DNA bending mechanism of compounds like ET-743. Compounds **1**, **2** and **3** included structural features of the minor groove binding agent netropsin and like netropsin, our molecules bound to AT-rich DNA sequences (Fig. 41). Compounds **1-3** were also designed to present bulky group into the minor groove of their target in a manner analogous to ET-743. We chose to use a phenyl ring since the phenyl

group would adopt a perpendicular orientation relative to the planar shape of the rest of the molecule which would allow it to sterically interact with walls of the minor groove. The compounds were synthesized in a straightforward manner analogous to other known minor groove binding agents.

An ethidium bromide displacement assay was used to determine the binding constants and sequence preferences of our compounds. Compounds **1** and **3** displayed the expected preference for AT-rich DNA sequences with **1** displaying about a 7-fold preference for the AT-rich target while **3** showed a 6-fold preference. Compound **2** demonstrated similar binding constants for all three DNA targets indicating that it was most likely a non-specific binding agent. The presence of a bulky group in the N-terminal position in compound **2** must have been great enough to prevent the compound from correctly positioning itself in the minor groove of an AT base pair.

The observed binding constants for all three compounds were 300 to 2000-fold less than the reported binding constant of netropsin for AT-rich DNA.[48] Our compounds differ from the model molecule, netropsin, by the substitution of 4-phenyl-thiazole for an *N*-methylpyrrole and utilization of different N- and C-terminal tail groups. The discrepancy in binding affinities between that reported for netropsin and our compounds could be attributed to these replacements.

The thiazole of our compounds incorporated a sulfur atom in place of a vinyl at the 1 position of the ring. The sulfur atom was larger than the original carbon atom and may have prevented optimal alignment of our compounds in the minor groove of the target DNA. The incorporation of thiazoles into the minor groove binding region of a polyamide compound had previously been used to selectively target the adenine over the

thymine of an AT base pair, due to alterations in the alignment of the N-terminal amide –NH in the minor groove.[55] This sequence preference may also have affected the binding of our compounds to an AT-rich DNA target by our compounds not aligning with the minor groove in the most advantageous manner.

In addition, the phenyl ring attached to the minor groove binding region of our compounds protruded perpendicularly from the minor groove, as seen in molecular modeling (Fig. 25). The lower binding affinities of compounds **1**-**3** could have been the result of steric interactions between the phenyl ring and the minor groove, shifting the position of the compounds in the minor groove resulting in less favorable alignment with the hydrogen bond acceptors in the minor groove.

Netropsin possessed a guanidine tail at its N-terminus. This tail interacted with the minor groove of DNA by forming hydrogen bonds between the protonated tail and the N3 position of adenines.[134] The replacement of this group with a proton or a bulky chain would prevent the formation of this hydrogen bond, resulting in the poorer binding affinities observed for these compounds toward the minor groove of AT-rich DNA sequences. The formamide group at the N-terminus of compound **3** also lacks this additional hydrogen bond, but maintains an amide –NH at this terminus to form a hydrogen bond with adenine N3 and thymine O2 groups in the minor groove, a feature missing from compound **1**.

Like the N-terminal guanidine, the C-terminal amidine of netropsin also played a role in binding to the minor groove of AT-rich DNA sequences, specifically to the N3 position of adenines.[134] Dimethylaminopropylamide, like the C-terminal amidine tail of netropsin, is protonated under physiological conditions allowing it to interact with the

minor grooves of AT base pairs.[126] The replacement of this amidine tail with a dimethylaminopropylamide tail has been shown to result in decreased binding affinities.[126-127] This decrease, however, was not severe enough to explain the 300 to 2000-fold difference in binding affinities between our compounds and netropsin. The dimethylaminopropyl-amide tail has been utilized in a number of polyamide and lexitropsin derivatives due to the increased hydrophobicity which increases membrane transport and forms favorable hydrophobic controls with the methylenes of the sugar group in the minor groove.[125]

The inclusion of a bulky group into compounds **1** and **3** was done to induce a bend in DNA by widening the minor groove through steric interactions. Using FRET, we had expected to see an increase in the efficiency of energy transfer in the presence of our compounds, indicating the induction of a bend. This would correlate to DNA bending. However, we observed the opposite effect. All three compounds decreased the efficiency of energy transfer between the fluorescent donor and acceptor. These lower efficiencies corresponded to an increase in the end-to-end distance with the drug:DNA complex being 1 to 5 Å̋ longer than the uncomplexed DNA.

The longer end-to-end distance indicates that the binding of compounds made the DNA targets more linear. Consequently, this suggests that the unliganded DNA target was bent. This conclusion is validated based upon the efficiencies of energy transfer of the uncoupled target DNA fragment. These measurements reveal that the end-to-end distance of our target DNA fragments is about 58.9 Å, shorter than a completely linear DNA strand with the same number of base pairs (66.5 Å). We believe that the AT-rich

DNA sequence in the center of the DNA target fragment may have been the source of the bend. Previous researchers have noted that A-rich DNA sequences can bend.

Compounds **1** and **3**, and to a lesser degree compound **2**, lessened the intrinsic DNA bend of the DNA target (Table IV). Complexes of the DNA and compounds **1** and **3** displayed DNA bends of 25° and 29°, respectively, compared to 55° for the uncomplexed DNA. Complexes of compound **2** lessened the intrinsic bend to 48°.

There are three possible mechanisms for how our compounds could have converted the intrinsic bend of the DNA into a more linear conformation (Fig. 42). The first is that these agents bind only to the linear conformation of DNA. This would stabilize the linear DNA:drug complex which in turn would shift equilibrium to the linear complex resulting in a decreased FRET signal. However, for this to occur, our compounds would have had to cause minor groove compression at the binding site (Fig. 42A). The presence of the phenyl ring of the compounds seems incompatible with a compression of the minor groove.

The second possibility is that the intrinsic bend observed in our target DNA is due to induced minor groove compression, where binding of the drug into the minor groove results in widening the minor groove and straightening the DNA (Fig. 42B). This mechanism would be consistent with the observation of minor groove compression in intrinsic bends and would also be compatible with our proposed mechanism of action for our agents. Such a conclusion would have to be validated by additional experiments to determine the mechanism of the intrinsic bend in our target DNA fragments and also a more detailed study of different steric groups in the compounds.

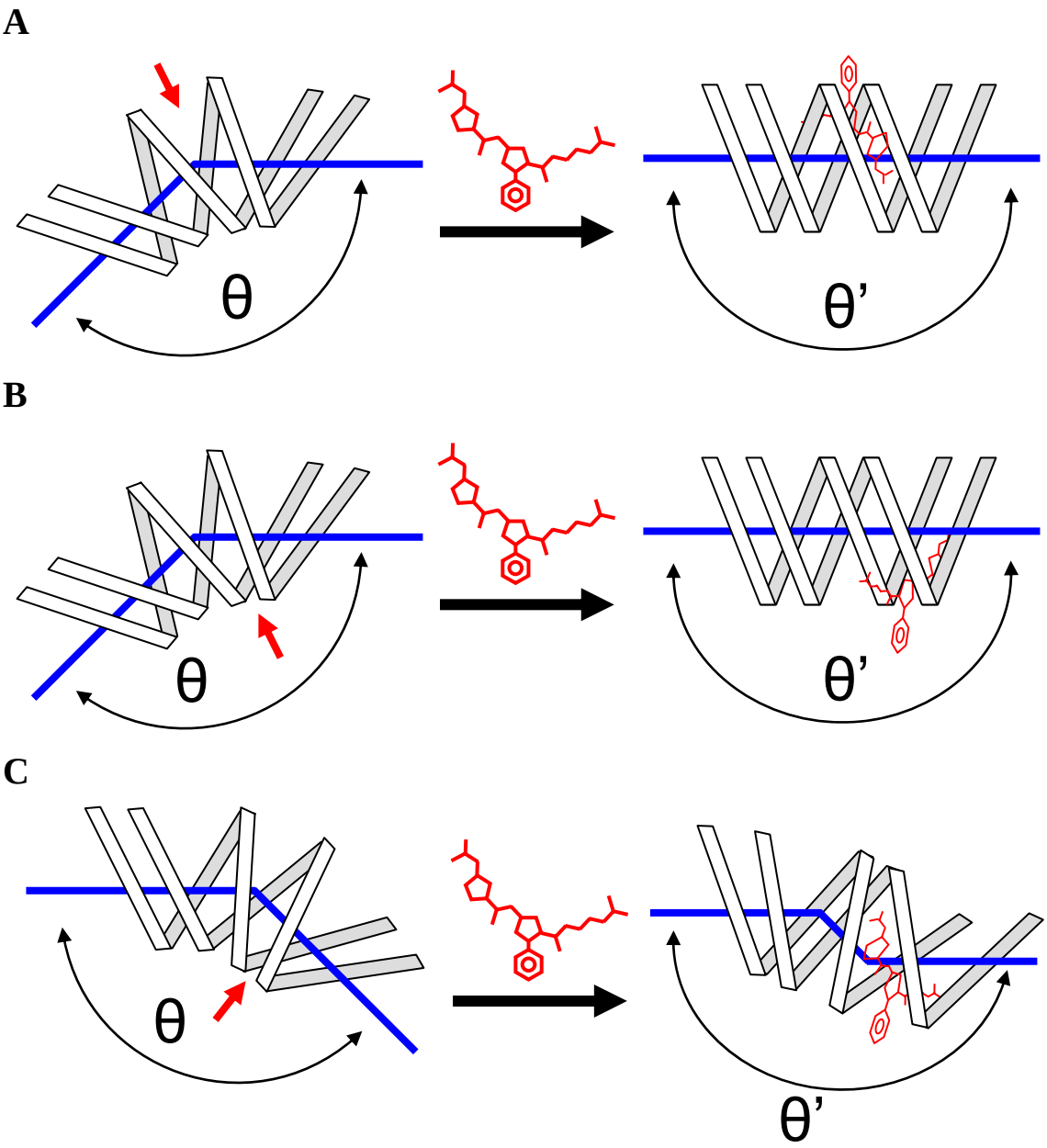


Figure 42. Our compounds could straighten the intrinsically bent 19-bp FRET DNA target by narrowing a widened minor groove (**A**), by widening a narrowed minor groove (**B**) or by creating a second, opposing, bend in the DNA at a different position from the intrinsic bend (**C**). θ represents the bend angle of the DNA target in the absence of compound, the red arrow indicates where the compound binds to the DNA and θ' represents the closer to linear bend angle of the DNA target coupled to compound.

The third possible mechanism is that the intrinsic bends of the DNA target fragments may have occurred at a location other than the binding site of compounds **1**, **2**

and **3** (Fig. 42C). This model would be especially likely in the event of the intrinsic bend occurring at a location other than the AT-rich central sequence. In this situation, the end-to-end distance of the DNA fragment was lengthened by the compound-induced second bend in the opposite direction as the intrinsic bend (Fig 42C). The calculations of the magnitude of the bends induced by our compounds in this model are far more complex and would require numerous other experiments.

The goal of this project was to create small molecules able to selectively bind to a DNA sequence and induce a DNA bend upon binding. Using the polyamide netropsin as a parent compound we synthesized three compounds that combined the sequence-selective binding of netropsin-related compounds and a DNA bending mechanism similar to that of ET-743. Our compounds displayed a preference for AT-rich sequences, though they lacked the binding affinity of netropsin toward this sequence. The conformation of DNA, as measured by the end-to-end distance of the DNA target changed when the compounds were bound to DNA. This suggests that our agents have the ability to alter the conformation of DNA. However, this alteration was more complex than expected. The DNA target possessed an intrinsic bend that was straightened when compounds **1**, **2** and **3** were bound. This may have been accomplished through the DNA bending activity of the compounds, with bends induced counter to the intrinsic bends already present. In this situation our compounds functioned as intended; compounds **1**, **2** and **3** were able to selectively bind to a DNA target and induce a DNA bend upon binding.

V. MATERIALS & METHODS

V.A. Project 1: Tethered-TFOs and Gene Expression

Chemicals and Enzymes. The deoxynucleotides and linker phosphoramidites were purchased from Glen Research. The unbroken TFOs, STFO and LTFO (Fig. 45), were purchased from TriLink Biotechnologies (San Diego, CA). The oligonucleotides were purified by extraction from a polyacrylamide gel. Restriction endonucleases, T4 polynucleotide kinase, Quick T4 DNA ligase, *Taq* DNA polymerase and pBR322 were purchased from New England BioLabs (Ipswich, MA). The EcoPro T7 kit was purchased from Novagen (San Diego, CA). Recombinant firefly luciferase was purchased from Calbiochem (San Diego, CA). The β-galactosidase was from Sigma-Aldrich. The anti-luciferase-HRP conjugated antibodies and the anti-β-galactodidase-HRP-labeled antibodies were from Research Diagnostics, Inc. (Concord, MA).

Oligonucleotide Syntheses. The tethered-TFOs, STFO and LTFO, the broken TFOs, bSTFO and bLTFO (Fig. 43), the PCR primers (Fig. 43) and the oligonucleotides used to construct the plasmid inserts (Fig. 44 and 45) were synthesized on a PerSeptive Biosystems Expedite Nucleic Acid Synthesis System using standard conditions. The oligonucleotides were purified by 13% denaturing PAGE followed by solid phase extraction.

A

STFO: 5′-GGTGGGGTGGGTGGG-**PgPgPg**-GGGTGGGTGGGGTGG-3′
LTFO: 5′-GGTGGGGTGGGTGGG-**TgTgTgTg**-GGGTGGGTGGGGTGG-3′
bSTFO:5′-GGTGGGGTGGGTGGG-3′ and 5′-**PgPgPg**-GGGTGGGTGGGGTGG-3′
bLTFO:5′-GGTGGGGTGGGTGGG-3′ and 5′-**TgTgTgTg**-GGGTGGGTGGGGTGG-3′

B

C1:5′-CTAGCTAGCTAGTAGGAGGG-3′
C2:5′-CGCGGATCCAGATCTGCTCG-3′
E1:5′-CGTGCTGCTAGCTAGT-3′
E2:5′-TGTAGGAGCTATAGGC-3′

Figure 43. (**A**) The TFOs used in the assays. **Pg** represents a propylene glycol phosphodiester unit, while **Tg** represents a triethylene glycol phosphodiester unit. (**B**) Sequences of PCR primers C1, C2, E1 and E2.

Plasmid Insert Construction. The plasmid inserts (Fig. 44) were synthesized in four parts (Fig. 45) as described below. Each oligonucleotide (100 pmol) was treated with T4 polynucleotide kinase (10 units) at 37°C for 1h and the four oligonucleotides for each insert were combined, annealed by a heating-cooling cycle of 95°C for 5 min followed by slow cooling to 25°C at a rate of 1°C/min. The plasmid inserts were ligated into whole units by treatment with Quick T4 DNA ligase (10 units) at room temperature for 5 min.

