Chapter 1

Introduction

The text of this chapter regarding polyamide uptake was taken in part from a manuscript coauthored with Timothy P. Best, Benjamin S. Edelson, and Peter B. Dervan (Caltech).

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

1.1 Rationale

It is estimated that the human genome contains some 20,000 to 25,000 genes (1, 2). Information encoding the structure, regulation, and expression of these genes and their expressed products is encoded in the nucleotide sequences of the base pairs of the DNA molecules themselves. In general, expression of genes involves transcription of sequences of DNA into sequences of RNA that are translated to direct the synthesis of a polypeptide chain. The products of genes catalyze chemical reactions, interact with other cellular proteins and the environment, and sense and react to environmental cues within and beyond the cell. The activation of specific gene programs in concert or opposition with others is what drives the development of a zygote to a person and from normal tissue to malignancy. With the exception of the genetic material itself, most molecules that make up a living cell are fleeting and must be renewed. Therefore, any lasting response to an environmental or internal cue must be initiated through changes in expressed genes. By corollary, disease states are initiated and maintained through aberrations in gene expression. Technologies that enable direct control over the expression of target genes in a predictable manner could have profound applications in human medicine.

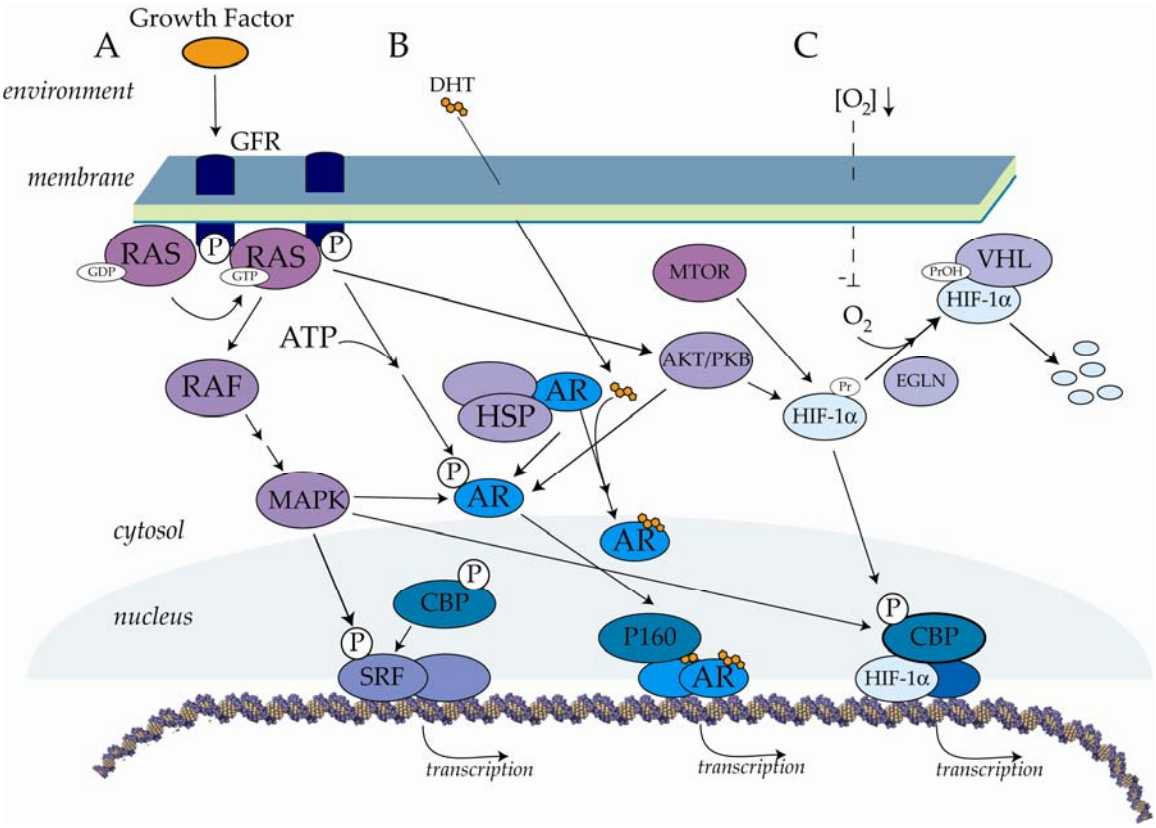

Figure 1.1 Activation of three representative signaling pathways and their down-stream transcription factor targets. (A) A growth factor binding to a membrane receptor induces auto-phosphorylation and activates the G-protein RAS, which activates RAF and mitogen activated protein kinase (MAPK), ultimately phorphorylating and activating resident nuclear transcription activators including serum response factor, resulting in the expression of genes involved in cell proliferation. (B) A steroid hormone such as dihydrotestosterone (DHT) diffuses into the cell to bind androgen receptor, which is released from cytoplasmic chaperones into the nucleus to activate transcription of target genes. (C) Changes in O2 concentration are sensed by an O_2 dependent proline hydroxylase (EGLN) that targets HIF-1 for degradation. Low O_2 levels result in HIF-1α accumulation and activation of target genes. Significant cross-talk exists between signaling pathways. For example: growth factor receptors can also activate AR independent of DHT; and protein kinase B (PKB/AKT) can activate HIF-1α and AR. The phosphorylation state of transcriptional co-activators such as CBP, which can be recruited by SRF, HIF-1, and AR, is influenced by several protein kinases.

