Chapter 1

Introduction to Protein-DNA Dimerizers

1.1 Background and Significance

Since the sequencing of the human genome[1,2] it has been estimated that humans possess 20,000 to 30,000 genes that are expressed in a highly controlled manner by DNA-binding proteins called transcription factors.[3,4] The evolution of more complex organisms is thought to be due not only to gene proliferation, but also increasingly sophisticated transcriptional regulation.[5] This is particularly evident from the observation that prokaryotes often utilize single homodimeric transcription factors to regulate target genes, in contrast to eukaryotes, in which gene expression is frequently controlled in a combinatorial fashion by multiprotein-DNA complexes.[6] Errors in regulating gene expression caused by mutant transcription factors often lead to diseases such as cancer or birth defects, and therapies that counteract these faulty proteins would have a significant impact on human medicine.[7] This thesis describes the design of chemicals called protein-DNA dimerizers that mimic endogenous transcription factors in that they facilitate the formation of protein-DNA complexes. Protein-DNA dimerizers are a promising class of compounds that suggest we may someday engage or replace errant transcription factors and neutralize their deleterious effects for therapeutic purposes.

1.2 Eukaryotic Gene Transcription

The protein-coding regions of the eukaryotic genomes are transcribed by RNA polymerase II. The polymerase physically associates with a number of proteins including TFIIA, TFIIB, TFIID, TFIIE, TFIIF, and TFIIH before initiating transcription.[8] TFIID, also called TATA-binding protein (TBP), binds to the TATA DNA sequence of the core promoter region adjacent to the start site of genes (Figure 1.1).[3,4,8] Additional transcription factors also bind close to the core promoter and further away at up- and downstream enhancer regions which influence the rate of transcription.[4,8] Large mutiprotein complexes mediate the interactions between the these transcription factors and the RNA polymerase II machinery[9] and a unified nomenclature for such mediator proteins has recently been

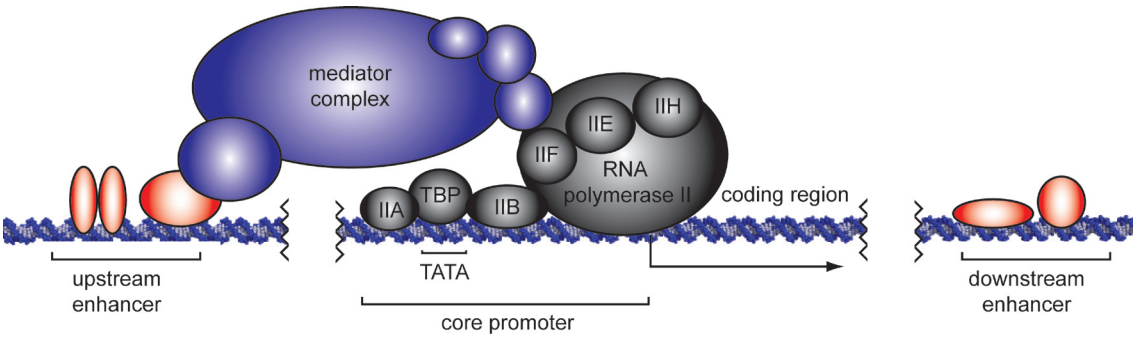

Figure 1.1 A schematic of a gene promoter. The RNA polymerase II transcriptional complex is shown bound at the gene start site near the core promoter. Upstream and downstream enhancer elements are shown to be much farther away from the start site of the gene. Enhancers contain binding sites for transcription factors which often interact indirectly with the RNA polymerase II complex through large multiprotein mediator complexes.

proposed.[10] Gene expression is also affected by the local chromatin structure which is regulated by proteins such as histone-modifying enzymes.[11]

1.3 DNA Recognition by Natural Transcription Factors

Central to the regulation of gene transcription is the sequence specific recognition of DNA by transcription factors. DNA contains a wide major groove and a narrow minor groove that expose unique surfaces for interaction with these proteins (Figure 1.2). Transcription factors are capable of binding to both the major and minor grooves by making defined electrostatic, hydrogen-bonding, and van der Waals contacts with the composite surface of nitrogenous bases and sugar-phosphate backbone of DNA (Figure 1.3). Some proteins, such as TBP[12-15] and the zinc-finger Zif268,[16] bind to DNA as monomers, but transcription factors often bind as part of homo- and heterodimeric complexes. For instance, the leucine-zipper homodimer of GCN4[17] or the heterodimer of Jun-Fos[18] interact through their two coiled-coil domains that position their α-helical DNA-binding domains into the major groove. These protein-DNA complexes are often reiterated in a combinatorial fashion which enables the integration of diverse cellular signals to regulate individual genes. This combinatorial mechanism is exemplified by the interferon-β (IFN-β) enhanceosome in which a total of

eight proteins, including ATF-2/c-Jun, IRF-3/IRF-7, and NF-κB, are required to bind to a conserved stretch of 55 base pairs of DNA in order to activate expression in response to viral infection.[19] Recently, an atomic model of the complete IFN-β enhanceosome has been assembled from three separate x-ray crystal structures (Figure 1.4).[19]