Plasmid Construction. The plasmids were digested with restriction enzymes corresponding to the plasmid insert to be incorporated. The cut plasmids were purified on an agarose gel. The cut plasmid (50 ng) and the plasmid insert (5x the molarity of the cut plasmid) were ligated together with Quick T4 DNA ligase (10 units) at room temperature for 5 minutes. Ligated plasmids were then amplified by transformation of DH5α *E. coli* cells cultured on ampicillin containing agar plates. The colonies were then grown overnight via inoculation into ampicillin containing LB media. Minipreps of the inoculations were performed to acquire the amplified plasmid DNA. pBend was

constructed by inserting the TFO target DNA fragment (Fig. 44 and 45) into the *Nhe*I/*Bam*HI site of pBR322. Insertion was confirmed by *Xho*I digestive analysis. pBLP77 was constructed by first inserting the promoter DNA fragment (Fig. 44 and 45) into the *Bgl*II/*Sal*I site of pBend to create pBP77. pBP77 was verified by *Xba*I digestive analysis. The pBLP77 was constructed by inserting the luciferase gene of pGL3 (Promega) into the *Nco*I/*Xba*I site of pBP77. Insertion was verified by digestive analysis. pBLP83 was constructed by first inserting the promoter DNA fragment (Fig. 44 and 45) into the *Bam*HI/*Sal*I site of pBend to create pBP83. pBP83 was verified by *Xba*I digestive analysis. The pBLP83 was constructed by inserting the luciferase gene of pGL3 into the *Nco*I/*Xba*I site of pBP83. Insertion was verified by digestive analysis. pBLP80 was constructed by inserting the Phase 80 DNA fragment (Fig. 44 and 45), first cut with *Xho*I and *Nco*I, into the *Xho*I/*Nco*I site of pBLP83. Insertion was verified by *Kpn*I and *BamH*I digestive analysis. pBLP86 was constructed by inserting the Phase 86 DNA fragment (Fig. 44 and 45) , first cut with *Xho*I and *Nco*I, into the *Xho*I/*Nco*I site of pBLP83. Insertion was verified by *Kpn*I and *BamH*I digestive analysis.

A

```
   Nhe I                                                                Xho I  BglII BamH I
5′-CTAGCTAGTAGGAGGGGAGGGAGGGCGCCGGCGCGGGAGGGAGGGGAGGTGCTCGAGCAGATCTG     -3′
3′-    GATCATCCTCCCCTCCCTCCCGCGGCCGCGCCCTCCCTCCCCTCCACGAGCTCGTCTAGACCTAG-5′
```

B

```
Bgl II-BamH I                                        Nco I Xba I Sal I
5′-GATCCTATACGCTTGCATTAGAAAGGTTTCTGGCCGACCTTATAACCATGGTCTAGAG     -3′
3′-    GATATGCGAACGTAATCTTTCCAAAGACCGGCTGGAATATTGGTACCAGATCTCAGCT-5′
```

C

```
         Xho I    Kpn I                                                  Nco I
5′-CCGTCCTCGAGCATGGTACCCTATACGCTTGCATTAGAAAGGTTTCTGGCCGACCTTATAACCATGGCATG-3′
3′-GGCAGGAGCTCGTACCATGGGATATGCGAACGTAATCTTTCCAAAGACCGGCTGGAATATTGGTACCGTAC-5′
```

D

```
         Xho I    Kpn I                                                  Nco I
5′-CCGTCCTCGAGCATCGGTACCATCGACTATACGCTTGCATTAGAAAGGTTTCTGGCCGACCTTATAACCATGGCATG-3′
3′-GGCAGGAGCTCGTAGCCATGGTAGCTGATATGCGAACGTAATCTTTCCAAAGACCGGCTGGAATATTGGTACCGTAC-5′
```

Figure 44. (**A**) The TFO target fragment sequence. The *Nhe* I overhang is highlighted in yellow; the tethered-TFO target sites are in red text with the center of the bending region highlighted red; the *Xho* I site is highlighted in green; the *Bgl* II site is highlighted in pink; and the *BamH* I overhang is highlighted in blue. (**B**) The promoter fragment sequence. The *Bgl* II/*BamH* I compatible overhang is highlighted in blue; the *malT* promoter region is underlined with the -10 region highlighted in gray and the -35 region highlighted in yellow; the *Nco* I site is highlighted in red; the *Xba* I site is white text highlighted in dark green; and the *Sal* I overhang is highlighted in black with white text. (**C**) The Phase 80 fragment sequence. The *Xho* I site is highlighted in green; the *Kpn* I site is white text highlighted with purple; the *malT* promoter region is underlined with the -10 region highlighted in gray and the -35 region highlighted in yellow; and the *Nco* I site is highlighted in red. (**D**) The Phase 86 fragment sequence. The *Xho* I site is highlighted in green; the *Kpn* I site is white text highlighted with purple; the *malT* promoter region is underlined with the -10 region highlighted in gray and the -35 region highlighted in yellow; and the *Nco* I site is highlighted in red.

TFO Target DNA Fragment:

```
5′-CTAGCTAGTAGGAGGGGAGGGAGGGCGCCGGCGCGGGAGGGAGGGGAGGTGCTCGAGCAGATCTG    -3′
3′-    GATCATCCTCCCCTCCCTCCCGCGGCCGCGCCCTCCCTCCCCTCCACGAGCTCGTCTAGACCTAG-5′
```

Single-stranded oligonucleotides to create TFO-target fragment:

```
TFOt1: 5′-CTAGCTAGTAGGAGGGGAGGGAGGGCGCCGGCCCGGGAGG-3′
TFOt2: 5′-GAGGGGAGGTGCTCGAGCAGATCTG-3′
TFOt3: 5′-GATCCAGATCTGCTCGAGCACCTCCCCTCCCTCCCGCGCC-3′
TFOt4: 5′-GGCGCCCTCCCTCCCCTCCTACTAG-3′
```

Promoter DNA Fragment:

```
5′-GATCCTATACGCTTGCATTAGAAAGGTTTCTGGCCGACCTTATAACCATGGTCTAGAG    -3′
3′-    GATATGCGAACGTAATCTTTCCAAAGACCGGCTGGAATATTGGTACCAGATCTCAGCT-5′
```

Single-stranded oligonucleotides to create promoter fragment:

```
TT1: 5′-GATCCTATACGCTTGCATTAGAAAGGTTTCTGGCCGACC-3′
TT2: 5′-TTATAACCATGGTCTAGAG-3'
TT3: 5′-TCGACTCTAGACCATGGTTATAAGGTCGGCCAGAAACC-3′
TT4: 5′-TTTCTAATGCAAGCGTATAG-3′
```

Phase 80 DNA Fragment:

```
5′-CCGTCCTCGAGCATGGTACCCTATACGCTTGCATTAGAAAGGTTTCTGGCCGACCTTATAACCATGGCATG-3′
3′-GGCAGGAGCTCGTACCATGGGATATGCGAACGTAATCTTTCCAAAGACCGGCTGGAATATTGGTACCGTAC-5′
```

Single-stranded oligonucleotides to create Phase 80 fragment:

```
P84a: 5′-CCGTCCTCGAGCATGGTACCCTATACGCTTGCATTAGAAAG-3′
P84b: 5′-GTTTCTGGCCGACCTTATAACCATGGCATG-3′
P84c: 5′-CATGCCATGGTTATAAGGTCGGCCAGAAACCTTTCTAATGCAAGC-3′
P84d: 5′-GTATAGGGTACCATGCTCGAGCACGG-3′
```

Phase 86 DNA Fragment:

```
5′-CCGTCCTCGAGCATCGGTACCATCGACTATACGCTTGCATTAGAAAGGTTTCTGGCCGACCTTATAACCATGGCATG-3′
3′-GGCAGGAGCTCGTAGCCATGGTAGCTGATATGCGAACGTAATCTTTCCAAAGACCGGCTGGAATATTGGTACCGTAC-5′
```

Single-stranded oligonucleotides to create Phase 86 fragment:

```
P90a: 5′-CCGTCCTCGAGCATCGGTACCATCGACTATACGCTTGCATTAGAAAG-3′
P90b: 5′-GTTTCTGGCCGACCTTATAACCATGGCATG-3′
P90c: 5′-CATGCCATGGTTATAAGGTCGGCCAGAAACCTTTCTAATGCAAGC-3′
P90d: 5′-GTATAGTCGATGGTACCGATGCTCGAGCACGG-3′
```

Figure 45. pBLP plasmid inserts.

Gel Mobility Assay. Two 129 bp DNA fragments, containing the TFO-target sequence, were amplified by PCR from pBend using primers C1 and C2 to generate the centered TFO target fragment or E1 and E2 to generate the end-located TFO target fragment. The fragment was purified by 13% denaturing PAGE followed by extraction from the gel. The purified centered and end-located TFO target fragments (2 pmol) were incubated with TFOs (2pmol) in 10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$ and 10% sucrose at 37°C for 16h. The TFO-target complexes were analyzed by 8% nondenaturing PAGE (29:1 cross-linking) containing 10 mM $MgCl_2$. The gel was run in 89 mM Tris borate and 10 mM $MgCl_2$ buffer (TBM) for 4 h at 10 V/cm and stained with SYBR GOLD (Molecular Probes) for 45 min. The gel was imaged on a Kodak Digital Science Image Station.

Gene Expression Analysis. The plasmids pBLP77, pBLP80, pBLP83, and pBLP86 were linearized with *Nde*I. The linearized plasmids (2.7 nM) were incubated with excess TFOs (0.27 μM) in 10 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$ and 10% sucrose at 37°C for 16 h. Transcription and translation of the plasmid-TFO complexes was performed with the EcoPro T7 System (Novagen) and incubation for 1 h at 37°C. A sample of the expressed luciferase protein (10 μl) was removed and β-galactosidase (200 ng) was added to each sample as a standard. SDS-PAGE loading buffer was added to each sample and the samples were then denatured in a thermocycler by incubation at 90°C for 10 min. If the sample was not blue after the addition of the SDS-PAGE loading buffer, 1 μL increments of 1 M Tris (pH 9.5) were added until a blue color was attained. The samples were run by SDS-PAGE on 10% Tris-HCl Ready Gels (Bio-Rad Laboratories) in

a 25 mM Tris, 250 mM glycine and 0.1% SDS buffer at 33V/cm for 40 min. The proteins were transferred from the gel to PVDF membranes in 25 mM Tris, 190 mM glycine, 20% methanol and 0.1% SDS at 100 V for 1h. The membranes were blocked in PBS buffer (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4 \bullet 7H_2O$, 1.4 mM KH_2PO_4) containing 10% nonfat milk at 4°C with rocking for 16 h and then washed in a wash buffer (100 mM NaCl, 10 mM Tris, 1 mM EDTA and 0.1% Tween-20). The PVDF membranes were blotted with rabbit-derived anti-β-galactosidase-HRP conjugated antibodies (1:5000 from a 10 mg/mL stock) and goat-derived anti-luciferas-HRP conjugated antibodies (1:2000 for a 10 mg/mL stock) in a blotting buffer (100 mM NaCl, 10 mM Tris, 1 mM EDTA, 0.1% Tween-20 and 1% nonfat milk) at 4°C with rocking for 2 h. The membranes were then washed with the wash buffer and treated with an ECL Western blotting kit (Amersham Biosciences) for 1 min. The membranes were scanned on a Kodak Digital Science Image Station with a 30 min exposure without UV light or a filter. The image was analyzed with Kodak 1D Image Analysis software.

V.B. Project 2: A Sequence Selective Push Bender

Synthesis 2-[(1-Methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (3-dimethylamino-propyl)-amide (**1**).

2-[(1-Methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid ethyl ester (4). *N*-methylpyrrole-2-carboxylic acid (0.5 g; 4 mmol) and ethyl-2-amino-4-phenyl-5-thiazole carboxylate (1 g; 4 mmol) were dissolved, by stirring, into DMF (20

mL). EDCI (3 g; 16 mmol) and DMAP (2.44 g; 20 mmol) were added and the reaction mixture was stirred overnight at room temperature. Ethyl acetate (200 mL) was added to the reaction mixture and the reaction mixture was then washed with 10% aqueous hydrochloric acid (3 x 200 mL) and saturated aqueous sodium bicarbonate (3 x 200 mL). The organic layer was collected, dried with magnesium sulfate and was rotavapped to dryness. The resulting product was dried overnight under reduced pressure in the presence of P_2O_5. Yield 61.6% (0.875 g; 2.46 mmol). ^{1}H NMR 300 MHz (DMSO-d_6) 1.19-1.24 (3H, t, J = 6.6 Hz), 3.93 (3H, s), 4.15-4.22 (2H, q, J = 7.1 Hz), 6.15 (1H, s), 7.05 (1H, s), 7.18 (1H, s), 7.43 (2H, s), 7.71 (2H, s),12.62 (1H, s).

2-[(1-Methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (5). 2-[(1-Methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid ethyl ester, **4** (0.3123 g; 0.879 mmol) was dissolved in a solution containing 1N lithium hydroxide (21.66 mL) and methanol (29.06 mL). The reaction mixture was heated to 60°C for 2.5 hours. The reaction mixture was evaporated to half the original volume and ethyl acetate (75 mL) was added. The aqueous layer was removed and acidified with concentrated hydrochloric acid until the pH reached 3. The acidic solution was then extracted with ethyl acetate (4 x 75 mL). The organic layers were collected, dried with magnesium sulfate and filtered. The filtrate was rotavapped to dryness. The resulting product was dried overnight under reduced pressure in the presence of P_2O_5. Yield 74.8% (0.2034 g; 0.657 mmol). ^{1}H NMR 300 MHz (DMSO-d_6) 3.93 (3H, s), 6.14 (1H, s), 6.71 (1H, s), 6.94 (1H, s), 7.17 (1H, s), 7.36-7.42 (2H, t, J = 8.79 Hz), 7.72-7.73 (2H, d, J = 3.93 Hz), 11.70 (1H, s), 12.53 (1H, s).

2-[(1-Methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (3-dimethylamino-propyl)-amide (1). 2-[(1-Methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid, **5** (0.03 g; 0.916 mmol) was dissolved in DMF (4 mL). HBTU (0.086 g; 0.366 mmol) and NMM (0.037 g; 0.366 mmol) were added to the reaction and the reaction mixture was stirred at room temperaturefor 2 hrs. 3-(dimethylamino) propylamine (0.028 g; 0.275 mmol) was then added and the reaction mixture was stirred overnight at room temperature. Ethyl acetate (15 mL) was added to the reaction mixture and the reaction mixture was washed with water (3 x 15 mL). The organic layer was collected, dried with magnesium sulfate and filtered. The filtrate was rotavapped to dryness and purified using silica gel column chromatography using 5:1:0.2 ethyl acetate:methanol:triethylamine. Yield 51.3% (0.019 g; 0.047 mmol). ^{1}H NMR 300 MHz (DMSO-d_6) 1.65-1.72 (2H, quintet, J = 5.13 Hz), 2.42 (6H, s), 2.49 (1H, s), 2.58-2.62 (2H, t, J = 6.86 Hz), 3.19-3.21 (2H, t, J = 5.31 Hz), 3.82 (3H, s), 6.14 (1H, s), 7.15 (1H, s), 7.41-7.43 (2H, t, J = 6.81 Hz), 7.70-7.72 (2H, d, J = 7.06 Hz), 8.20 (1H, s). Anal. Calc for $C_{21}H_{25}N_5O_2S$.DMF: C 59.61; H 6.46; O 9.93 Found: C 59.11; H 6.02; O 9.33.