1.2 From environmental stimuli to programs of gene expression

As information passes from environment to cell nucleus it is transduced in a variety of forms: changes in the concentration of ions or second messenger molecules, the

presence or absence of a phosphate group on particular proteins, the association or disassociation of non-covalent interactions between proteins or between proteins and other molecules, to name a few (3). The interlacing networks of proteins in cells that receive and channel information is complex; input signals amplify, diverge, and converge, and significant cross-talk between multiple cascades of information flow contribute to a final integration of stimuli to evoke a response by modulating the expression of sets of genes (3). These responses are enacted by transcription activators and repressors. These proteins bind DNA sequence specifically, and direct, or prevent direction of, transcriptional machinery to the promoters of target genes (3). The activator/repressor-DNA interface is the point where information from protein signaling is converted into programs of gene expression.

Activators and repressors are regulated by upstream signals through diverse mechanisms. For example, activation can result from an allosteric structural change as the result of phosphorylation, from ligand binding, or from a dynamic change in the stability of the activator. As an example, Figure 1.A-C summarizes three signaling pathways each activated by a different environmental stimulus: growth-factor binding to a cell surface receptor, hormone binding to an intracellular receptor, and changes in the concentrations of O_2. Binding of growth factor ligands to cell surface receptors results in the activation of a protein kinase cascade that ultimately phosphorylates the DNA-bound serum response factor (SRF), and other transcription factors, inducing a conformational change that results in gene activation (4). The androgen receptor (AR) is sequestered in the cytoplasm by heat shock proteins (5). Binding of the hormone dihydrotestosterone to AR results in release from the cytoplasm and binding to enhancer sites in the genome.

AR activity is also, in part, influenced by phosphorylation by kinases that are activated by different stimuli. Under conditions of normal O_2 concentration, the α unit of the hypoxia inducible factor (HIF-1α) is targeted for degradation (6, 7). Under hypoxia, HIF-1α accumulates and affects expression of a set of genes involved in cell and tissue adaptation to hypoxia (6).

Many oncogenes, which constitute the driving force in malignancy, are involved in cell signaling pathways (8). Because numerous signaling pathways converge on a smaller number of transcription factors to exert their effects on gene expression, it has been proposed that transcription factors could be among the most appropriate drug targets in oncology (8). Molecules that bind DNA sequence specifically offer one approach toward modulating the activity of specific transcription factors (9).

1.3 Mechanisms of Gene Expression

Activators may bind nearby or even several thousand base pairs up- or down-stream from the promoter region of a target gene, or even on a different chromosome, and can access DNA sites that are associated with histones (10-12). Gene activation by these factors proceeds through a number of mechanisms (Figure 1.2). Activators, or proteins that they recruit, can exert a direct effect on the local nucleosome architecture. Co-activators such as Creb Binding Protein (CBP) that do not bind DNA directly, but are recruited through interaction with a DNA binding activator, have histone acetyl-transferase activity. The acetylation of histones promotes a loose chromatin architecture that is conducive to additional protein recruitment and transcription (13, 14). Transcriptional co-repressors, conversely, can recruit histone deacetylase complexes that

promote a tighter chromatin architecture. In addition, activators bind and recruit mediator proteins, which in turn recruit the general transcriptional machinery (10, 15). In cases where the activators are bound to distal enhancers, the physical association of the activator-mediator-transcriptional machinery brings the enhancer region in close proximity to the promoter when the gene is actively being transcribed. The general transcriptional machinery of protein-coding genes is composed of several proteins including Tata binding protein (TBP) and TFIIA, B, D, E, F, and H, and RNA polymerase II (10). A generalized model for the expression of a protein coding gene is depicted in figure 1.2.