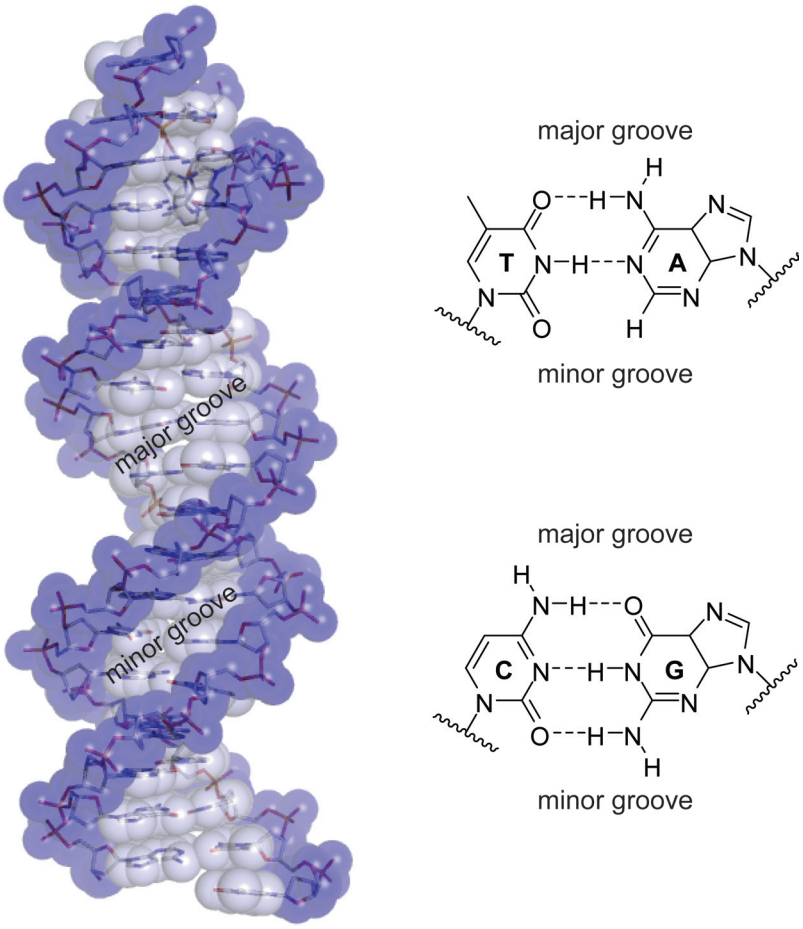

Figure 1.2 The structure of DNA. On the left a crystal structure of a DNA double helix is shown (PDB code 1YSA). On the right, the four nitrogenous bases are shown as Watson-Crick base pairs where A complexes with T and G with C. The exposed edges of the base pairs in the major and minor grooves are indicated which are important for the sequence-specific recognition of DNA by proteins and small molecules.

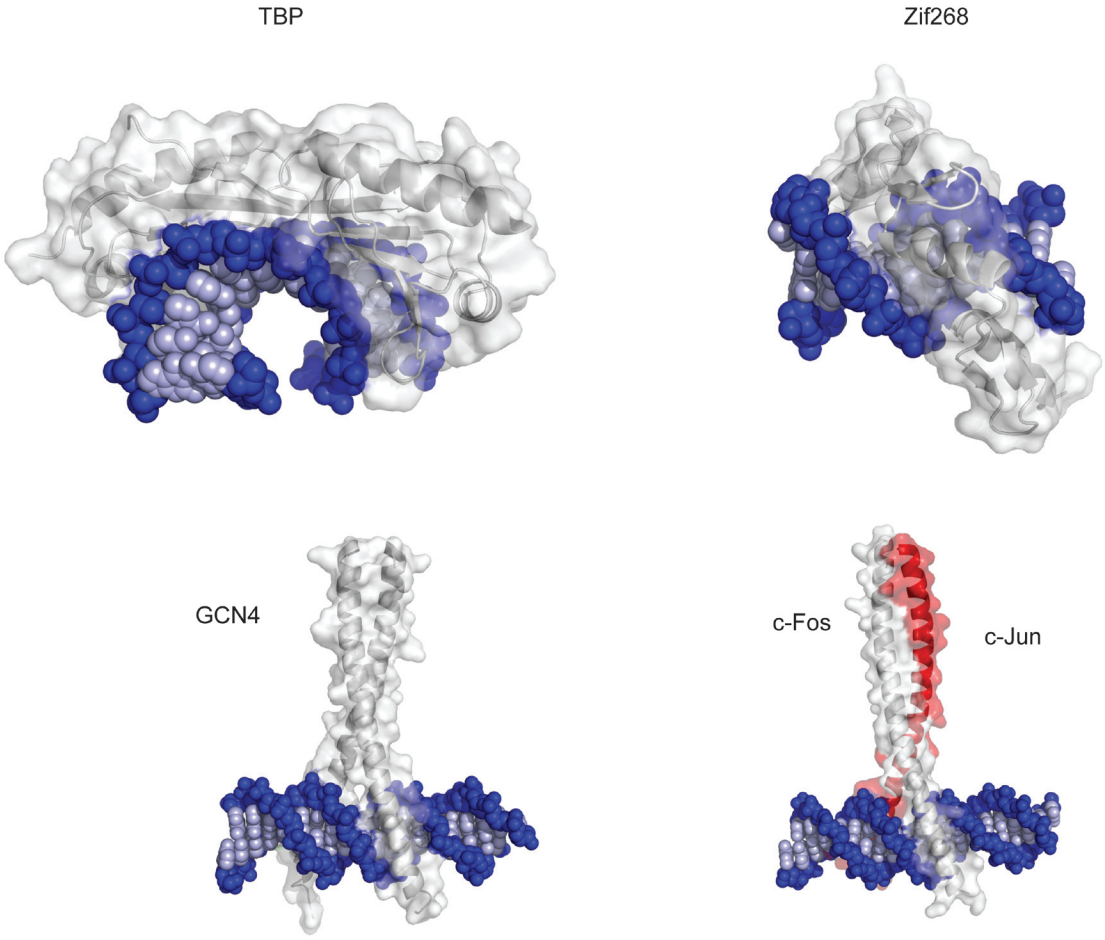

Figure 1.3 Crystal structures of four transcription factor-DNA complexes. The top left shows a structure of TATA-binding protein (TBP) bound to the DNA minor groove as a monomer leading to significant bending and widening of the groove (PDB code 1TGH). The top right shows a structure of zinc-finger protein, Zif268, which consists of three tandem major-groove recognition domains (PDB code 1ZAA). The bottom left shows a GCN4 leucine-zipper homodimer, which inserts its extended α-helical coiled-coil domains into the DNA major grooves (PDB code 1YSA). The bottom right shows a c-Jun/c-Fos heterodimer, which binds to DNA in a similar manner to the GCN4 homodimer (PDB code 1FOS).