Synthesis {5-[5-(3-Dimethylamino-propylcarbamoyl)-4-phenyl-thiazol-2-ylcarbamoyl]-1-methyl-1H-pyrrol-3-yl}-carbamic acid tert-butyl ester (**2**).

2-[(4-tert-Butoxycarbonylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid ethyl ester (6). 4-tert-Butoxycarbonylamino-1-methyl-1H-pyrrole-2-carboxylic acid (0.96 g; 4 mmol) from Sigma-Aldrich (St. Louis, MO) and ethyl-2-amino-4-phenyl-5-thiazole carboxylate (1 g; 4 mmol) were dissolved, by stirring,

into DMF (20 mL). EDCI (3 g; 16 mmol) and DMAP (2.44 g; 20 mmol) were added to the mixture and the reaction was stirred overnight at room temperature. Ethyl acetate (200 mL) was then added and the organc layer was then washed with 10% aqueous hydrochloric acid (3 x 200 mL) and saturated aqueous sodium bicarbonate (3 x 200 mL). The organic layer was collected, dried with magnesium sulfate and filtered. The filtrate was rotavapped to dryness. Yield 51.8% (0.982 g; 2.09 mmol). ^{1}H NMR 300 MHz (DMSO-d_6) 1.19-1.24 (3H, t, J = 7.07 Hz), 1.45 (9H, s), 3.87 (3H, s), 4.15-4.22 (2H, q, J = 6.72 Hz), 7.17 (1H, s), 7.25 (1H, s), 7.31 (1H, s), 7.40-7.45 (2H, t, J = 7.53 Hz), 7.71-7.73 (2H, d, J = 7.08 Hz), 12.67 (1H, s).

2-[(4-tert-Butoxycarbonylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (7). 2-[(4-tert-Butoxycarbonylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid ethyl ester, **6** (0.9824 g; 2.09 mmol) was dissolved into a solution containing 1N lithium hydroxide (51.47 mL) and methanol (69.04 mL). The reaction mixture was heated to 60°C for 3 hours, cooled and evaporated to half the original volume. The resulting solution was extracted with ethyl acetate (175 mL) and the aqueous layer of the extraction was acidified with concentrated hydrochloric acid to pH 3. The acidic aqueous layer was then extracted with ethyl acetate (4 x 175 mL), the organic layers were collected, dried with magnesium sulfate and filtered. The filtrate was rotavapped to dryness. Yield 71.9% (0.665 g; 1.51 mmol). ^{1}H NMR 300 MHz (DMSO-d_6) 1.45 (9H, s), 3.87 (3H, s), 7.04 (1H, s), 7.16 (1H, s), 7.35 (2H, s), 7.71-7.73 (2H, d, J = 4.14 Hz), 9.05 (1H, s), 9.21 (1H, s), 12.57 (1H, s).

{5-[5-(3-Dimethylamino-propylcarbamoyl)-4-phenyl-thiazol-2-ylcarbamoyl]-1-methyl-1H-pyrrol-3-yl}-carbamic acid tert-butyl ester (2). 2-[(4-tert-butoxycarbonylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid, **7** (0.03 g; 0.068 mmol) was dissolved in DMF (4 mL). HBTU (0.064 g; 0.272 mmol) and NMM (0.028 g; 0.272 mmol) were added to the reaction. The reaction mixture stirred at room temperature for 2 hrs followed by the addition of 3-(dimethylamino) propylamine (0.021 g; 0.203 mmol). The reaction mixture was stirred overnight at room temperature followed by addition of ethyl acetate (15 mL) to the reaction mixture. The organic layer was washed with water (3 x 15 mL), collected, dried with magnesium sulfate and filtered. The filtrate was rotavapped to dryness and the resulting product was purified by silica gel column chromatography using 5:1:0.2 ethyl acetate:methanol:triethylamine. Yield 60.3% (0.022 g; 0.041 mmol). ^{1}H NMR 300 MHz (DMSO-d_6) 1.46 (9H, s), 1.53-1.57 (2H, t, J = 6.21 Hz), 2.08 (6H, s), 2.13-2.19 (2H, quintet, J = 4.65 Hz), 3.17-3.19 (2H, t, J = 5.62 Hz), 3.87 (3H, s), 7.14 (1H, s), 7.20 (1H, s), 7.40-7.42 (2H, d, J = 7.22 Hz), 7.69-7.72 (2H, d, J = 6.69 Hz), 8.09 (1H, s), 9.18 (1H, s). Anal. Calc for $C_{26}H_{34}N_6O_4S$.DMF: C 59.30; H 6.51; O 12.15 Found: C 58.99; H 6.33; O 12.47.

Synthesis of 2-[(4-Formylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (3-dimethylamino-propyl)-amide (**3**).

2-[(4-Formylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-4-phenyl-thiazole-5-carboxylic acid (3-dimethylamino-propyl)-amide (3). {5-[5-(3-Dimethylamino-propylcarbamoyl)-4-phenyl-thiazol-2-ylcarbamoyl]-1-methyl-1H-pyrrol-3-yl}-carbamic

acid tert-butyl ester, **2** (0.070 g; 0.133 mmol) was dissolved in dichloromethane (5 mL). Trifluoroacetic acid (0.3 mL) and benzenethiol (0.1 mL) were added and the reaction mixture was stirred at room temperature for 1 hour. The reaction was rotavapped to dryness and the residue was washed with diethyl ether (3 x 15 mL) to remove non-polar impurities. The residue was dissolved in ethanol (10 mL) and ethyl formate (20 mL) was added. The reaction was heated to reflux for 48 hours, cooled and evaporated to dryness. The resulting residue was then purified by silica gel column chromatography using 5:1:0.2 ethyl acetate:methanol:triethylamine. Yield 73.4% (0.044 g; 0.098 mmol). ^{1}H NMR 300 MHz (DMSO-d_6) 1.72-1.81 (2H, quintet, J = 5.43 Hz), 2.64 (6H, s), 2.84-2.90 (2H, t, J = 8.28 Hz), 3.07-3.09 (2H, t, J = 6.87 Hz), 3.90 (3H, s), 7.33 (1H, s), 7.41-7.45 (2H, d, J = 9.21 Hz), 7.70-7.72 (2H, d, J = 6.33 Hz), 8.14 (1H, s), 8.24 (1H, s), 10.17 (1H, s). Anal. Calc for $C_{22}H_{26}N_6O_3S.CH_3CH_2OH$: C 57.58; H 6.44; O 12.78 Found: C 58.01; H 6.12; O 12.50.

Ethidium bromide displacement assay. This assay was performed on a Perkins Elmer L555 Fluorimeter equipped with FL WinLab software. The assay was conducted as follows. Ethidium bromide (4.4 μM final concentration) was added to a quartz cuvette containing Tris buffer (0.1 M Tris, 0.1 M NaCl, pH 8). The fluorescence was measured (ex. 545 nm, em. 595 nm). The DNA target of interest was added (8.8 μM in DNA base pairs final concentration) and the fluorescence was again measured. Compounds **1**, **2** and **3** were then titrated into the cuvette, measuring the fluorescence 5 min after every addition. Each compound was titrated until the relative fluorescence had decreased to less than 50%.

FRET analysis. Two 19-bp DNA targets were diluted to 0.5 μM in 10 mM sodium phosphate buffer (pH 7). One duplex contained a 5'-terminal fluorescein, while the other duplex contained both a 5'-terminal fluorescein and a terminal tetramethylrhodamine on the other strand. The emission spectrum for each of these duplexes was analyzed in both the absence and presence of compounds **1**, **2** and **3**. The samples were analyzed on a PerkinElmer Instruments LS55 Luminescence Spectrometer with corresponding FL WinLab software. The excitation and emission slits of the machine were set to 3 nm and 10 nm, respectively. The samples were excited at 495 nm and the emissions were scanned from 500 nm to 700 nm. Data analysis utilized the integrated area under of the various spectrums from 510 nm to 530 nm as calculated with the FL WinLab software. The emission spectrums for 1 ml of each DNA target, at 0.5 μM in a quartz cuvette, were first scanned without any of the synthesized compounds present. The samples were then scanned after the addition of compounds **1**, **2** and **3**, in excess at 2.0 μM, and a 15 minute incubation period at room temperature.

VI. ^{1}H-NMR Spectra

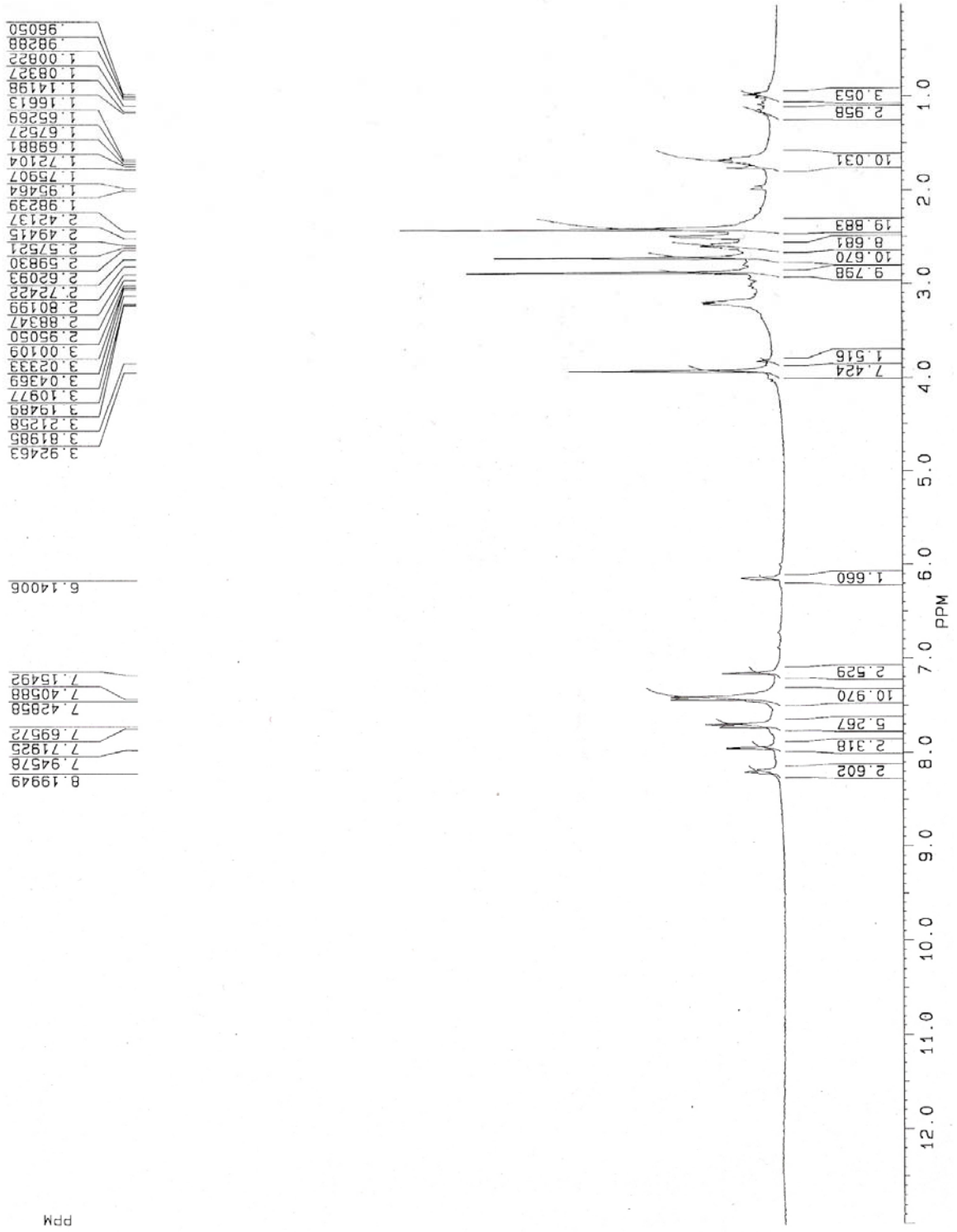


Figure 46. 1H-NMR Spectrum of compound **1** in DMSO-d_6.

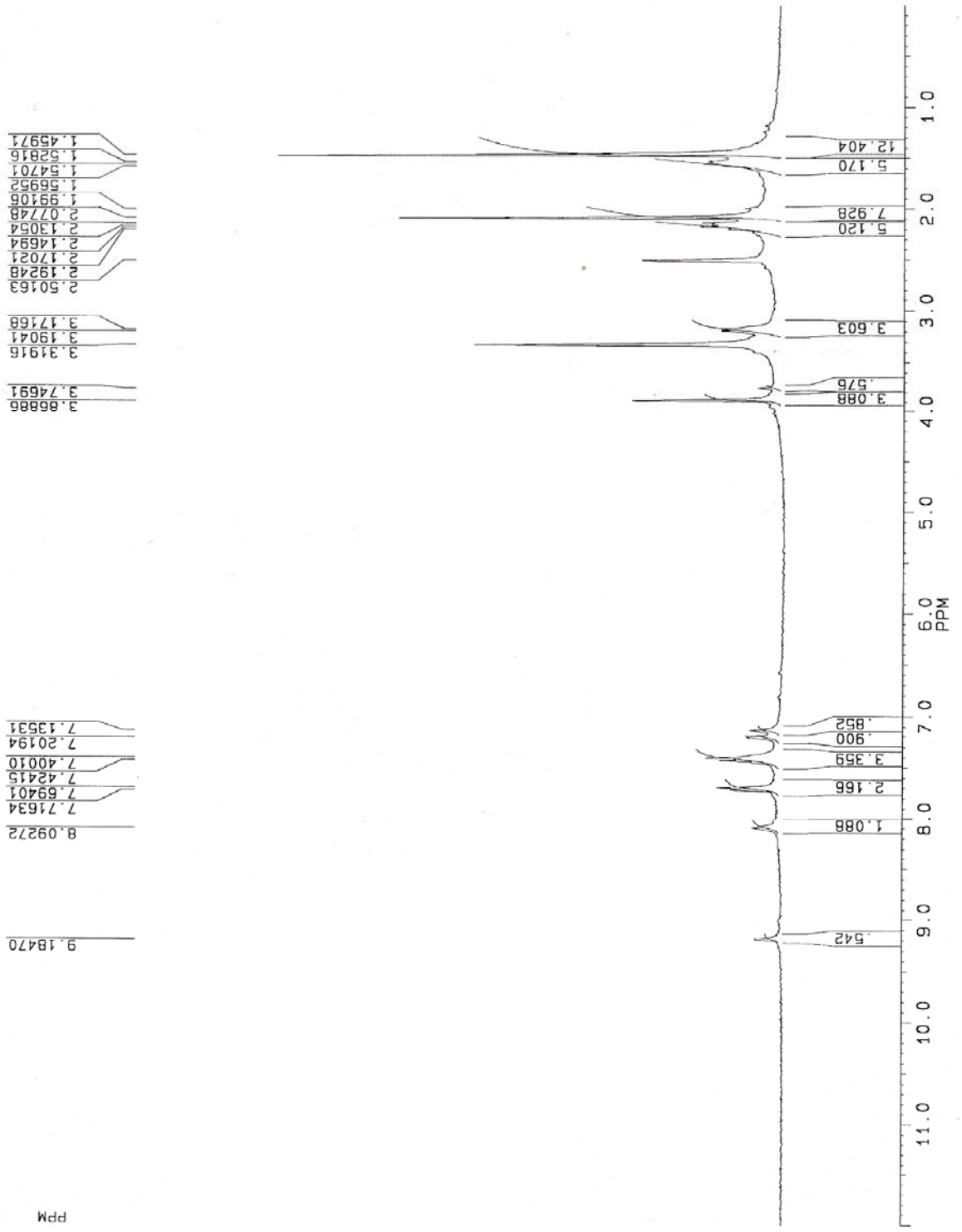


Figure 47. 1H-NMR Spectrum of compound **2** in DMSO-d_6.