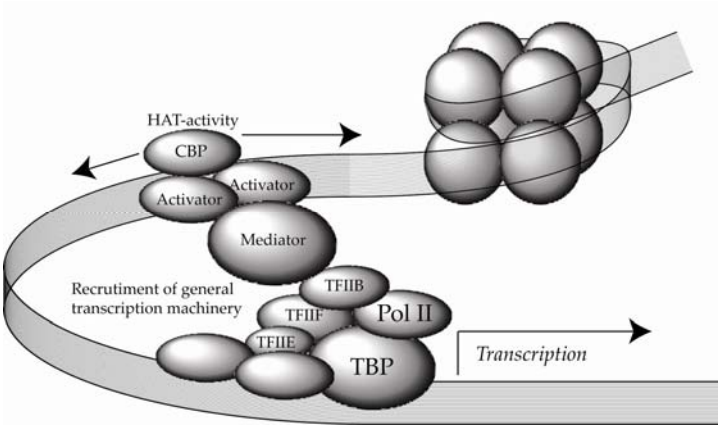

Figure 1.2 Depiction of an active transcription complex. An activator bound to an enhancer recruits co-activators that affect histone acetylation, relaxing nucleosome associations. Activators contact other co-activators and mediator proteins that in turn contact and recruit the general transcriptional machinery to the promoter. The promoter and enhancer regions are often located proximal through space but separated by a variable number of base pairs.

1.4 DNA binding proteins and small molecules

DNA consists of two complimentary, antiparallel polydeoxyribonucleotide strands intertwined as a double helix (16). The strands are held together by specific hydrogen bonds between nucleotide bases: adenine pairs with thymine, and guanine pairs

with cytosine (16). In the biologically relevant B-form, the sugar phosphate backbones of each strand define two grooves: a major groove that is broad and shallow, and a minor groove that is narrow and deep (16). The differences in molecular surfaces defined by the edges of each base pair at the floors of the major and minor grooves form the basis for sequence-specific binding by proteins and other molecules.

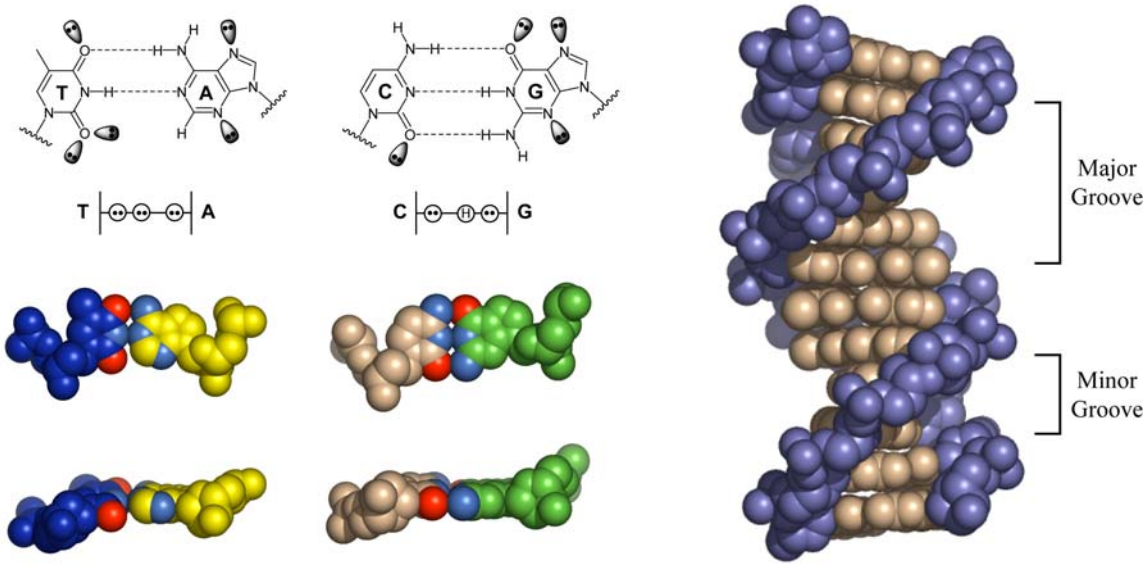

Figure 1.3 The structure of DNA. The structure of T•A and C•G base pairs. B-form double helical DNA noting the major and minor grooves.

Transcription factors utilize a diverse array of structural motifs to bind DNA (figure 1.4). Sequence specific binding is achieved through combinations of electrostatic interactions with the negatively charged phosphate backbone and van der Waals interactions with the bases at the floor of either groove (17-22). A majority of DNA binding proteins interact with DNA predominantly in the major groove, though many bind the minor groove, or both. In many cases, DNA binding proteins bind as homo- or hetero-dimers to their target sites. The zinc-finger motif is composed of 30 amino acid domains that are arranged as two antiparallel β strands and an α helix stabilized by a zinc

ion complexed by two histidines and two cysteines. Due to the modular structures of the zinc-finger motif, it has been possible to design polydactyl zinc finger proteins that bind predetermined DNA sequences (23).