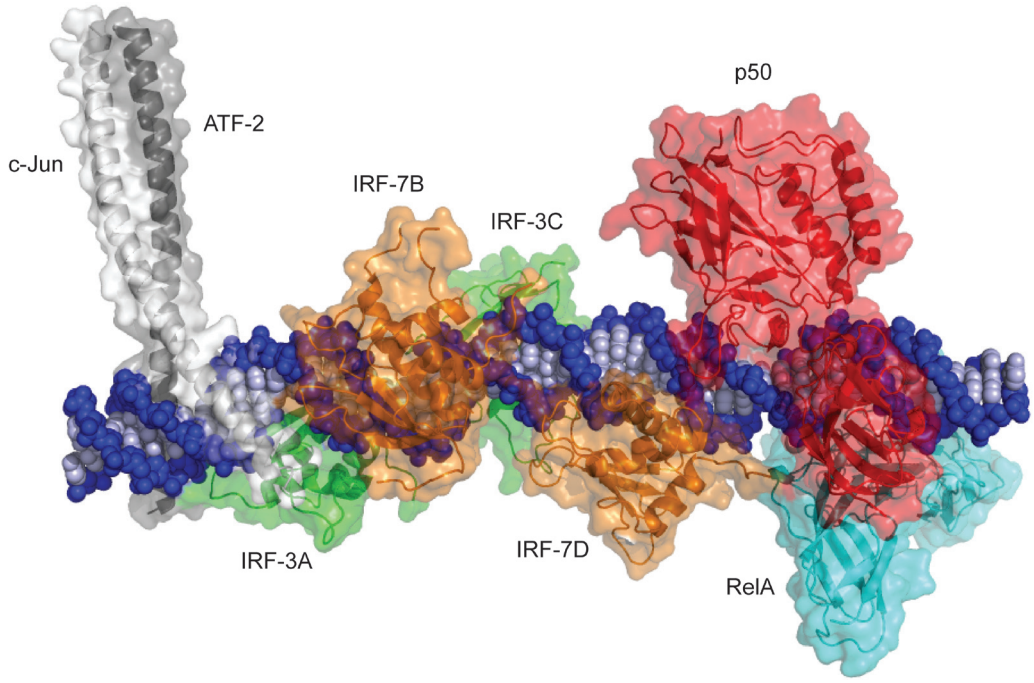

Figure 1.4 An atomic model of the IFN-β enhanceosome in which a total of eight transcription factors bind to a highly conserved 55 base pair segment of DNA (PDB from supplemental of reference 19). All eight transcription factors must bind to the enhanceosome to activate IFN-β expression in response to viral infection.

1.4 HOX Transcription Factors

Homeobox (HOX) proteins are a particular class of transcription factors central to this thesis that are critical for the proper execution of developmental programs of higher organisms.[20-22] HOX genes have been intensely studied in the fruit fly *D. melanogaster* in which they were identified by the abnormal body segment phenotypes caused by mutations to single genes.[20-22] For instance, a mutation to the gene ultrabithorax (Ubx) leads to the development of an additional thorax in place of the fly's balance-sensing organs called halteres. In humans, mutations to HOX proteins are implicated in several forms of cancer and limb deformation diseases.[21,23-26] Thus, HOX transcription factors serve as master regulators of the downstream genes that must be activated or deactivated at precise times during the course of organism development.

In contrast to their specific roles during development, HOX proteins only bind to relatively nonspecific 4 to 6 base pair sequences of DNA.[27-29] By binding to DNA with

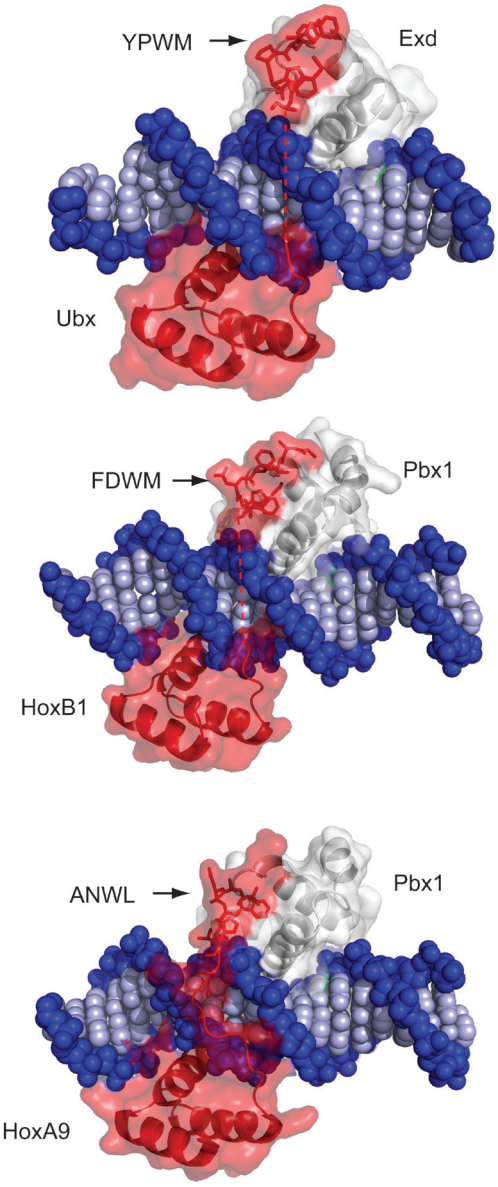

Figure 1.5 Crystal structures of three HOX/TALE/DNA ternary complexes. The top shows the structure of the *D. Melanogaster* developmental transcription factors ultrabithorax (Ubx) and extradenticle (Exd) (PDB code 1B8I). The proteins bind to opposite sides of the DNA duplex primarily in the major groove with their 3 α-helix homeodomains. The conserved YPWM peptide motif of Ubx spans the intervening minor groove and mediates a direct protein-protein contact with an Exd binding pocket between the 1st and 2nd α-helices. The middle structure shows the homologous human transcription factors HoxB1 and Pbx1 in which the FDWM peptide of HoxB1 binds to Pbx1 in a similar manner as the YPWM of the fruit fly proteins (PDB code 1B72). A 4th α-helix was discovered in the Pbx1 structure that stabilizes homeodomain-DNA binding. The bottom structure shows the human proteins HoxA9 and Pbx1 bound to DNA (PDB code 1PUF). The divergent ANWL peptide of HoxA9 contains the conserved W that dominates the protein-protein interface. Unique to this structure, the linking amino acids between the ANWL peptide and the homeodomain are ordered and contact the intervening DNA minor groove.