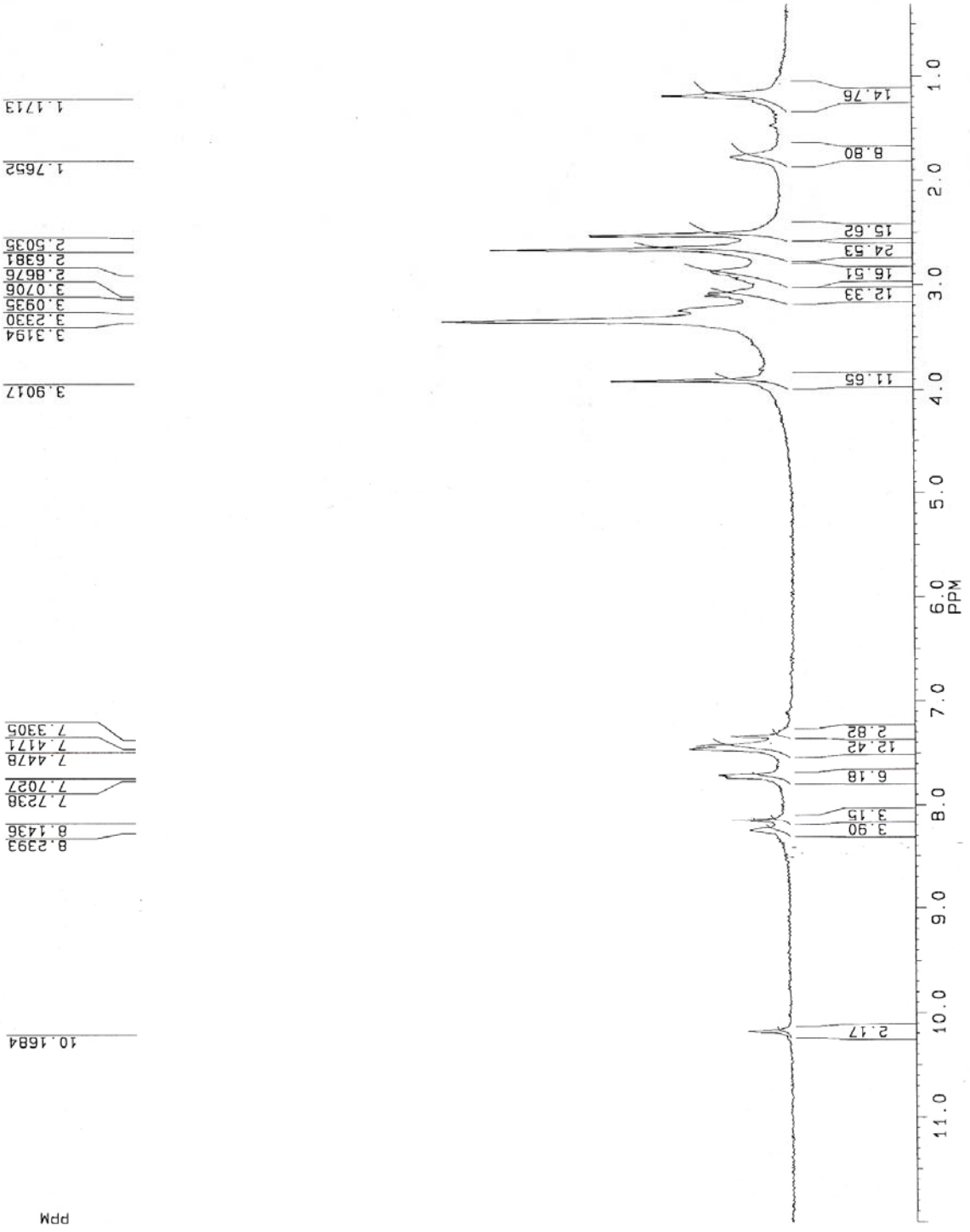


Figure 48. 1H-NMR Spectrum of compound **3** in DMSO-d_6.

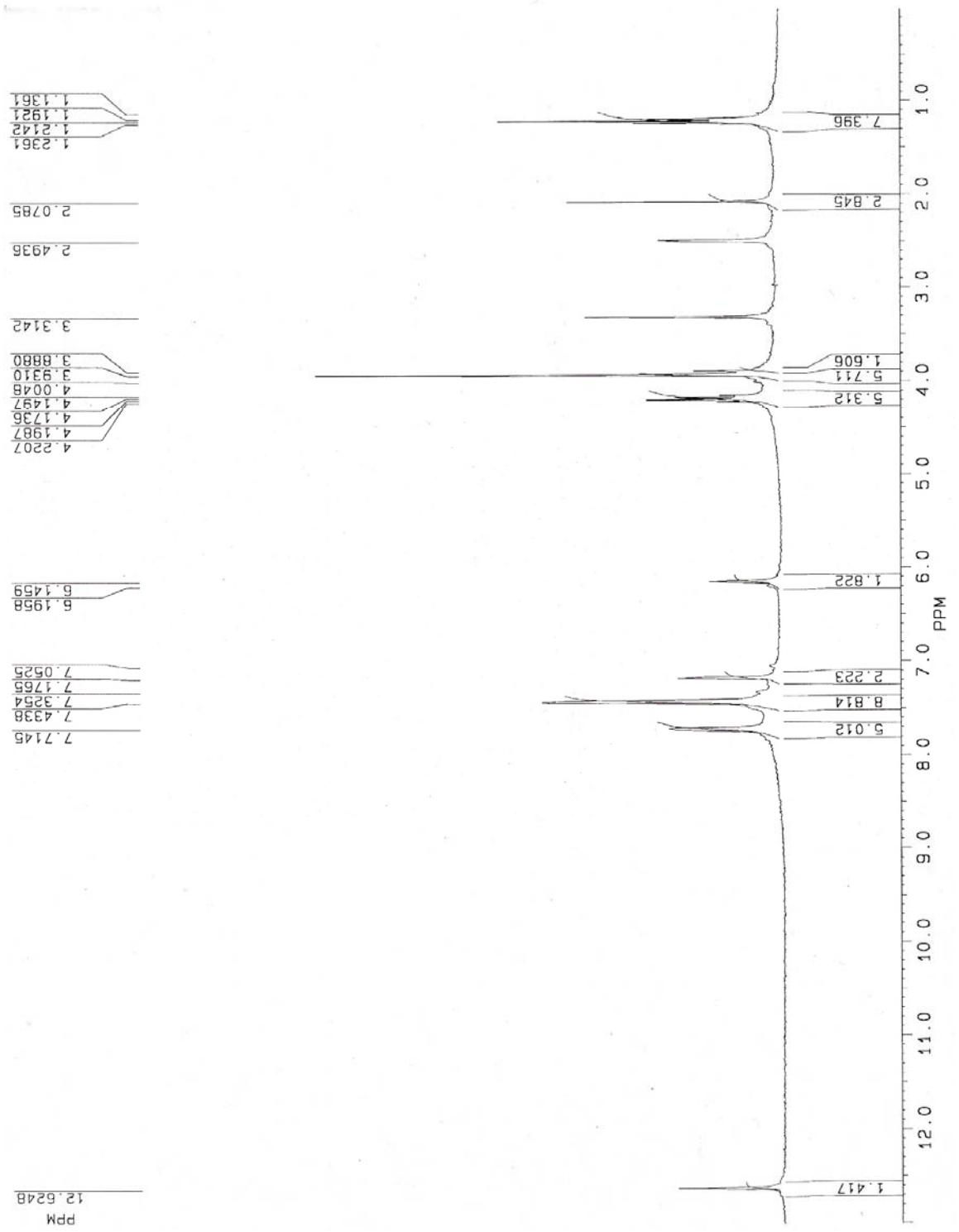


Figure 49. 1H-NMR Spectrum of compound **4** in DMSO-d_6.

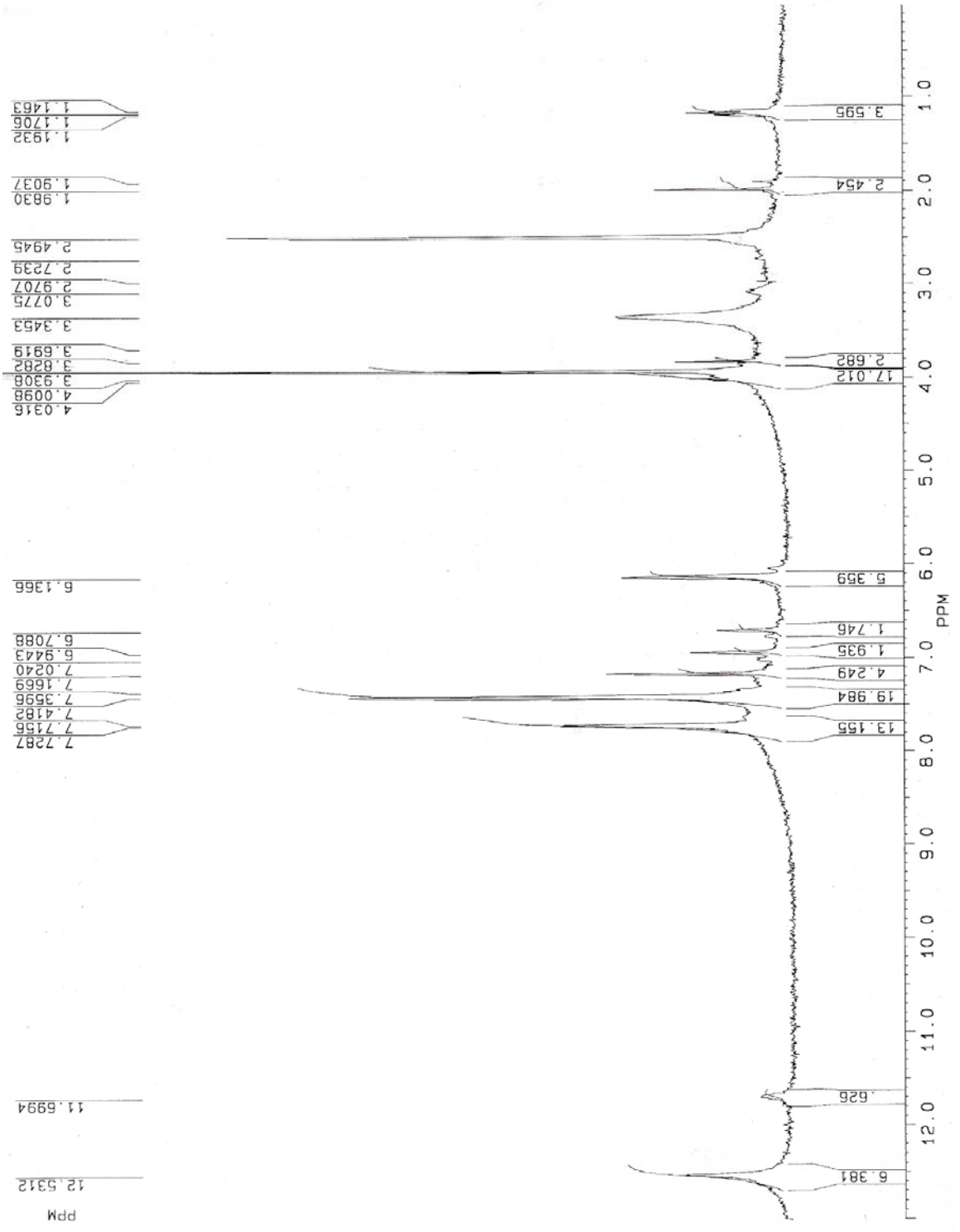


Figure 50. 1H-NMR Spectrum of compound **5** in DMSO-d_6.