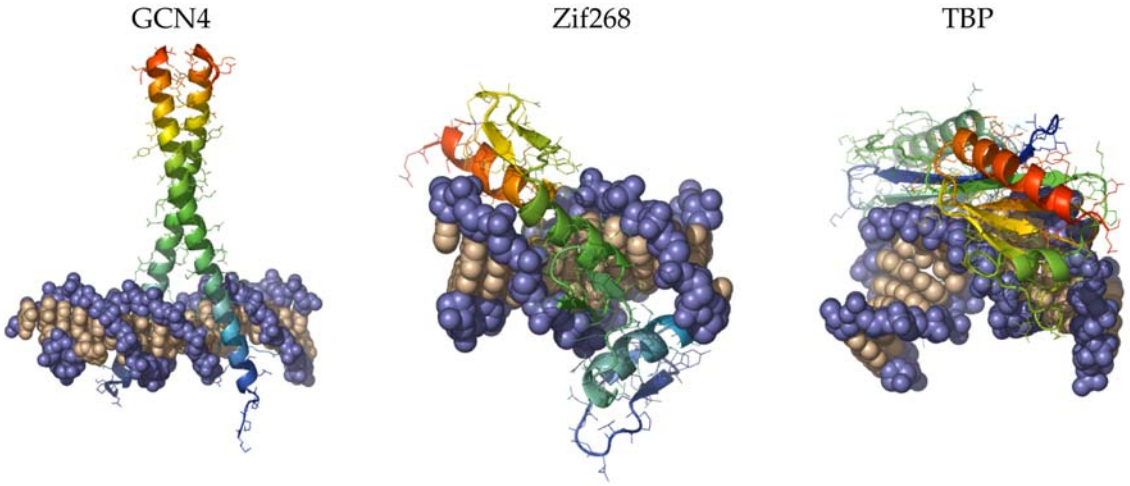

Figure 1.4 Crystal structures of three DNA binding proteins. GCN4 binds as a homodimer to the major groove. Zif268 contains three zinc finger modules and interacts primarily in the major groove. TBP binds the minor groove predominantly and induces a severe bend into the DNA structure.

Aside from proteins, a number of natural products exist that bind DNA with some degree of sequence specificity (figure 1.5). Among these molecules are calicheamicin, Hoechst 3258, and distamycin (24-27). The crescent-shaped natural product distamycin binds A•T tracts of DNA in the minor groove in either a 1:1 or 2:1 stoichiometry. The relatively small size, simple structure, and modularity of the repeating pyrrole units make distamycin an attractive lead molecule for the development of programmable DNA binding ligands.

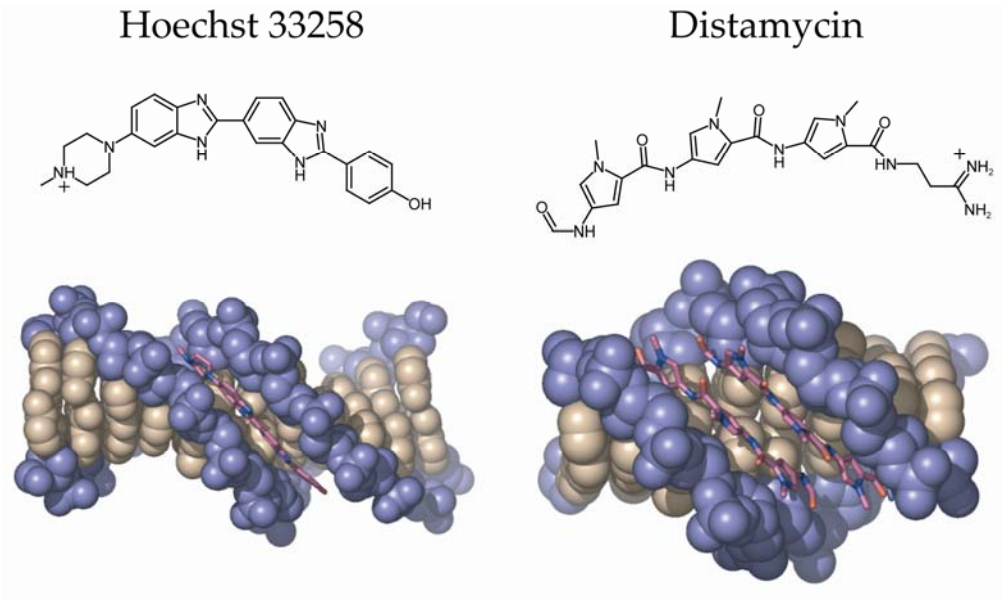

Figure 1.5 Crystal structures of two DNA binding natural products. Hoechst 33258 binds A•T and T•A sequences. Distamycin also recognizes A•T and T•A tracts. It is shown here in a 2:1 binding stoichiometry.