cofactor proteins, they increase their composite recognition site which is thought to be important for their function in developmental pathways.[27] For instance, Ubx binds the site 5′-TTAT-3′ alone, but in complex with extradenticle (Exd) the composite site increases by four base pairs to 5′-TGATTTAT-3′.[30] A crystal structure of the ternary Ubx/Exd/DNA complex revealed how both Ubx and Exd's major groove binding homeodomains interact via Ubx's N-terminal YPWM peptide motif (Figure 1.5).[30] In fact, all HOX proteins contain a conserved homeodomain which adopts a three α-helix bundle that binds to the DNA major groove. The third α-helix, aptly named the recognition helix, makes the majority of the specific DNA contacts with the edges of nitrogenous bases. The YPWM motif of Ubx spans the DNA minor groove between the two proteins and inserts a hydrophobic tryptophan into the complementary Exd binding pocket. Additional crystal structures of HoxB1/Pbx/DNA and HoxA9/Pbx/DNA which contain human homologues of Ubx and Exd have also been reported in which the conserved tryptophan of the YPWM motif mediates the majority of the protein-protein contacts.[31,32] The YPWM motif is sometimes referred to as a penta- or hexapeptide since two flanking amino acids are also highly conserved yielding an overall consensus sequence of ϕYPWM(K/R) where ϕ is a hydrophobic residue (Figure 1.6).[33] Both Exd and Pbx, sometimes called TALE proteins, contain a three-amino acid loop extension between the first and second α-helices of their divergent homeodomains which line the binding pocket for the YPWM motif. In fruit flies an Exd mutation leads to the increase in the number of denticles (i.e., extra denticles)[34] and in humans a t(1;19) chromosomal translocation at Pbx and E2A loci leads to pre-B cell leukemia (i.e., pre-B cell transforming factor).[35]

The t(1;19) chromosomal translocation leads to an E2A-Pbx1 fusion protein that is observed in 5% of pediatric and 3% of adult B-cell acute lymphoblastic leukemia cases.[36] The E2A-Pbx1 chimeric protein turns Pbx into a constitutive transcriptional activator by fusing the activation domain of E2A with the DNA-binding homeodomain of Pbx1.[37] Like the wild-type Pbx1, E2A-Pbx1 is capable of cooperatively binding to DNA

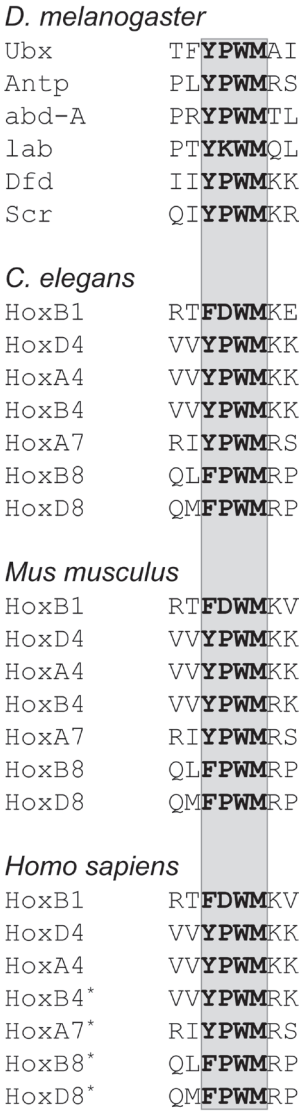

Figure 1.6 Evolutionary conservation of the YPWM motifs from fruit flies to humans. Note that all the motifs shown contain W and M. W is absolutely conserved in natural HOX proteins, but different residues are sometimes found instead of M. (Sequences were compiled from reference 33 except for entries with an asterisk (*) which were obtained from the National Center for Biotechnology Information: www.ncbi.nlm.nih.gov)

with other HOX proteins to the motif 5′-ATCAATCAA-3′.[38] Although the exact mechanism of leukemogenesis is still a matter of some debate, E2A-Pbx1 activates the expression of a novel WNT gene, Wnt16,[39] and a tyrosine kinase, EB-1,[40] which are believed to contribute to oncogenesis. Recently, siRNA inhibition of E2A-Pbx1 and subsequent downstream deactivation of Wnt16 and EB-1 has been shown to increase apoptosis in 697 cells which harbor the t(1;19) translocation.[41] This suggests that interfering with the abnormal function of E2A-Pbx1 may be a possible therapeutic approach for this type of leukemia.

1.5 Programmable DNA-Binding Small Molecules

Nature produces many small molecules which bind to DNA, but the polypyrrole compounds netropsin (**1**) and distamycin (**2**) are of particular relevance to the protein-DNA dimerizers designed in this thesis. Netropsin and distamycin bind to the DNA minor groove of AT-rich tracts and make specific H-bonds with