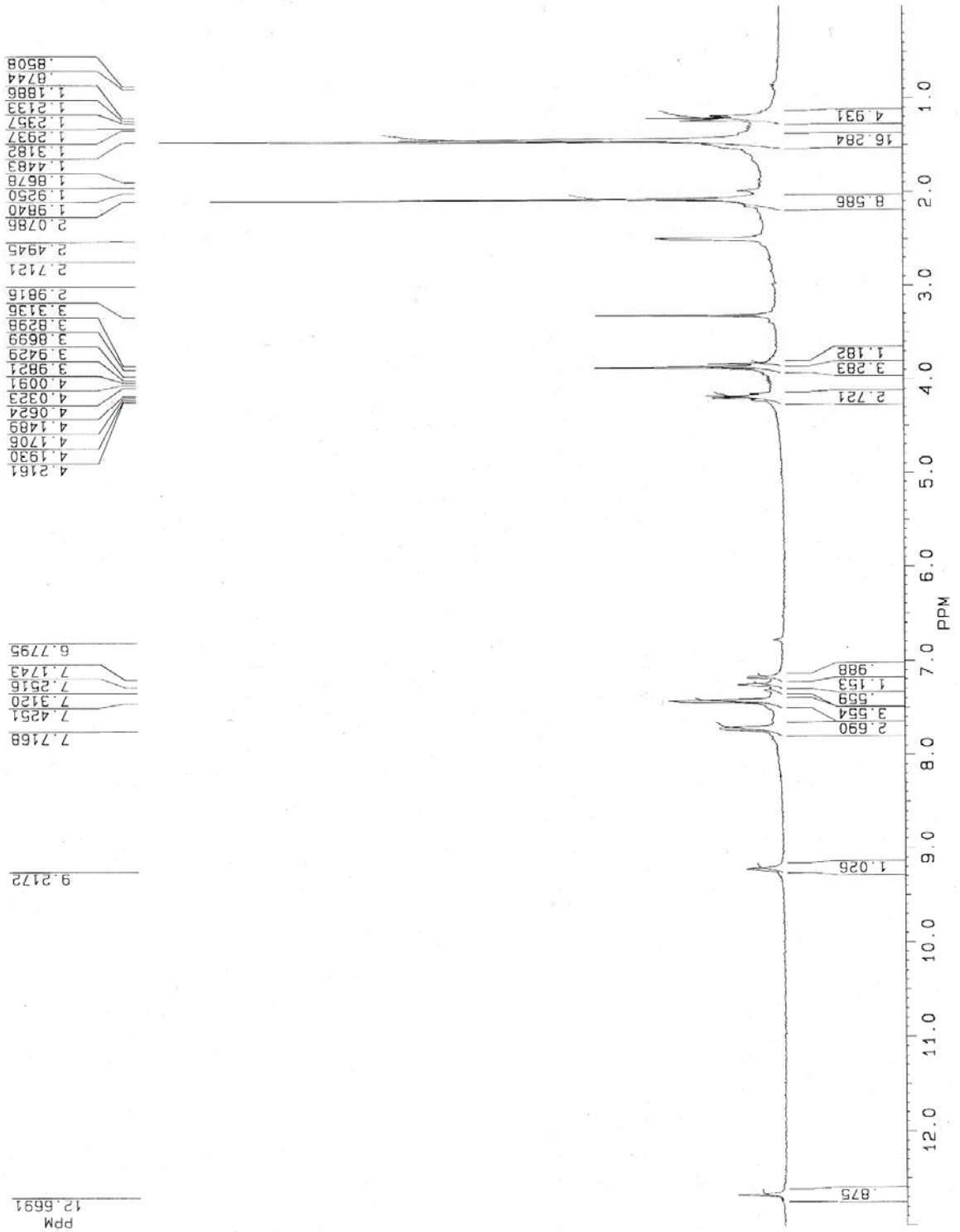


Figure 51. 1H-NMR Spectrum of compound **6** in DMSO-d_6.

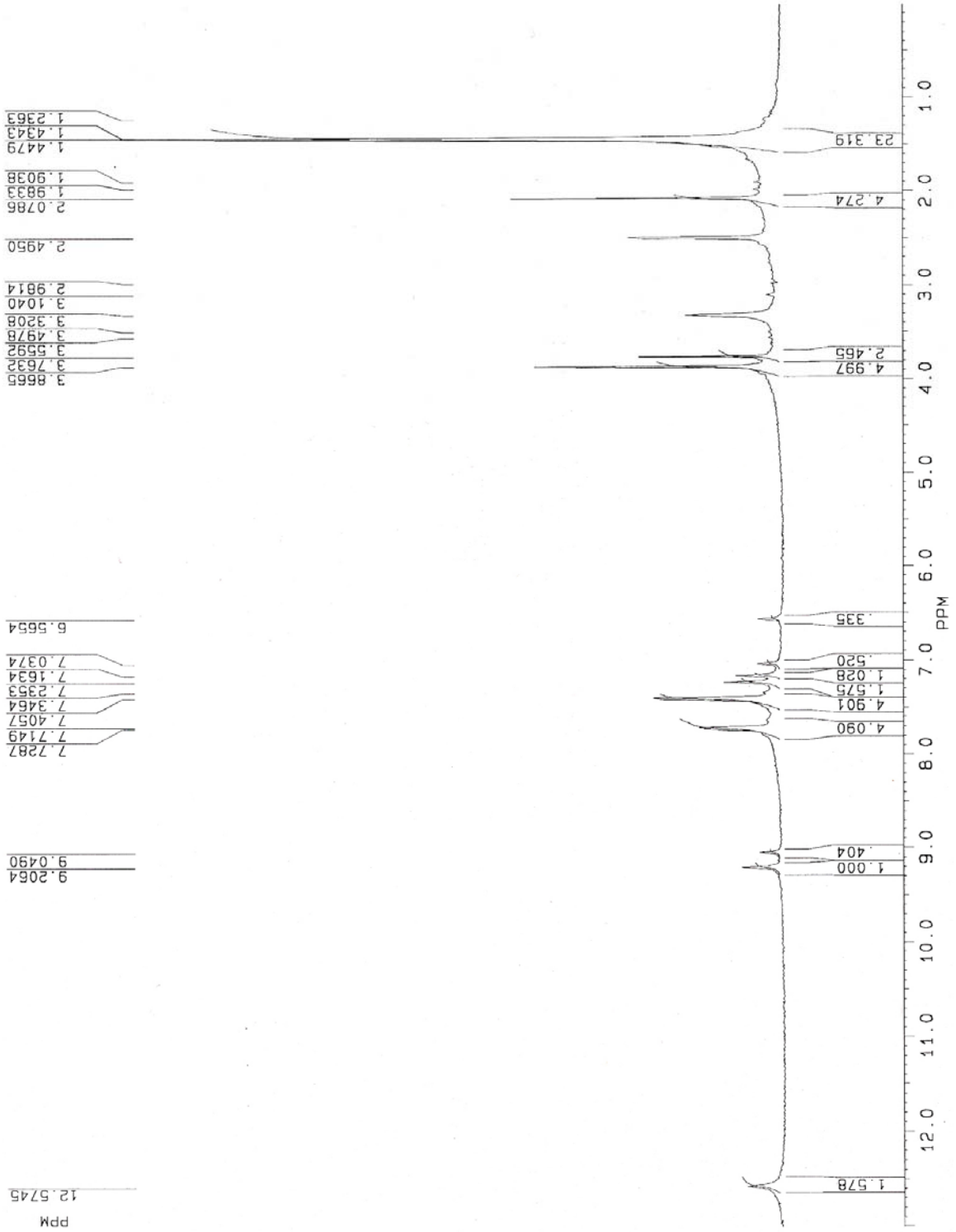


Figure 52. 1H-NMR Spectrum of compound **7** in DMSO-d_6.

VII. Appendix A: Statistical Difference in Percent Change in Luciferase Expression

Table V. Significant Difference between the Percent Change in Luciferase Expression Levels Caused by the Various Tethered-TFOs for Individual pBLP Plasmids at a 95% Confidence Level

		none	STFO	LTFO	bSTFO	bLTFO
pBLP77	none	-	Y	Y	N	N
	STFO	Y	-	Y	Y	Y
	LTFO	Y	Y	-	N	N
	bSTFO	N	Y	N	-	N
	bLTFO	N	Y	N	N	-
pBLP80	none	-	Y	N	N	N
	STFO	Y	-	Y	Y	Y
	LTFO	N	Y	-	N	N
	bSTFO	N	Y	N	-	N
	bLTFO	N	Y	N	N	-
pBLP83	none	-	Y	Y	Y	N
	STFO	Y	-	Y	Y	Y
	LTFO	Y	Y	-	N	N
	bSTFO	Y	Y	N	-	N
	bLTFO	N	Y	N	N	-
pBLP86	none	-	Y	N	N	N
	STFO	Y	-	Y	Y	Y
	LTFO	N	Y	-	N	N
	bSTFO	N	Y	N	-	N
	bLTFO	N	Y	N	N	-

Y = Yes; significant difference in the change in expression levels with 95% confidence
N = No; there was no significant difference

Table VI. Significant Difference between the Percent Change in Luciferase Expression Levels Between the pBLP Plasmids for the Various Tethered-TFOs at a 95% Confidence Level

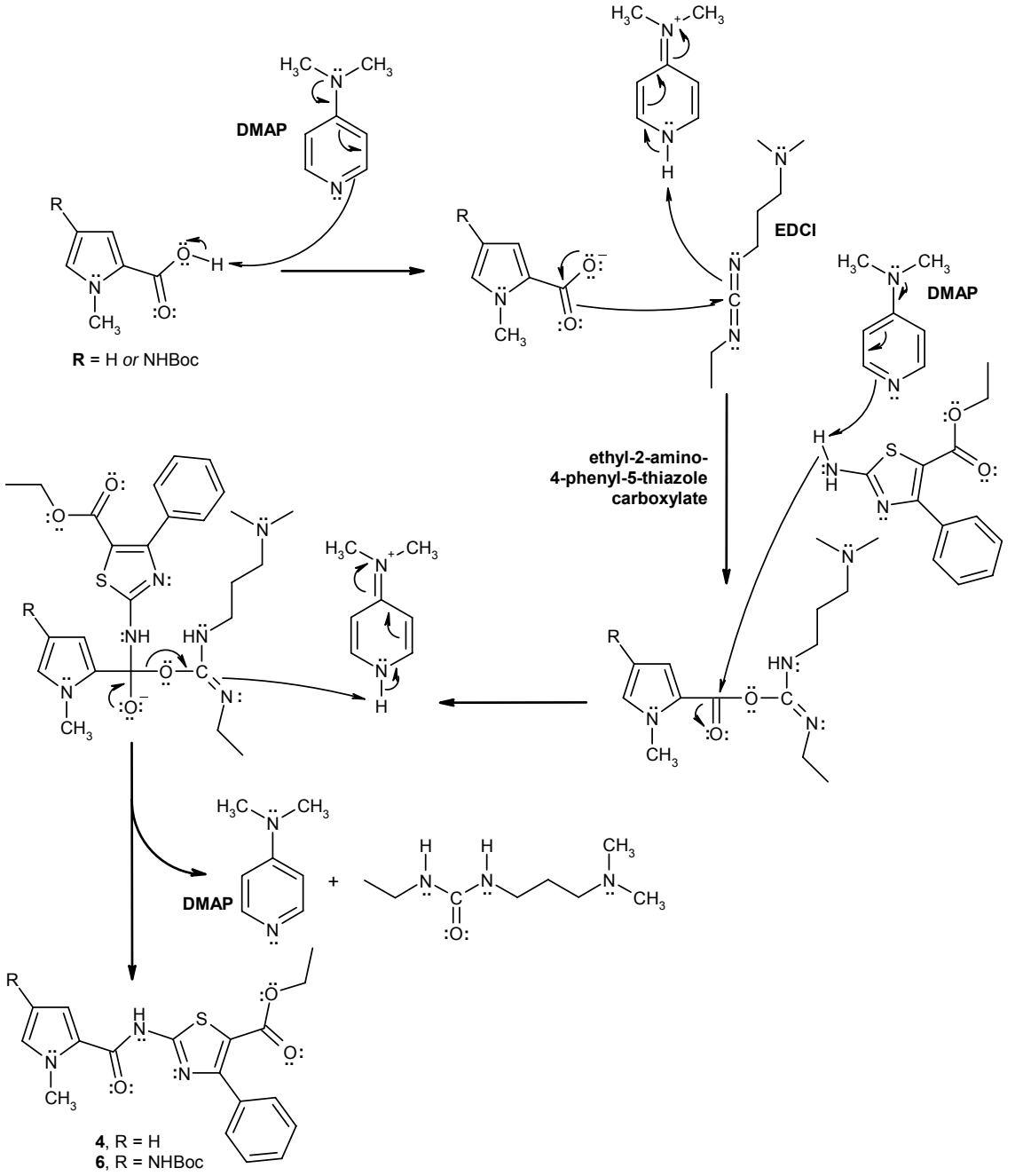
		pBLP77	pBLP80	pBLP83	pBLP86
STFO	**pBLP77**	-	Y	Y	N
	pBLP80	Y	-	Y	Y
	pBLP83	Y	Y	-	Y
	pBLP86	N	Y	Y	-
LTFO	**pBLP77**	-	N	Y	N
	pBLP80	N	-	N	N
	pBLP83	Y	N	-	N
	pBLP86	N	N	N	-
bSTFO	**pBLP77**	-	N	Y	N
	pBLP80	N	-	N	N
	pBLP83	Y	N	-	N
	pBLP86	N	N	N	-
bLTFO	**pBLP77**	-	N	N	N
	pBLP80	N	-	N	N
	pBLP83	N	N	-	N
	pBLP86	N	N	N	-

Y = Yes; significant difference in the change in expression levels with 95% confidence
N = No; there was no significant difference

VIII. Appendix B: Reaction Mechanisms

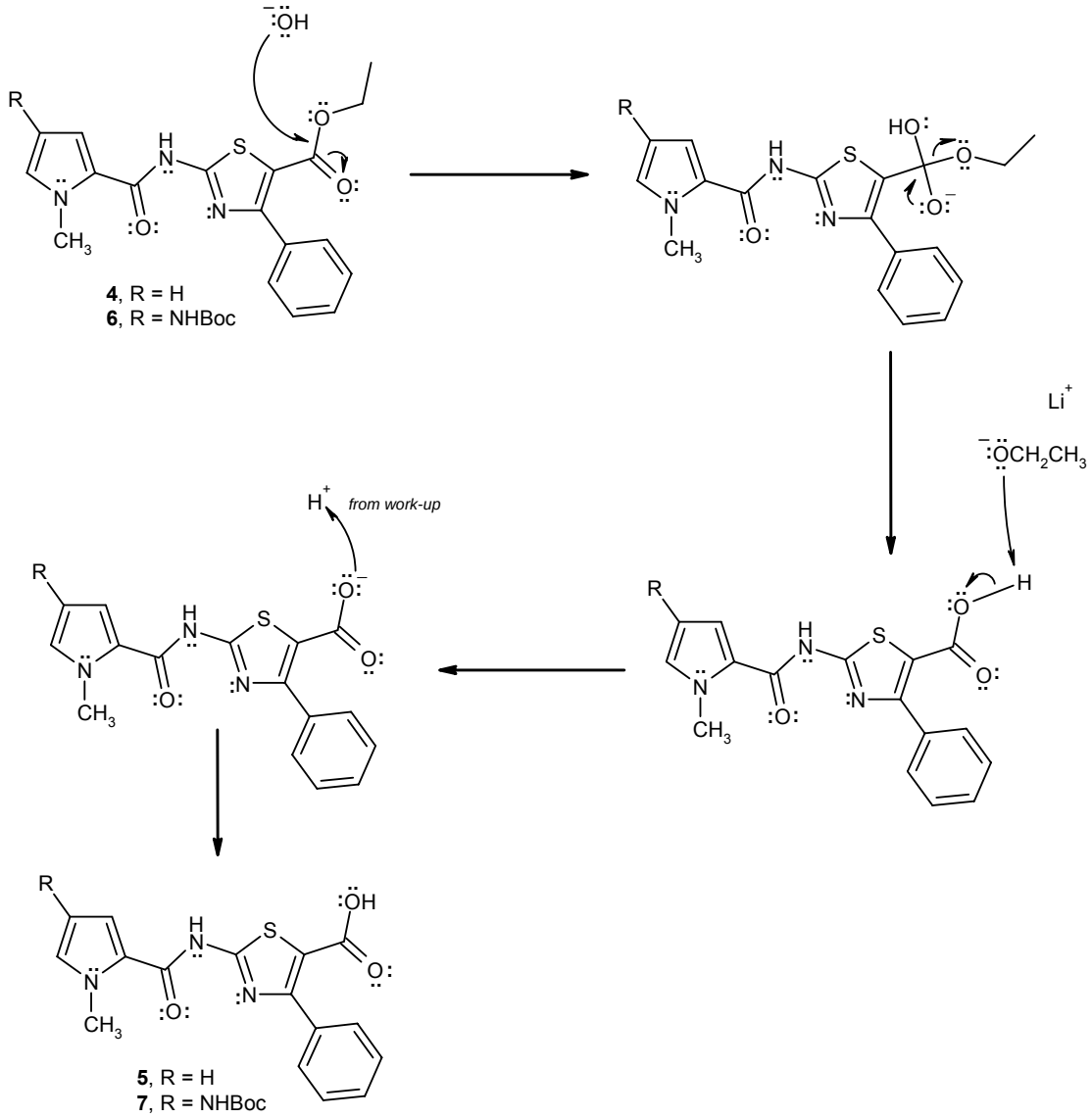
VIII.A. Amide bond formation of Compounds 4 and 6 in Schemes I and II

Condensation of *N*-methylpyrrole-2-carboxylic acid (where R was a proton, H) or 4-tert-butoxycarbonylamino-1-methyl-1H-pyrrole-2-carboxylic acid (where R was –NHBoc) and ethyl-2-amino-4-phenyl-5-thiazole carboxylate with EDCI and DMAP, to form **4** or **6**, schemes I and II.