1.5 Sequence-specific recognition of DNA by polyamides

It was hypothesized that replacement of the first N-methylpyrrole of distamycin with an N-methylimidazole would impart different sequence specificity, through accommodation of the exocyclic amine of the guanine base with the nitrogen of imidazole. Footprinting experiments revealed that this polyamide bound the sequence 5'-WGWCW-3' (28). Structural studies using NMR revealed that this sequence specificity results from a 2:1 binding mode in which the three-ring polyamide bound in the minor groove as an anti-parallel dimmer (28). Co-facial pairings of Im and Py could distinguish G•C from C•G.

The distamycin scaffold has inspired development of a class of synthetic, programmable DNA binding molecules that bind the minor groove of DNA with high affinity and specificity. These molecules are oligomers of linked N-methylpyrrole (Py), N-methylimidazole (Im), N-methyl-3-hydroxypyrrole (Hp), benzimidazole (Bi),

hydroxybenzimidazole (Hz), and 3-chlorothiophene (Ct) carboxamides (9, 29-35). Similar to distamycin, these polyamides can bind DNA in both 1:1 and 2:1 stoichiometries. Each polyamide strand orients N→C along the 5'→3' direction of the DNA helix in most cases. The crescent shape of the ring pairs fits the overall shape of the DNA minor groove with regard to width, curvature, and depth. The individual ring pairs recognize subtle differences in molecular shape and hydrogen bond donor and acceptor patterns along the edges of the Watson-Crick nucleotide bases pairs at the floor of the minor groove. Polyamide-DNA association is driven by a combination of van der Waals interactions and specific hydrogen bonds.

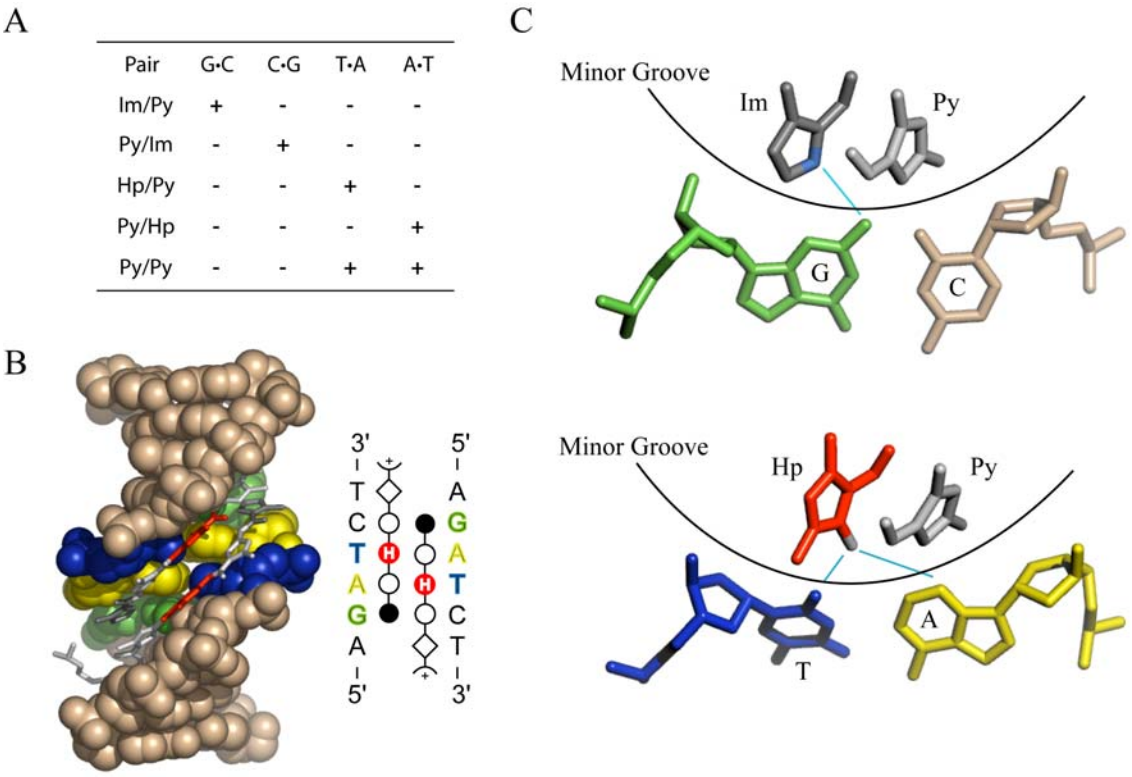

Figure 1.6 (A) Pairing rules for minor groove recognition by polyamides. (B) Crystal structure of a polyamide homodimer bound to a 5'-GTAC-3' core sequence. Closed circles represent Im, open circles represent Py, red circles containing an "H" represent Hp, and diamonds represent β-alanine. (C) Recognition of G•C and T•A by Im/Py and Hp/Py.