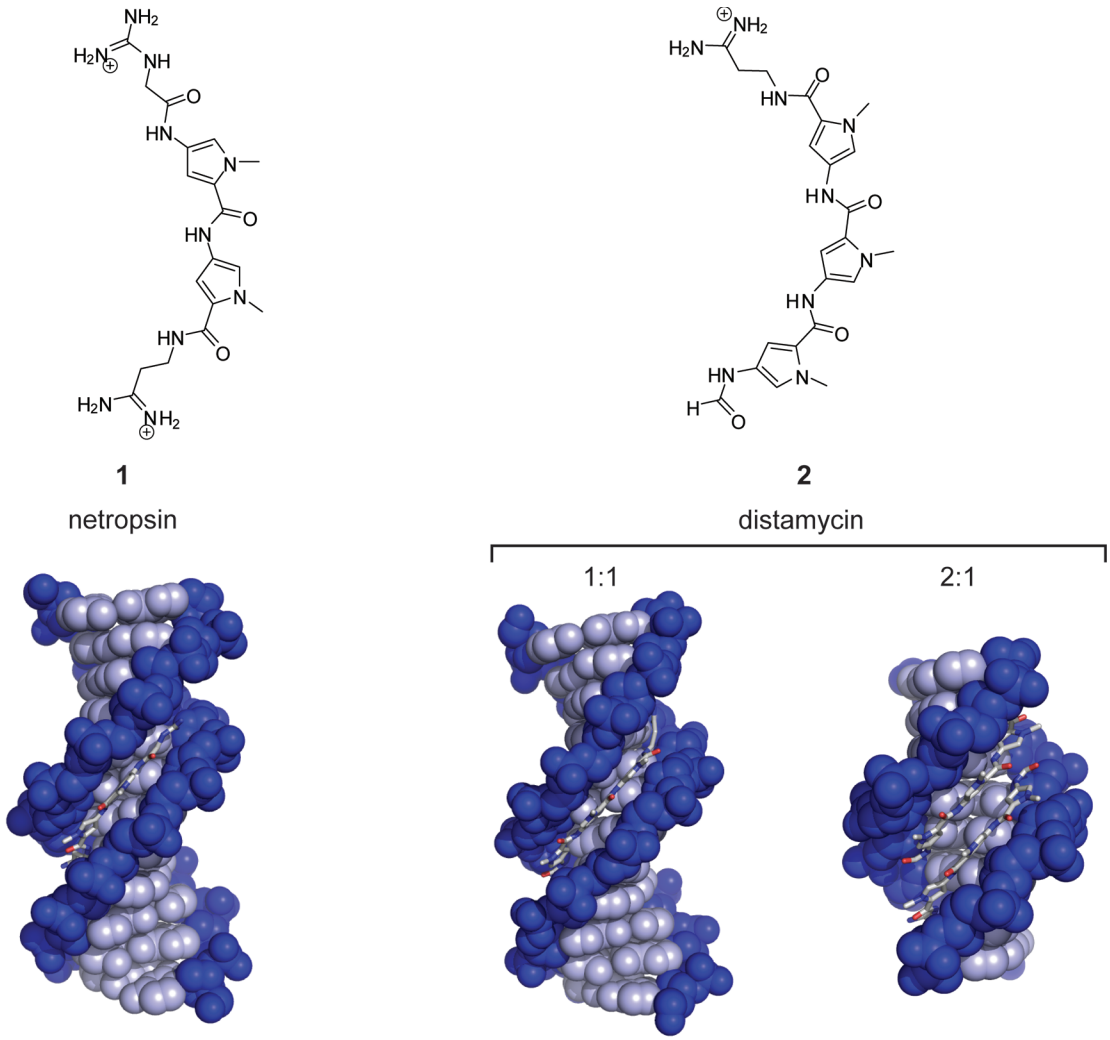

Figure 1.7 Structures illustrating DNA recognition by the natural products netropsin (**1**) and distamycin (**2**). The dicationic dipyrrole netropsin binds to narrow AT tracts of the DNA minor groove as a monomer (PDB code 6BNA). The monocationic tripyrrole distamycin can bind to AT tracts of the DNA minor groove as either a monomer or a dimer. (PDB codes 2DND and 378D).

the edges of the nitrogenous bases (Figure 1.7).[42-43] It was postulated in the mid-1980s that replacement of an *N*-methylpyrrole (Py) of distamycin by an *N*-methylimidazole (Im) would lead to a compound that would bind preferentially to G instead of A or T by making specific contacts to exocyclic amine of G in the minor groove.[44-45] Synthesis of such an ImPyPy distamycin analogue in the Dervan group led to the discovery that this compound actually bound preferentially to 5′-WGWCW-3′ instead of the expected 5′-GWW-3′ (where W = A or T).[46] This result was compellingly rationalized following the publication of

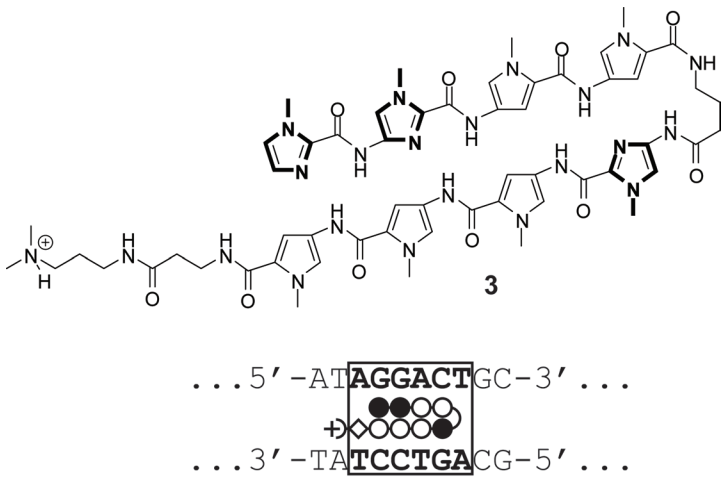

Figure 1.8 The chemical structure and schematic of a typical pyrrole-imidazole hairpin polyamide (**3**). *N*-methylpyrrole (Py) is represented by an open circle, *N*-methylimidazole (Im) by a filled circle, *γ*-aminobutyic acid (γ) by a half circle, *β*-alanine (β) by a diamond, and *N,N*-dimethylaminopropylamine (Dp) by a half circle with a plus. The ring pairings prefer binding to different base pairs: Im/Py for G·C, Py/Im for C·G, Py/Py for A·T or T·A. The turn residue (i.e., γ) and the C-terminal β, sometimes referred to as the tail, prefer binding over A·T or T·A base pairs.

an NMR structure of a 2:1 distamycin:DNA complex,[47] which suggested Im could stack across from a Py generating an Im/Py pair that specifically bound to G·C base pairs in the context of a noncovalent ImPyPy dimer.

In the early 1990s it was established that the 2:1 distamycin:DNA complex could be enforced by covalently attaching the antiparallel Im and Py carboxamides through a γ-aminobutyric acid linker to generate a hairpin polyamide such as compound **3** (Figure 1.8).[48] In this context, a set of pairing rules were determined, in which an Im/Py pair binds to a G·C, a Py/Im binds to a C·G, and Py/Py binds to both A·T and T·A (i.e., W). The W degeneracy of the Py/Py pairing was broken with the *N*-methyl-3-hydroxypyrrole (Hp) heterocycle, in which a Py/Hp pairing binds to A·T and Hp/Py to T·A (Figure 1.9).[49-50] For example, the ImImPyPy-γ-ImPyPyPy-β-Dp sequence of compound **3** has been programmed to target 5′-WGGWCW-3′.[51-52] The programmable DNA binding properties of polyamides have been thoroughly studied by a variety of physical techniques, including DNase and MPE footprinting,[53] x-ray crystallography[50,54-55] and multidimensional NMR,[56-57] DNA microarrays,[58] and fluorescence spectroscopy.[59-61] Recently a 27-compound