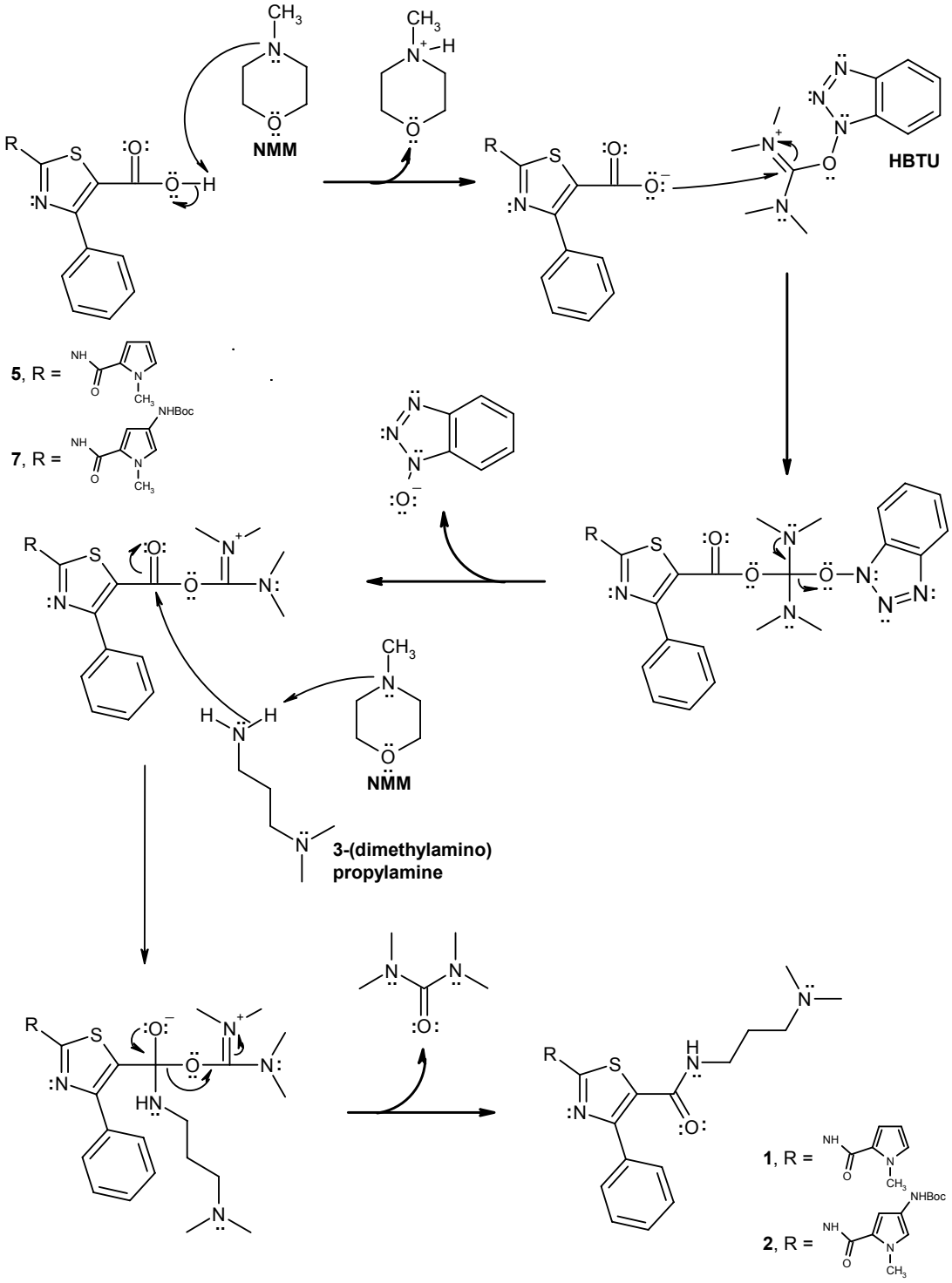
VIII.B. C-terminal ester hydrolysis to form 5 and 7 in Schemes I and II

Hydrolysis of C-terminal esters of compounds **4** and **6** with 1N LiOH to generate free carboxylic acid, compounds **5** and **7**, schemes I and II.



VIII.C. Addition of C-terminal tail to form 1 and 2 in Schemes I and II

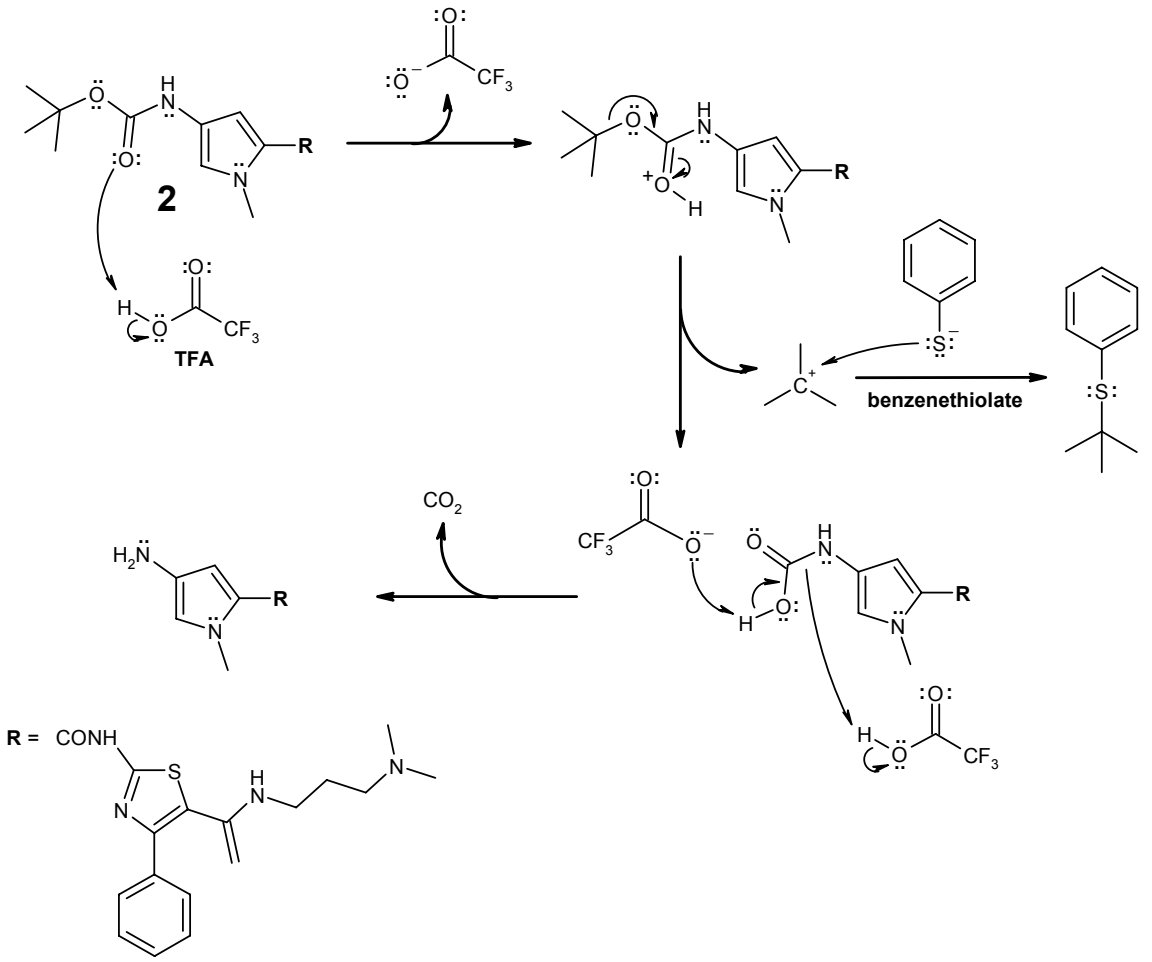
HBTU and NMM coupling conditions were used to form an amide bond between the C-terminal carboxylic acids of compounds **5** and **7** and the primary amine of 3-(dimethylamino)propylamine to form the final products, **1** and **2**, schemes and II.



VIII.D. Scheme III reactions

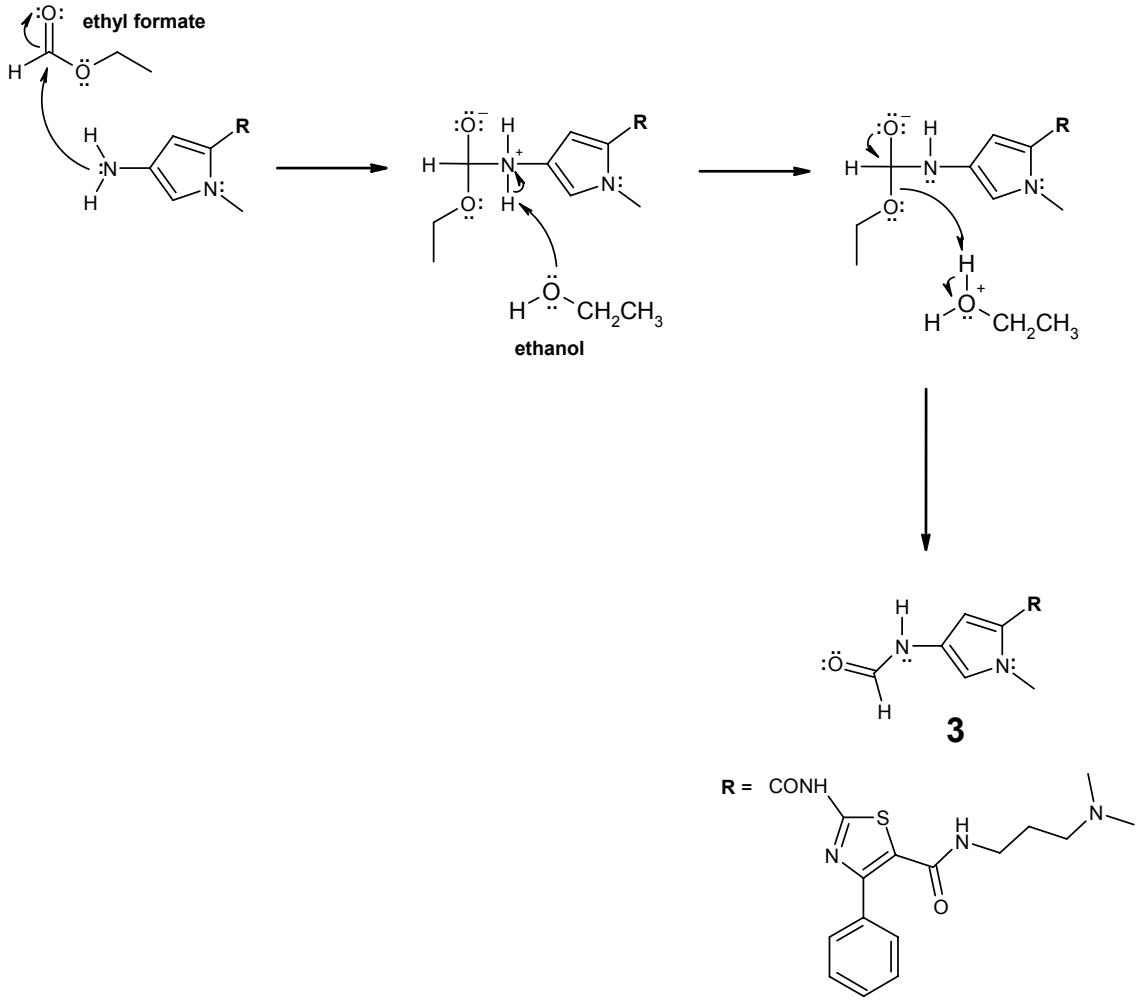
VIII.D.1. *Deprotection of N-terminal Boc group of compound 2*

Deprotection of Boc group of compound **2** with TFA and benzenethiol, as a carbocation scavenger, to yield a free amine at the N-terminus of the molecule, scheme III.



VIII.D.2. ***Ethyl formate reaction to form N-terminal formamide of 3***

The N-terminal free amine generated by Boc group deprotection was immediately reacted with ethyl formate to generate the formamide, compound **3**, scheme III.