Sequence specificity is programmed by side-by-side pairings of the heterocyclic amino acids in the minor groove of DNA: Im/Py distinguishes G•C from C•G; Py/Py binds both A•T and T•A; Hp/Py and Hz/Py both distinguish T•A from A•T; and Ct/Py prefers T•A over A•T at the amino terminus position (Figure 1.6, 1.7). These pairing guidelines have been identified based on NMR, crystallographic, and fluorescence studies, as well as hundreds of quantitative DNase I footprinting titrations. A number of structural studies have investigated the nature of the sequence-specific binding nature of the polyamide ring pairs. Crystal structures of polyamide-DNA complexes have revealed the presence of a hydrogen bond between the nitrogen of imidazole and the exocyclic amine of guanine, as well as between the amide bonds of the polyamide backbone and the edges of the base pairs (36-38). Recognition of T•A by Hp/Py is thought to be driven by a hydrogen bond between the hydroxyl and the thymine-O2 and shape selective recognition of an asymmetric cleft between the thymine-O2 and adenine-C2 (37, 38). Recognition of T•A at the amino terminus by Ct/Py is thought to depend on projection of the 3-chloro substituent into the minor groove floor (39).

Covalent linkage of antiparallel polyamide strands with a flexible alkyl linker was hypothesized to reduce the entropic cost of polyamide-DNA binding. Linking two polyamide strands N- to C- with such an aliphatic linker results in a hairpin-shaped polyamide with increased affinity (>100-fold) and specificity as compared to the corresponding unlinked dimer (40). This “turn” unit also enforces side-by-side pairing and prevents the slipped binding modes possible for unlinked dimers (40, 41). The most well-studied hairpin polyamides consist of a total of eight heterocyclic rings linked by γ-amino butyric acid, with or without a chiral amine moiety at the α carbon (42). These

polyamides bind a DNA sequence six base pairs in length, typically with an affinity and specificity that is comparable to those of natural DNA binding proteins (43). The turn residue itself exhibits strong selectivity for T•A and A•T pairs over G•C and C•G pairs (40).

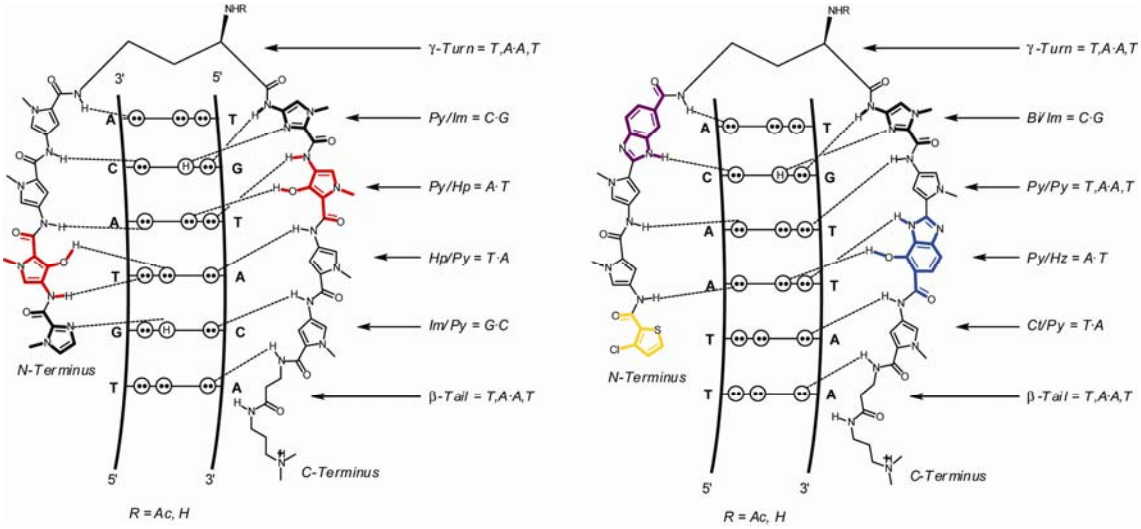

Figure 1.7 Structure of a hairpin polyamide targeted to 5'-WGTACW-3' consisting of Py, Im, and Hp with putative hydrogen bonds (left) and to 5'-WTAACW-3' (right) incorporating benzimidazole 6:5 rings and a Ct N-terminus. A circle with two dots represents a hydrogen bond acceptor; a circle with an "H" represents a hydrogen bond donor.

Other structural motifs that use a variety of covalent linkages between polyamide strands have also been examined (Figure 1.8). Covalent linkage of the N- and C- termini of a hairpin polyamide results in a cyclic polyamide that typically exhibits increased affinity, albeit with slight decreased specificity, over the corresponding hairpin (44). Two hairpins linked in tandem result in a molecule with high affinity and greatly increased binding site size, but at the cost of single base pair mismatch specificity (45, 46). Polyamide strands can also be linked with an aliphatic chain via the ring nitrogens. The U-pin motif, linked at the terminal ring nitrogens, relaxes the requirement for a T•A or A•T base pair at the turn position (47). The H-pin motif, which links the polyamide

strands through internal ring nitrogens, exhibits affinity comparable to an analogous hairpin polyamide (48).