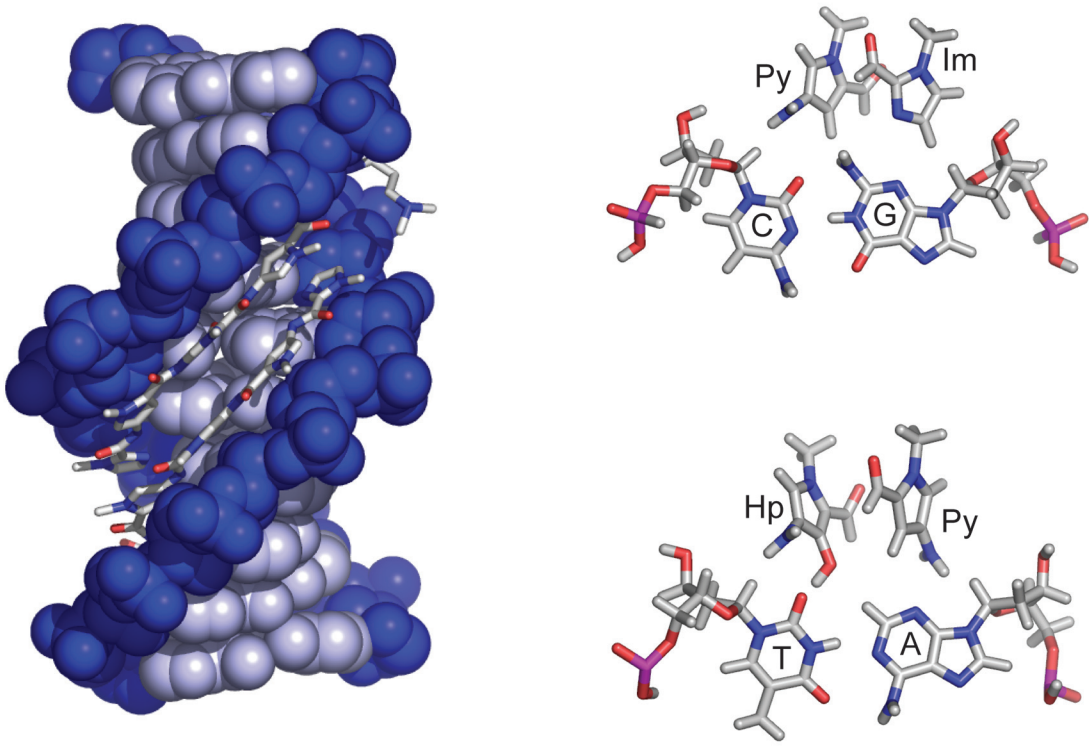

Figure 1.9 A structure illustrating the DNA recognition by a synthetic polyamide. Similar to the distamycin homodimer, the polyamide binds to DNA with the stacked heterocycles making contacts with nitrogenous base pairs in the minor groove (PDB code 1CVX). On the right, the structural basis of the pairing rules is shown.

polyamide library that can target all sequences of the type 5′-WWGNNNW-3′ has been completed.[62] Polyamides can bind to DNA as it exists in cell nuclei on the nucleosome core particle[63-64] and fluorescently labeled derivatives often localize to the nucleus of living cells.[65-67] Certain polyamides have shown activity in cell culture experiments. Notable recent examples inhibit HIF1-α[68] and androgen receptor[69] DNA interactions in the promoter regions of target genes and down-regulate expression.

1.6 Synthetic Mimics of HOX Transcription Factors

Small-molecule synthetic mimics of HOX transcription factors have been created by combining the technology of polyamides with an analysis of HOX proteins.[70] The initial idea was to mimic Ubx as observed in the Ubx/Exd/DNA ternary complex structure.[30] Polyamide-peptide conjugate **4** was designed to bind the DNA minor groove at 5′-TGGTCA-3′ and project an FYPWMKG peptide towards Exd bound to the adjacent major groove at 5′-ATCA-3′ (Figure 1.10). Electrophoretic mobility shift assays showed that conjugate **4** enhanced the binding affinity of Exd to DNA by approximately three orders of magnitude at the composite site of 5′-TGGTCAATCA-3′ but not mismatch sites. Subsequently, unbiased DNA microarray experiments demonstrated that the intended cognate DNA site enabled complex formation better than other sequences.[58] Thus, this conjugate serves as a protein-DNA dimerizer that mimics Ubx and other HOX proteins by binding cooperatively to specific DNA sites with the cofactor protein Exd.[70] Some compounds with longer linkers have also been shown to be make temperature sensitive protein-DNA dimerizers.[71]

1.7 Artificial Transcription Factors

More generally, the covalent attachment of functional peptides to DNA-binding small molecules has led to the development of compounds called artificial transcription factors.[72-73] An early example of such a compound utilized a triple-helix forming oligonucleotide (TFO) conjugated to a VP16 activation peptide that increased luciferase gene expression *in vitro*.[74] Polyamides conjugated to GCN4 dimerization and amphipathic helix peptide activation domains also lead to activation of gene expression on a g-less cassette *in vitro*.[75] The peptide domains have been reduced down to as little as 8 and 16 amino acids and still activate gene expression.[76] Experiments using a poly-proline helix as a structurally rigid spacer domain have defined the optimal distance between the polyamide and activation peptides.[77] Only in the past couple of years has a non-peptidic activation domain been conjugated to a polyamide and shown to activate gene expression *in vitro*.[78] A polyamide-