IX. REFERENCES

1. Liao, D.J. and Dickson, R.B. C-Myc in breast cancer. *Endocr. Relat. Cancer* **2000**, *7*, 143-164.
2. Yu, D. and Hung, M.C. Overexpression of ErbB2 in cancer and ErbB2-targeting strategies. *Oncogene* **2000**, *19*, 6115-6121.
3. Chen, J.S.; Coustan-Smith, E.; Suzuki, T.; Neale, G.A.; Mihara, K.; Pui, C.H. and Camoana, D. Identification of novel markers for monitoring minimal residual disease in acute lymphoblastic leukemia. *Blood* **2001**, *97*, 2115-2120.
4. Ho, L.; Guo, Y.; Spielman, L.; Petrescu, O.; Haroutunian, V.; Purohit, D.; Czernik, A.; Yemul, S.; Aisen, P.S.; Mohs, R. and Pasinetti, G.M. Altered expression of a-type but not b-type synapsin isoform in the brain of patients at high risk for Alzheimer's disease assessed by DNA microarrays. *Neurosci. Lett.* **2001**, *298*, 191-194.
5. Brooks, W.M.; Lynch, P.J.; Ingle, C.C.; Hatton, A.; Emson, P.C.; Faull, R.L. and Starkey, M.P. Gene expression profiles of metabolic enzyme transcripts in Alzheimer's disease. *Brain Res.* **2007**, *1127*, 127-135.
6. Wahlestedt, C.; Salmi, P.; Good, L.; Kela, J.; Johnsson, T.; Hökfelt, T.; Broberger, C.; Porreca, F.; Lai, J.; Ren, K.; Ossipov, M.; Koshkin, A.; Jakobsen, N.; Skouv, J.; Oerum, H.; Jacobsen, M.H. and Wengel, J. Potent and nontoxic antisense oligonucleotides containing locked nucleic acids. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 5633-5638.
7. Braasch, D.A.; Liu, Y. and Corey, D.R. Antisense inhibition of gene expression in cells by oligonucleotides incorporating locked nucleic acids: effect of mRNA target sequence and chimera design. *Nucleic Acids Res.* **2002**, *30*, 5160-5167.
8. Summerton, J. and Weller, D. Morpholino antisense oligomers: design, preparation and, properties. *Antisense Nucleic Acid Drug Dev.* **1997**, *7*, 187-195.
9. Stancheva, I.; Collins, A.L.; Van den Veyver, I.B.; Zoghbi, H. and Meehan, R.R. A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos. *Mol. Cell* **2003**, *12*, 425-435.
10. Draper, B.W.; Morcos, P.A. and Kimmel, C.B. Inhibition of zebrafish fgf8 pre-mRNA splicing with morpholinio oligos: A quantifiable method of gene knockdown. *Genesis* **2001**, *30*, 154-156.
11. Matter, N. and König, H. Targeted 'knockdown' of spliceosome function in mammalian cells. *Nucleic Acids Res.* **2005**, *33*, e41.
12. Elbashir, S.M.; Harborth, J.; Lendeckel, W.; Yalcin, A.; Weber, K.; Tuschi, T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **2001**, *411*, 494-498.

13. Andersson, M.; Melander, M.; Pojmark, P.; Larsson, H.; Bulow, L.; Hofvander, P. Targeted gene suppression by RNA interference: An efficient method for production of high-amylose potato lines. *J. Biotechnol.* **2006**, *123*, 137-148.

14. Ho, S.N.; Biggar, S.R.; Spencer, D.M.; Schreiber, S.L.; Crabtree, G.R. Dimeric ligands define a role for transcriptional activation domains in reinitiation. *Nature* **1996**, *382*, 822-826.

15. Blancafort, P.; Segal, D.J.; Barbas, C.F. Designing transcription factor architectures for drug discovery. *Mol. Pharmacol.* **2004**, *66*, 1361-1371.

16. Klug, A. Towards therapeutic applications of engineered zinc finger proteins. *FEBS Lett.* **2005**, *579*, 892-894.

17. Elrod-Erickson, M.; Benson, T.E.; Pabo, C.O. High resolution structures of variant Zif268-DNA complexes: implications for understanding zin finger-DNA recognition. *Structure* **1998**, *6*, 451-464.

18. Kim, J.S.; Kim, J.; Cepek, K.L.; Sharp, P.A.; Pabo, C.O. Design of TATA box-binding protein/zinc finger fusions for targeted regulation of gene expression. *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3616-3620.

19. Beerli, R.R.; Segal, D.J.; Dreier, B.; Barbas, C.F. Toward controlling gene expression at will: Specific regulation of the *erbB-2*/*HER-2* promoter by using polydactyl zinc finger proteins constructed from modular building blocks. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14628-14633.

20. Blancafort, P.; Magnenat, L.; Barbas, C.F. Scanning the human genome with combinatorial transcription factor libraries. *Nat. Biotechnol.* **2003**, *21*, 269-274.

21. Graslund, T.; Li, X.; Magnenat, L.; Popkov, M.; Barbas, C.F. Exploring strategies for the design of artificial transcription factors. *J. Biol. Chem.* **2005**, *280*, 3707-3714.

22. Fry, C.J.; Farnham, P.J. Context –dependent transcriptional regulation. *J. Biol. Chem.* **1999**, *274*, 29583-29586.

23. Chandler, S.P.; Fox, K.R. Specificity of antiparallel DNA triple helix formation. *Biochemistry* **1996**, *35*, 15038-15048.

24. Frank-Kamenetskii, M.D.; Mirkin, S.M. Triplex DNA structures. *Annu. Rev. Biochem.* **1995**, *64*, 65-95.

25. Keppler, M.D.; Neidle, S.; Fox, K.R. Stabilisation of TG- and AG-containing antiparallel DNA triplexes by triplex-binding ligands. *Nucleic Acids Res.* **2001**, *29*, 1935-1942.

26. Besch, R.; Giovannangeli, C.; Degitz, K. Triplex-forming oligonucleotides – Sequence-specific DNA ligands as tools for gene inhibition and for modulation of DNA-associated functions. *Curr. Drug Targets* **2004**, *5*, 691-703.

27. Cooney, M.; Czernuszewicz, G.; Postel, E.H.; Flint, S.J.; Hogan, M.E. Site-specific oligonucleotide binding represses transcription of the human c-*myc* gene in vitro. *Science* **1988**, *241*, 456-459.

28. Joseph, J.; Kandala, J.C.; Veerapanane, D.; Weber, K.T.; Guntka, R.V. Antiparallel polypurine phosphothioate oligonucleotides form stable triplexes with the rat α1(I) collagen gene promoter and inhibit transcription in cultured rat fibroblasts. *Nucleic Acids Res.* **1997**, *25*, 2182-2188.

29. Carbone, G.M.; McGuffie, E.M.; Collier, A.; Catapano, C.V. Selective inhibition of transcription of the Ets2 gene in prostate cancer cells by a triplex-forming oligonucleotide. *Nucleic Acids Res.* **2003**, *31*, 833-843.

30. Carbone, G.M.; Napoli, S.; Valentini, A.; Cavalli, F.; Watson, D.K.; Catapano, C.V. Triple DNA-mediated downregulation of Ets2 expression results in growth inhibition and apoptosis in human prostate cancer cells. *Nucleic Acids Res.* **2004**, *32*, 4358-4367.

31. Mayfield, C.; Ebbinghaus, S.; Gee, J.; Jones, D.; Rodu, B.; Squibb, M.; Miller, D. Triplex formation by the human Ha-*ras* promoter inhibits Sp1 binding and *in vitro* transcription. *J. Biol. Chem.* **1994**, *269*, 18232-18238.

32. Ritchie, S.; Boyd, F.M.; Wong, J., Bonham, K. Transcription of the human c-Src promoter is dependent on Sp1, a novel pyrimidine binding factor Spy, and can be inhibited by triplex-forming oligonucleotides. *J. Biol. Chem.* **2000**, *275*, 847-854.

33. Bello-Roufai, M.; Roulon, T.; Escude C. Ligand-mediated transcription elongation control using triplex-based padlock oligonucleotides. *Chem. Biol.* **2004**, *11*, 509-516.

34. Mojzisek, M. Triplex forming oligonucleotides – Tool for gene targeting. *Acta Medica (Hradec Kralove)* **2004**, *47*, 151-156.

35. Besch, R.; Marschall, C.; Schuh, T.; Giovannangeli, C.; Kammerbauer, C.; Degitz, K. Triple helix-mediated inhibition of gene expression is increased by PUVA. *J. Invest. Dermatol.* **2004**, *122*, 1114-1120.

36. Kalish, J.M. and Glazer, P.M. Targetd genome modification via triple helix formation. *Ann. N.Y. Acad. Sci.* **2005**, *1058*, 151-161.

37. Rogers, F.A.; Lloyd, J.A.; Glazer, P.M. Triplex-forming oligonucleotides as potential tools for modulation of gene expression. *Curr. Med. Chem.- Anti-Cancer Agents* **2005**, *5*, 319-326.

38. Ghosh, M.K.; Katyal, A.; Chandra, R.; Brahmachari, V. Targeted activation of transcription *in vivo* through hairpin-triplex forming oligonucleotides in *Saccharomyces cerevisiae*. *Mol. Cell. Biochem.* **2005**, *278*, 147-155.

39. Seidman, M.M.; Puri, N.; Majumdar, A.; Cuenoud, B.; Miiller, P.S.; Alam, R. The development of bioactive triple helix-forming oligonucleotides. *Ann. N.Y. Acad. Sci.* **2005**, *1058*, 119-127.

40. Wang, G.; Xu, X.; Pace, B.; Dean, D.A.; Glazer, P.M.; Chan, P.; Goodman, S.R.; Shokolenko, I. Peptide nucleic acid (PNA) binding-mediated induction of human γ-globin gene expression. *Nucleic Acids Res.* **1999**, *27*, 2806-2813.

41. Cogoi, S.; Codognotto, A.; Rapozzi, V.; Meeuwenoord, N.; van der Marel, G.; Xodo, L.E. Transcription inhibition of oncogenic *KRAS* by a mutation-selective peptide

nucleic acid conjugated to the PKKKRKV nuclear localization signal peptide. *Biochem.* **2005**, *44*, 10510-10519.

42. Boffa, L.C.; Cutrona, G.; Cilli, M.; Matis, S.; Damonte, G.; Mariani, M.R.; Millo, E.; Moroni, M.; Roncella, S.; Fedeli, F. and Ferrarini, M. Inhibition of Burkitt's lymphoma cells growth in SCID mice by a PNA specific for a regulatory sequence of the translocated c-*myc*. *Cancer Gene Therapy* **2007**, *14*, 220-226.
43. Mollegaard, N.E.; Buchard, O.; Egholm, M.; Nielsen, P.E. Peptide nucleic acid•DNA strand displacement loops as artificial transcription promoters. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3892-3895.
44. Janowski, B.A.; Kaihatsu, K.; Huffman, K.E.; Schwartz, J.C.; Ram, R.; Hardy, D.; Mendelson, C.R.; Corey, D.R. Inhibiting transcription of chromosomal DNA with antigene peptide nucleic acids. *Nat. Chem. Biol.* **2005**, *1*, 210-215.
45. Turner, J.J.; Ivanova, G.D.; Verbeure, B.; Williams, D.; Arzumanov, A.A.; Abes, S.; Lebleu, B.; Gait, M.J. Cell-penetrating peptide conjugates of peptide nucleic acids (PNA) as inhibitoes of HIV-1 Tat-dependent *trans*-activation in cells. *Nucleic Acids Res.* **2005**, *33*, 6837-6849.
46. Kritzer, J.A.; Lear, J.D.; Hodson, M.E.; Schepartz, A. Helical β-peptide inhibitors of the p53-hDM2 interaction. *J. Am. Chem. Soc.* **2004**, *126*, 9468-9469.
47. Vassilev, L.T.; Vu, B.T.; Graves, B.; Carvajal, D.; Podlaski, F.; Filipovic, Z.; Kong, N.; Kammlott, U.; Lukacs, C.; Klein, C.; Fotouhi, N.; Liu, E.A. In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science* **2004**, *303*, 844-848.
48. Uytterhoeven, K.; Sponer, J.; Van Meervelt, L. Two 1:1 binding modes for distamycin in the minor groove of d(GGCCAATTGG). *Eur. J. Biochem.* **2002**, *269*, 2868-2877.
49. Pelton, J.G.; Wemmer, D.E. Structural characterization of a 2:1 distamycin A•d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5723-5727.
50. Baliga, R.; Crothers, D.M. On the kinetics of distamycin binding to its target sites on duplex DNA. *Proc. Natl. Acad. Sci.USA* **2000**, *97*, 7814-7818.
51. Dervan, P.B. Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2215-2235.
52. Renneberg, D.; Dervan, P.B. Imidazopyridine/Pyrrole and Hydroxypyrrole/Pyrrole Pairs for DNA Minor Groove Recognition. *J. Am. Chem. Soc.* **2003**, *125*, 5707-5716.
53. White, S.; Szewczyk, J.W.; Turner, J.M.; Baird, E.E.; Dervan, P.B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998**, *391*, 468-471.
54. Marques, M.A.; Doss, R.M.; Foister, S.; Dervan, P.B. Expanding the repertoire of heterocyclic ring pairs for programmable minor groove DNA recognition. *JACS* **2004**, *126*, 10339-10349.

55. Bailly, C.; Chaires, J.B. Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues. *Bioconj. Chem.* **1998**, *9*, 513-538.

56. Melander, C.; Burnett, R.; Gottesfeld, J.M. Regulation of gene expression with pyrrole-imidazole polyamides. *J. Biotechnol.* **2004**, *112*, 195-220.

57. Gottesfeld, J.M.; Neely, L.; Trauger, J.W.; Baird, E.E.; Dervan, P.B. Regulation of gene expression by small molecules. *Nature* **1997**, *387*, 202-205.

58. Dervan, P.B.; Doss, R.M.; Marques, M.A. Programmable DNA binding oligomers for control of transcription. *Curr. Med. Chem.- Anti-Cancer Agents* **2005**, *5*, 373-387.

59. Lai, Y.M.; Fukuda, N.; Ueno, T.; Matsuda, H.; Saito, S.; Matsumoto, K.; Ayame, H.; Bando, T.; Sugiyama, H.; Mugishima, H.; Serie, K. Synthetic pyrrole-imidazole polyamide inhibits expression of the human transforming growth factor-β1 gene. *J. Pharmacol. Exp. Ther.* **2005**, *315*, 571-575.

60. Olenyuk, B.Z.; Zhang, G.J.; Klco, J.M.; Nickols, N.G.; Kaelin, W.G. Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16768-16773.

61. Mapp, A.K.; Ansari, A.Z.; Ptashne, M.; Dervan, P.B. Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. USA*, **2000**, *97*, 3930-3935.

62. Ansari, A.Z.; Mapp, A.K.; Nguyen, D.H.; Dervan, P.B.; Ptashne, M. Towards a minimal motif for artificial transcriptional activators. *Chem. Biol.* **2001**, *8*, 538-592.

63. Minter, A.R.; Brennan, B.B.; Mapp, A.K. A small molecule transcriptional activation domain. *J. Am. Chem. Soc.* **2004**, *126*, 10504-10505.

64. Kwon, Y.; Arndt, H.D.; Mao, Q.; Choi, Y.; Kawazoe, Y.; Dervan, P.B.; Uesugi, M. Small molecule transcription factor mimic. *J. Am. Chem. Soc.* **2004**, *126*, 15940-15941.