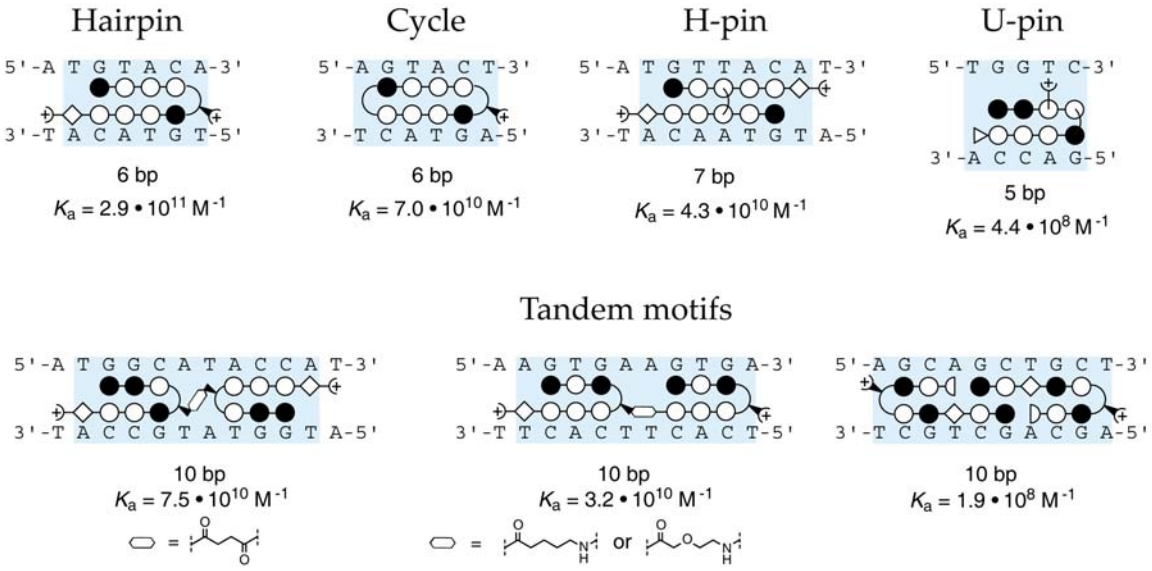

Figure 1.8 Ball-and-stick models and representative K_a values for various polyamide structural motifs.

Due to their high affinity, polyamides can often competitively displace DNA binding proteins when targeted to the same or adjacent sites (Figure 1.9) (31). Numerous *in vitro* studies have shown successful inhibition of the binding of proteins to target sites on DNA that include many of the general transcription factors and transcriptional activators through steric and allosteric mechanisms (31). Some exclusively major-groove binding proteins, such as the basic-region leucine zipper (bzip) GCN4, have been shown to co-occupy sequences of DNA with minor groove binding polyamides (49). However, polyamide-acridine conjugates that deliver an intercalator to a DNA site sequence

specifically have been shown to inhibit the binding of proteins that were unaffected by unmodified polyamides (50).