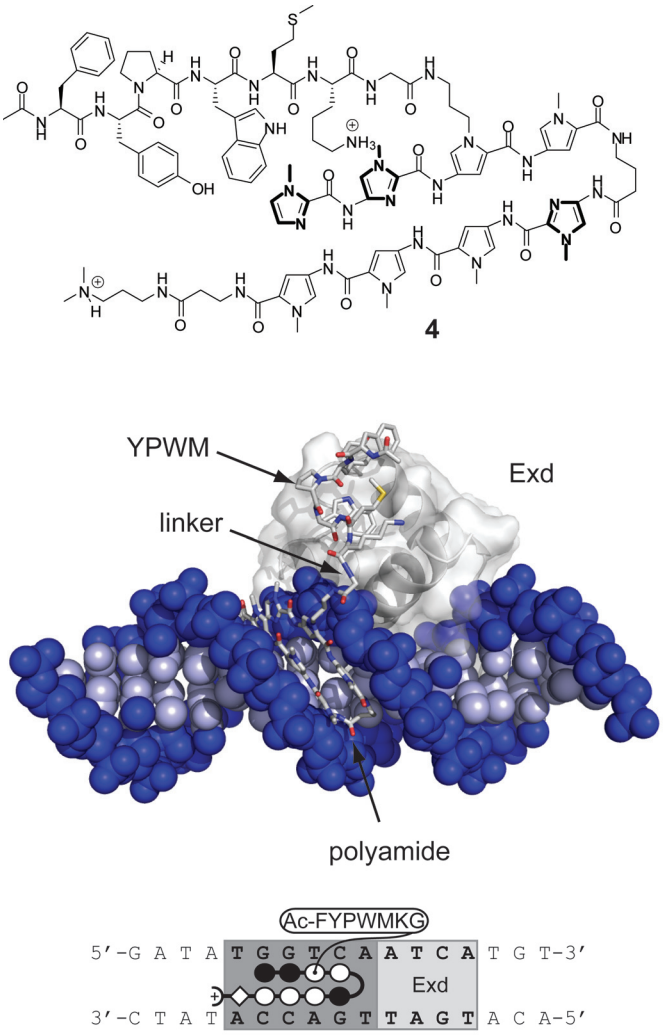

Figure 1.10 The chemical structure and model of a synthetic protein-DNA dimerizer (**4**). The dimerizer consists of a DNA-binding pyrrole-imidazole polyamide conjugated to a heptapeptide containing a YPWM motif that recruits Exd to the adjacent major groove. The illustrative model was constructed from a representative polyamide-DNA structure (PDB code 365D) and the Ubx/Exd/DNA structure (PDB code 1B8I).

peptoid activation domain has very recently been discovered to activate gene expression of a transiently transfected luciferase gene in cell culture,[79] but a cell-permeable, artificial transcription factor that activates an endogenous gene has yet to be reported.

1.8 Scope of this Work

This thesis describes efforts toward determining the essential components of protein-DNA dimerizer transcription factor mimics that should enable them to function in living cells. Chapter 2 details experiments involving that minimization of protein-DNA dimerizer **4** paying particular attention to the protein-binding domain. This led to the discovery that the orginial FYPWMKG heptapeptide can be reduced to a WM dipeptide. Inverting the tryptophan stereochemistry, from the natural L-from to the unnatural D-form, made the dimerizer less sensitive to temperature increases. Chapter 3 shows that the shape of protein-DNA dimerizers can be changed from branched to linear which expands the range of DNA sites and orientations for dimerization. The optimal DNA site spacings and necessary protein-binding elements were determined, which lead to the discovery that a WMK tripeptide works as a protein-binding domain in this context. In chapter 4 fluorescent, cell-permeable protein-DNA dimerizers are presented that incorporate the minimization efforts from the previous chapters.

Chapter 5 details the progress towards the development of a general, cell-permeable artificial transcription factor that mimics the natural HIF-1α protein. In the native system, endogenous HIF-1α binds to the hypoxia response element (HRE) of the VEGF promoter in response to hypoxia.[80] HIF-1α then recruits the enormous co-activator CREB binding protein (CBP) which then activates gene expression. The design strategy for a small molecule HIF-1α mimic employs a polyamide known to bind the VEGF HRE conjugated to compounds that bind CBP. Thus, an effort was made to artificially recruit CBP to the VEGF promoter and activate gene expression in cell culture experiments.

1.9 References

(1) Lander, E. S. et al. *Nature* **2001**, *409*, 860-921.

(2) Venter, J. C. et al. *Science* **2001**, *291*, 1304-1351.

(3) Lee, T. I. and Young, R. A. *Annu. Rev. Genet.* **2000**, *34*, 77-137.

(4) Kadonaga, J. T. *Cell* **2004**, *116*, 247-257.

(5) Levine, M. and Tjian, R. *Nature* **2003**, *424*, 147-151.

(6) Wolberger, C. *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 29-56.

(7) Darnell, J. E. *Nat. Rev. Cancer* **2002**, *2*, 740-749.

(8) Matson, G. A., Evans, S. K., Green, M. R. *Annu. Rev. Genomics Hum. Genet.* **2006**, *7*, 29-59.

(9) Näär, A. M., Lemon, B. D., Tjian, R. *Annu. Rev. Biochem.* **2001**, *70*, 475-501.

(10) Bourbon, H-M. et al. *Mol. Cell* **2004**, *14*, 553-557.

(11) Williams, S. K. and Tyler, J. K. *Curr. Opin. Genet. Develop.* **2007**, *17*, 88-93.

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

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

(14) Kim, J. L., Burley, S. K. *Nat. Struct. Biol.* **1994**, *1*, 638-653.

(15) Juo, Z. S., Chiu, T. K., Leiberman, P. M., Baikalov, I., Berk, A. J., Dickerson, R. E. *J. Mol. Biol.* **1996**, *261*, 239-254.

(16) Pavletich, N. P., Pabo, C. O. *Science* **1991**, *252*, 809-817.

(17) Ellenberger, T. E., Brandl, C. J., Struhl, K., Harrison, S. C. *Cell* **1992**, *71*, 1223-1237.

(18) Glover, J. N. M., Harrison, S. C. *Nature* **1995**, *373*, 257-261.

(19) Panne, D., Maniatis, T., Harrison, S.C. *Cell* **2007**, *129*, 1111-1123.

(20) Lewis, E. B. *Nature* **1978**, *276*, 565-570.

(21) Veraksa, A., Del Campo, M., McGinnis, W. *Mol. Genet. Metab.* **2000**, *69*, 85-100.

(22) Pearson, J. C., Lemons, D., McGinnis, W. *Nat. Rev. Genet.* **2005**, *6*, 893-904.

(23) Abate-Shen, C. *Nat. Rev. Cancer* **2002**, *2*, 777-785.

(24) Chang, C., Immaculata, D., Clearly, M. L. *Mol. Cell. Biol.* **1997**, *17*, 81-88.

(25) Owens, B. M., Hawley, R. G. *Stem Cells* **2002**, *20*, 364-379.

(26) Mortlock, D. P., Innis, J. W. *Nat. Genet.* **1997**, *15*, 179-180.

(27) Mann, R. S., Chan, S. K. *Trends Genet.* **1996**, *12*, 258-262.

(28) Mann, R. S., Affolter, M. *Curr. Opin. Genet. Dev.* **1998**, *8*, 423-429.

(29) Ryoo, H. D., Marty, T. Casares, F., Affolter, M., Mann, R. S. *Development* **1999**, *126*, 5137-5148.

(30) Passner, J. M., Ryoo, H. D., Shen, L. Y., Mann, R. S., Aggarwal, A. K. *Nature* **1999**, *397*, 714-719.

(31) Piper, D. E., Batchelor, A. H., Chang, C. P., Clearly, M. L., Wolberger, C. *Cell* **1999**, *96*, 587-597.

(32) LaRonde-LeBlanc, N. A.; Wolberger, C. *Genes Dev.* **2003**, *17*, 2060-2072.