65. Sinden, R.R. DNA Structure and Function. Academic Press **1994**.

66. Shishkin, O.V.; Sponer, J.; Hobza, P. Intramolecular flexibility of DNA bases in adenine-thymine and guanine-cytosine Watson-Crick base pairs. *J. Mol. Struct.* **1999**, *477*, 15-21.

67. Bash, R.C.; Vargason, J.M.; Cornejo, S.; Ho, P.S.; Lohr, D. Intrinsically bent DNA in the promoter regions of the yeast GAL1-10 and GAL80 genes. *J. Biol. Chem.* **2001**, *276*, 861-866.

68. Grove, A.; Galeone, A.; Mayol, L.; Geiduschek, E.P. Localized DNA flexibility contributes to target site selection by DNA-bending proteins. *J. Mol. Biol.* **1996**, *260*, 120-125.

69. Maher, L.J. Mechanisms of DNA bending. *Curr. Opin. Chem. Biol.* **1998**, *2*, 688-694.

70. Virstedt, J.; Berge, T.; Henderson, R.M.; Waring, M.J.; Travers, A.A. The influence of DNA stiffness upon nucleosome formation. *J. Struct. Biol.* **2004**, *148*, 66-85.

71. Allemand, J.-F.; Cocco, S.; Douarche, N.; Lia, G. Loops in DNA: An overview of experimental and theoretical approaches. *Eur. Phys. J. E*, **2006**, *19*, 293-302.

72. Hagerman, P.J. Sequence-directed curvature of DNA. *Annu. Rev. Biochem.* **1990**, *59*, 755-781.

73. Bailly, C.; Minnock, A.; Waring, M.J. A simple ligation assay to detect effects of drugs on the curvature/flexibility of DNA. *FEBS Lett.* **1996**, *396*, 253-256.

74. MacDonald, D.; Herbert, K.; Zhang, X.; Polgruto, T.; Lu, P. Solution structure of an A-tract DNA bend. *J. Mol. Biol.* **2001**, *306*, 1081-1098.

75. Liberles, D.A.; Dervan, P.B. Design of artificial sequence-specific DNA bending ligands. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 9510-9514.

76. Roque, A.; Orrego, M.; Ponte, I.; Suau, P. The preferential binding of histone H1 to DNA scaffold-associated regions is determined by its C-terminal domain. *Nucleic Acids Res.* **2004**, *32*, 6111-6119.

77. Blomquist, P.; Belikov, S.; Wrange, O. Increased nuclear factor 1 binding to its nucleosomal site mediated by sequence-dependent DNA structure. *Nucleic Acids Res.* **1999**, *27*, 517-525.

78. Crothers, D.M.; Drak, J. Global features of DNA structure by comparison gel electrophoresis. *Methods in Enzymology* **1992**, *212*, 46-71.

79. Bobst, E.V.; Keyes, R.S.; Cao, Y.Y.; Bobst, A.M. Spectroscopic probe for the detection of local DNA bending at an AAA triplet. *Biochemistry* **1996**, *35*, 9309-9313.

80. Tchernaenko, V.; Halvorson, H.R.; Lutter, L.C. Topological measurement of an A-tract bend angle: effect of magnesium. *J. Mol. Biol.* **2004**, *341*, 55-63.

81. Barbic, A.; Zimmer, D.P.; Crothers, D.M. Structural origins of adenine-tract bending. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 2369-2373.

82. Perez-Martin, J.; de Lorenzo, V. Clues and consequences of DNA bending in transcription. *Annu. Rev. Microbiol.* **1997**, *51*, 593-628.

83. Bolshoy, A.; Nevo, E. Ecological genomics of DNA: Upstream bending in prokaryotic promoters. *Genome Research* **2000**, *10*, 1185-1193.

84. Yokota, H.; Fung, K.; Trask, B.J.; van den Engh, G.; Sarikaya, M.; Aebersold, R. Sharp DNA bends as landmarks of protein-binding sites on straightened DNA. *Anal. Chem.* **1999**, *71*, 1663-1667.

85. Hizver, J.; Rozenberg, H,; Frolow, F.; Rabinovich, D.; Shakked, Z. DNA bending by an adenine-thymine tract and its role in gene expression. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 8490-8495.

86. Tomky, L.A.; Straus-Soukup, J.K.; Maher, L.J. Effects of phosphate neutralization on the shape of the AP-1 transcription factor binding site in duplex DNA. *Nucleic Acids Res.* **1998**, *26*, 2298-2305.

87. Tijan, R.; Maniatis, T. Transcriptional activation: A complex puzzle with few easy pieces. *Cell* **1994**, *77*, 5-8.

88. Kim, J.; de Haan, G.; Shapiro, D.J. DNA bending between upstream activator sequences increases transcriptional synergy. *Biochemical and Biophysical Research Communications* **1996**, *226*, 638-644.

89. Werner, M.H.; Gronenborn, A.M.; Clore, G.M. Intercalation, DNA kinking and the control of transcription. *Science* **1996**, *271*, 778-784.

90. Sjottem, E.; Andersen, C.; Johansen, T. Structural and functional analyses of DNA bending induced by Sp1 family transcription factors. *J. Mol. Biol.* **1997**, *267*, 490-504.

91. Kerppola, T.K. Transcriptional cooperativity: bending over backwards and doing the flip. *Structure* **1998**, *6*, 549-554.

92. Ross, E.D.; Keating, A.M.; Maher, L.J. DNA constraints on transcription activation in virto. *J. Mol. Biol.* **2000**, *297*, 321-334.

93. Scaffidi, P.; Bianchi, M.E. Spatially precise DNA bending is an essential activity of the Sox2 transcription factor. *J. Biol. Chem.* **2001**, *276*, 47296-47302.

94. Rivetti, C.; Codeluppi, S.; Dieci, G.; Bustamante, C. Visulaizing RNA extrusion and DNA wrapping in transcription elongation complexes of bacterial and eukaryotic RNA polymerases. *J. Mol. Biol.* **2003**, *326*, 1413-1426.

95. Coulombe, B. DNA wrapping in transcription initiation by RNA polymerase II. *Biochem. Cell Biol.* **1999**, *77*, 257-264.

96. Rivetti, C.; Guthold, M.; Bustamante, C. Wrapping of DNA around the *E. coli* RNA polymerase open promoter complex. *EMBO J.* **1999**, *18*, 4464-4475.

97. Starr, D.B.; Hoopes, B.C.; Hawley, D.K. DNA bending is an important component of site-specific recognition by the TATA binding protein. *J. Mol. Biol.* **1995**, *250*, 434-446.

98. Pérez-Lago, L.; Salas, M.; Camacho, A. A precise DNA bend angle is essential for the function of the phage ϕ29 transcriptional regulator. *Nucleic Acids Res.* **2005**, 126-134.

99. Powell, R.M.; Parkhurst, K.M.; Brenowitz, M.; Parkhurst, L.J. Marked stepwise differences within a common kinetic mechanism characterize TATA-binding protein interactions with two consensus promoters. *J. Biol. Chem.* **2001**, *276*, 29782-29791.

100. Wu, J.; Parkhurst, K.M.; Powell, R.M.; Brenowitz, M.; Parkhurst, L.J. DNA bends in TATA-binding protein•TATA complexes in solution are DNA sequence-dependent. *J. Biol. Chem.* **2001**, *276*, 14614-14622.

101. Gartenberg, M.R.; Crothers, D.M. Synthetic DNA bending sequences increase the rate of in vitro transcription initiation at the Escherichia coli lac promoter. *J. Mol. Biol.* **1991**, *219*, 217-230.

102. Bracco, L.; Koltarz, D.; Kolb, A.; Diekmann, S.; Buc, H. Synthetic curved DNA sequences can act as transcriptional activators in Escherichia coli. *EMBO J.* **1989**, *8*, 4289-4296.

103. Aiyar, S.E.; Gourse, R.L. & Ross, W. Upstream A-tracts increase bacterial promoter activity through interactions with the TNA polymerase α subunit. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14652-14657.

104. Perez-Martin, J.; Espinosa, M. Protein-induced bending as a transcriptional switch. *Science* **1993**, *260*, 805-807.

105. Déthiollaz, S.; Eichenberger, P.; Geiselmann, J. Influence of DNA geometry on transcriptional activation in Escherichia coli. *EMBO J.* **1996**, *15*, 5449-5458.

106. Uil, T.G.; Haisma, H.J.; Rots, M.G. Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities. *Nucleic Acids Res.* **2003**, *31*, 6064-6078.

107. Henderson, D.; Hurley, L.H. Molecular struggle for transcriptional control. *Nature Medicine* **1995**, *1*, 525-527.

108. Akiyama, T.; Hogan, M.E. The design of an agent to bend DNA. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 12122-12127.

109. Akiyama, T.; Hogan, M.E. Structural analysis of DNA bending induced by tethered triple helix forming oligonucleotides. *Biochemistry* **1997**, *36*, 2307-2315.

110. Akiyama, T.; Hogan, M.E. Microscopic DNA flexibility analysis. *J. Biol. Chem.* **1996**, *271*, 29126-29135.

111. Kessler, D.J.; Pettitt, B.M.; Cheng, Y.K.; Smith, S.R.; Jayaraman, K.; Vu, H.M.; Hogan, M.E. Triple helix formation at distant sites: Hybrid oligonucleotides containing a polymeric linker. *Nucleic Acids Res.* **1993**, *21*, 4810-4815.

112. Sidorova, N.Y.; Gazoni, P.; Rau, D.C. Competition between netropsin and restriction nuclease EcoRI for DNA binding. *J. Mol. Struct. Dyn.*, **1995**, *13*, 367-385.

113. Garcia-Nieto, R.; Manzanares, T.; Cuevas, C.; Gago, F. Increased DNA binding specificity for antitumor ecteinascidin 743 through protein-DNA interactions? *J. Med. Chem.* **2000**, *43*, 4367-4369.

114. Cuevas, C.; Perez, M.; Martin, M.J.; Chicarro, J.L.; Fernandez-Rivas, C.; Flores, M.; Franesch, A.; Gallego, P.; Zarzuelo, M.; de la Calle, F.; Garcia, J.; Polanco, C.; Rodriguez, I.; Manzanares, I. Synthesis of ecteinascidin ET-743 and phthalascidin Pt-650 from cyanosafracin B. *Org. Lett.* **2000**, *2*, 2545-2548.

115. Zewail-Foote, M.; Li, V.S.; Kohn, H.; Bearss, D.; Guzman, M.; Hurley, L.H. The inefficiency of incisions of ecteinascidin 743-DNA adducts by the UvrABC nuclease and the unique structural feature of the DNA adducts can be used to explain the repair-dependent toxicities of this antitumor agent. *Chemistry & Biology*, **2001**, *8*, 1033-1049.

116. D'Incalci, M.; Erba, E.; Damia, G.; Galliera, E.; Carrassa, l.; Marchini, S.; Mantovani, R.; Tognon, G.; Fruscio, R.; Jimeno, J.; Faircloth, G.T. Unique features of the mode of action of ET-743. *Oncologist*, **2002**, *7*, 210-216.

117. Ikemoto, N.; Kumar, R.A.; Ling, T.; Ellestad, G.A.; Danishefsky, S.J. Calicheamicin-DNA complexes: Warhead alignment and saccharide recognition of the minor groove. *Proc. Natl. Acad. Sci.USA*, **1995**, *92*, 10506-10510.

118. Sissi, C.; Aiyar, J.; Boyer, S.; Depew, K.; Danishefsky, S. Interaction of calicheamicin- ζ_1^I and its related carbohydrates with DNA-protein complexes. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 10643-10648.

119. Salzberg, A.A.; Dedon, P.C. DNA bending is a determinant of calicheamicin target recognition. *Biochemistry*, **2000**, *39*, 7605-7612.

120. Wu, J.Y.; Vlastos, A.T.; Pelte, M.F.; Caligo, M.A.; Bianco, A.; Krause, K.H.; Laurent, G.J.; Irminger-Finger, I. Aberrant expression of *BARD1* in breast and ovarian cancers with poor prognosis. *Int. J. Cancer* **2006**, *118*, 1215-1226.

121. Carmeliet, P.; Jain, R.K. Angiogenesis in cancer and other diseases. *Nature* **2000**, *407*, 249-256.

122. Bunn, H.F. Pathogenesis and treatment of sickle cell disease. *N. Engl. J. Med.* **1997**, *337*, 762-769.

123. Bank, A. Regulation of human fetal hemoglobin: new players, new complexities. *Blood* **2006**, *107*, 435-443.

124. Musso, M.; Van Dyke, M.W. Torsionally-strained DNA and intermolecular purine-purine-pyrimidine triple-helix formation. *Mol. Cell. Biochem.* **1996**, *154*, 65-70.

125. Weyermann, P.; Dervan, P.B. Recognition of ten base pairs of DNA by head-to-head hairpin dimers. *J. Am. Chem. Soc.* **2002**, *124*, 6872-6878.

126. Bartulewicz, D.; Markowska, A.; Wołczyński, S.; Dąbrowska, M.; Różański, A. Molecular modeling, synthesis and antitumor activity of carbocyclic analogues of netropsin and distamycin – new carriers of alkylating elements. *Acta Biochim. Pol.* **2000**, *47*, 23-35.

127. Wemmer, D.E. Designed sequence-specific minor groove ligands. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 439-461.

128. Lerman, L.S.; Frisch, H.L. Why does the electrophoretic mobility of DNA in gels vary with the length of the molecule? *Biopolymers* **1982**, *21*, 995-997.

129. Boger, D.L.; Fink, B.E.; Hedrick, M.P. Total synthesis of distamycin A and 2640 analogues: A solution-phase combinatorial approach to the discovery of new, bioactive DNA binding agents and development of a rapid, high-throughput screen for determining relative binding affinity or DNA binding sequence selectivity. *J. Am. Chem. Soc.* **2000**, *122*, 6382-6394.

130. Khalaf, A.I.; Waigh, R.D.; Drummond, A.J.; Pringle, B.; McGroarty, I.; Skellern, G.G.; Suckling, C.J. Distamycin analogues with enhanced lipophilicity: Synthesis and antimicrobial activity. *J. Med. Chem.*, **2004**, *47*, 2133-2156.

131. Wildeson, J. and Murphy, C.J. Intrinsic bending in GGCC tracts as probed by fluorescence resonance energy transfer. *Anal. Biochem.* **2000**, *284*, 99-106.

132. Workman, J.L. Nucleosome displacement in transcription. *Genes Dev.* **2006**, *20*, 2009-2017.

133. Richmond, T.J.; Davey, C.A. The structure of DNA in the nucleosome core. *Nature*, **2003**, *423*, 145-150.

134. Goodwin, K.D.; Long, E.C.; Georgiadis, M.M. A host-guest approach for determining drug-DNA interactions: an example using netropsin. *Nucleic Acids Res.* **2005**, *33*, 4106-4116.