(33) Duboule, D. editor in *Guidebook to the Homeobox Genes*; Oxford University Press: Oxford, 1994.

(34) Peifer, M., Wieschaus, E. *Genes. Dev.* **1990**, *4*, 1209-1223.

(35) Kamps, M. P., Murre, C., Sun, X., Baltimore, D. *Cell* **1990**, *60*, 547-555.

(36) Pui, C., Relling, M. V., Downing, J. R. *N. Eng. J. Med.* **2004**, *350*, 1535-1548.

(37) Lu, Q., Wright, D. D., Kamps, M. P. *Mol. Cell. Biol.* **1994**, *14*, 3938-3948.

(38) Knoepfler, P. S., Kamps, M. P. *Oncogene* **1997**, *14*, 2521-2531.

(39) McWhirter, J. R., Neuteboom, S. T. C., Wancewicz, E. V., Monia, B. P., Downing, J. R., Murre, C. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11464-11469.

(40) Fu, X., McGrath, S., Pasillas, M., Nakazawa, S., Kamps, M. P. *Oncogene* **1999**, *18*, 4920-4929.

(41) Casagrande, G., Kronnie, G., Basso, G. *Haematologica/The Hematology Journal* **2006**, *91*, 765-771.

(42) Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., Dickerson, R. E. *J. Mol. Biol.* **1985**, *183*, 553-563.

(43) Coll, M., Frederick, C. A., Wang, A. H., Rich, A. *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 8385-8389.

(44) Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., Dickerson, R. E. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1376.

(45) Lown, J. W., Krowicki, K., Bhat, U. G., Skorobogaty, A., Ward, B., Dabrowiak, J. C. *Biochemistry* **1986**, *25*, 7408.

(46) Wade, W. S., Mrksich, M., Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783-8794.

(47) Pelton, J. G., Wemmer, D. E. *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5723-5727.

(48) Mrksich, M., Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572.

(49) White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B. *Nature* **1998**, *391*, 468.

(50) Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B. *Science* **1998**, *282*, 111.

(51) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215-2235.

(52) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299.

(53) Trauger, J. W., Dervan, P. B. *Methods Enzymol.* **2001**, *340*, 450-466.

(54) Kielkopf, C. L., Baird, E. E., Dervan, P. B., Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104-109.

(55) Kielkopf, C. L., Bremer, R. E., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. D., Rees, D. C. *J. Mol. Biol.* **2000**, *295*, 557-567.

(56) Urbach, A. R., Love, J. J., Ross, S. A., Dervan, P. B. *J. Mol. Biol.* **2002**, *320*, 55-71.

(57) Zhang, Q., Dwyer, T. J., Tsui, V., Case, D. A., Cho, J., Dervan, P. B., Wemmer, D. E. *J. Am. Chem. Soc.* **2004**, *126*, 7958-7966.

(58) Warren, C. L., Kartochvil, N. C. S., Hauschild, K. E., Foister, S., Brezinski, M. L., Dervan, P. B., Phillips, G. N., Ansari, A. Z. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 867-872.

(59) Rucker, V. C., Foister, S., Melander, C., Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*,

1195-1202.

(60) Fechter, E. J., Olenyuk, B., Dervan, P. B. *J. Am. Chem. Soc.* **2005**, *127*, 16685-16691.

(61) Chenoweth, D. M., Viger, A., Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2216-2217.

(62) Hsu, C. F., Phillips, J. W., Trauger, J. W., Farkas, M. E., Belitsky, J. M., Heckel, A., Olenyuk, B. Z., Puckett, J. W., Wang, C. C. C. Dervan, P. B. *Tetrahedron* **2007**, *63*, 61416-6151.

(63) Gottesfeld, J. M., Melander, C., Suto, R. K., Raviol, H., Luger, K., Dervan, P. B. *J. Mol. Biol.* **2001**, *309*, 615-629.

(64) Suto, R. K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., Luger, K. *J. Mol. Biol.* **2003**, *326*, 371-380.

(65) Belistky, J. M., Leslie, S. J., Arora, P. S., Beerman, T. A., Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 3313-3318.

(66) Best, T. P., Edelson, B. S., Nickols, N. G., Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12063-12068.

(67) Edelson, B. S., Best, T. P., Olenyuk, B., Nickols, N. G., Doss, R. M., Foister, S., Heckel, A., Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802-2818.

(68) Olenyuk, B. Z., Zhang, G., Klco, J. M., Nickols, N. G., Kaelin, W. G., Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16768-16773.

(69) Nickols, N. G., Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10418-10423.

(70) Arndt, H.-D., Hauschild, K. E., Sullivan, D. P., Lake, K., Dervan, P. B., Ansari, A. Z. *J. Am. Chem. Soc.* **2003**, *125*, 13322-13323.

(71) Hauschild, K. E., Metzler, R. E., Arndt, H.-D., Moretti, R., Raffaelle, R., Dervan, P. B., Ansari, A. Z. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5008-5013.

(72) Ansari, A. Z. and Mapp, A. K. *Curr. Opin. Chem. Biol.* **2002**, *6*, 765-772.

(73) Arndt, H.-D. *Angew. Chem., Int. Ed.* **2006**, *45*, 4552-4560.

(74) Kuznetsova, S., Ait-Si-Ait, S., Nagibneva, I., Troalen, F., Le Villain, J., Harel-Bellan, A., Svinarchuk, F. *Nucleic Acids Res.* **1999**, *27*, 3995-4000.

(75) Mapp, A. K., Ansari, A. Z., Ptashne, M., Dervan, P. B. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3930-3935.

(76) Ansari, A. Z., Mapp, A. K., Nguyen, D. H., Dervan, P. B., Ptashne, M. *Chem. Biol.* **2001**, *8*, 583-592.

(77) Arora, P. S., Ansari, A. Z., Best, T. P., Ptashne, M., Dervan, P. B. *J. Am. Chem. Soc.* **2002**, *124*, 13067-13071.

(78) Kwon, Y., Arndt, H.-D., Mao, Q., Choi, Y., Kawazoe, Y., Dervan, P. B., Uesugi, M. *J. Am. Chem. Soc.* **2004**, *126*, 15940-15941.

(79) Xiao, X., Yu, P., Lim, H.-S., Sikder, D., Kodadek, T. *Angew. Chem. Int. Ed.* **2007**, *46*, 2865-2868.

(80) Kaelin, W. G. *Genes Dev.* **2002**, *16*, 1441-1445.