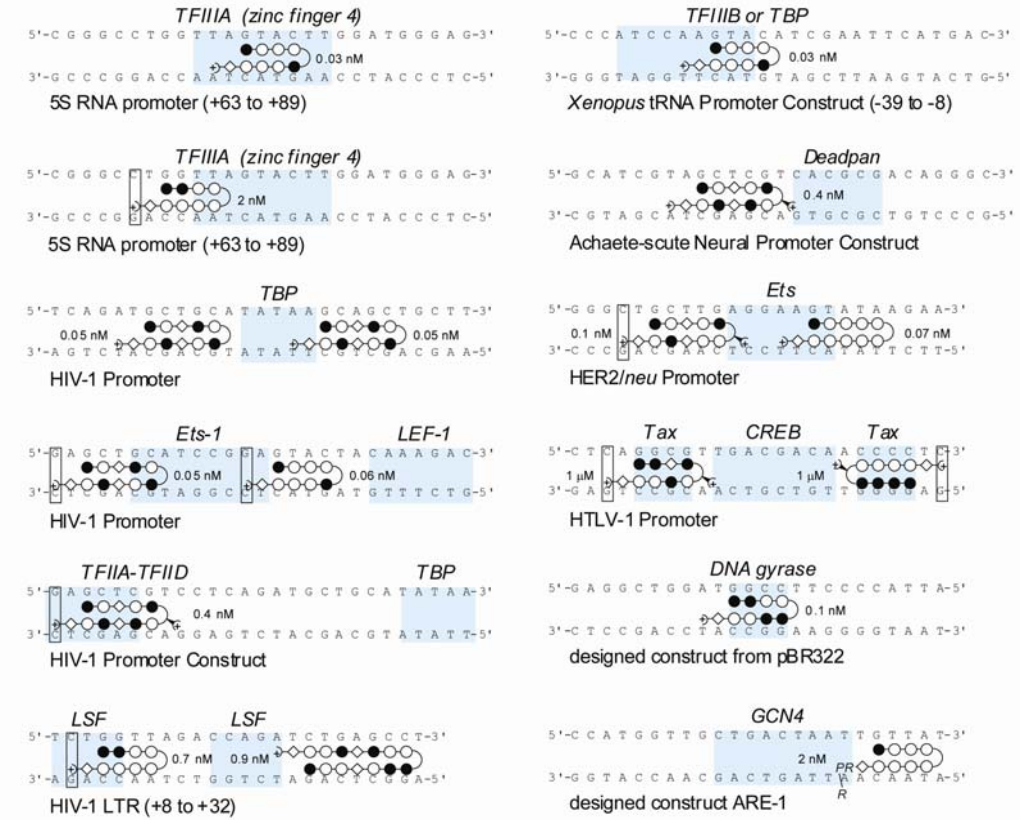

Figure 1.9 Polyamides have been shown to inhibit the binding of numerous DNA binding proteins. Dissociation constants for the polyamide displayed at that site are included.

Much of the DNA in cells is tightly compacted as chromatin in association with histone proteins. The basic repeating unit of chromatin is the nucleosome: a stretch of DNA 147 base pairs in length coiled twice around a protein core of eight histones. The precise structure of the nucleosome and the tightness of the compaction can be regulated by the activity of histone modifying enzymes that are recruited by transcriptional activators and repressors. The ability to target nucleosomal DNA with polyamides has been explored by biochemical and structural studies (51). It has been shown that

polyamides can readily access nucleosomal DNA that is facing away from the protein core. In addition, crystal structures of polyamides bound to a nucleosome core particle indicate that the histone octamer is not affected by polyamide binding (Figure 1.10) (52).

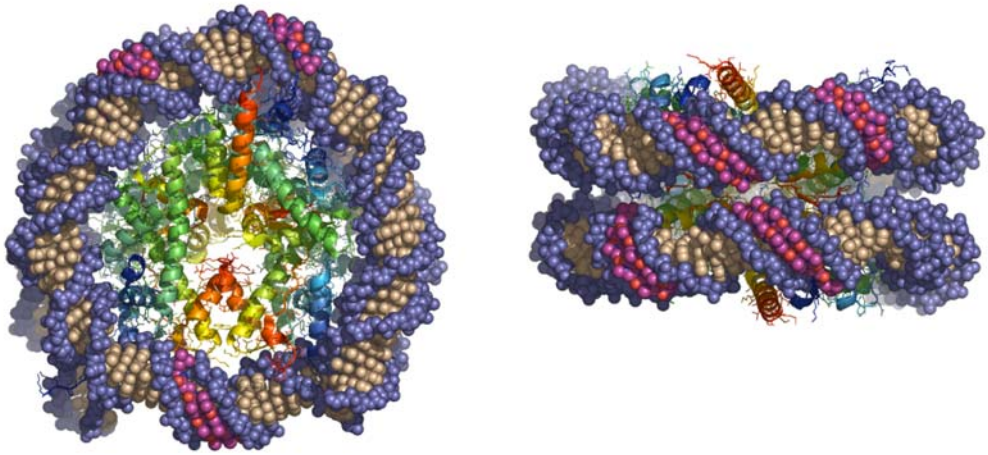

Figure 1.10 Crystal structure of a hairpin polyamide bound to a nucleosome core particle (52).

1.6 Nuclear localization of polyamides

A major goal is the use of polyamides to regulate the expression of endogenous genes by binding to DNA sites in critical gene-regulatory sequences of cells. This requires not only that the polyamides bind to chromosomal DNA site-specifically and with affinity sufficient to displace, inhibit, or recruit endogenous DNA binding proteins, but also that they be permeable to the outer membrane and gain access to the nucleus of living cells.

The cellular uptake and localization of hairpin polyamides has been extensively studied by confocol microscopy using fluorescently labeled polyamides in live cells (53-

55). Experiments involving hundreds of polyamide-dye conjugates and dozens of cell lines reveals that pyrrole–imidazole sequence and content, dye choice and position, linker composition, and molecular weight are determinants of nuclear localization (53-55). Uptake appears to be energy dependent (54). Although a significant number of eight-ring hairpin-polyamides with fluorescein at the C-terminus were shown to have good nuclear uptake in multiple cell lines, clear trends for predicting uptake have not yet emerged. Polyamides of greater than eight-rings and tandem polyamides have exhibited generally poor nuclear uptake, suggesting that molecular weight plays some role in uptake character (54).

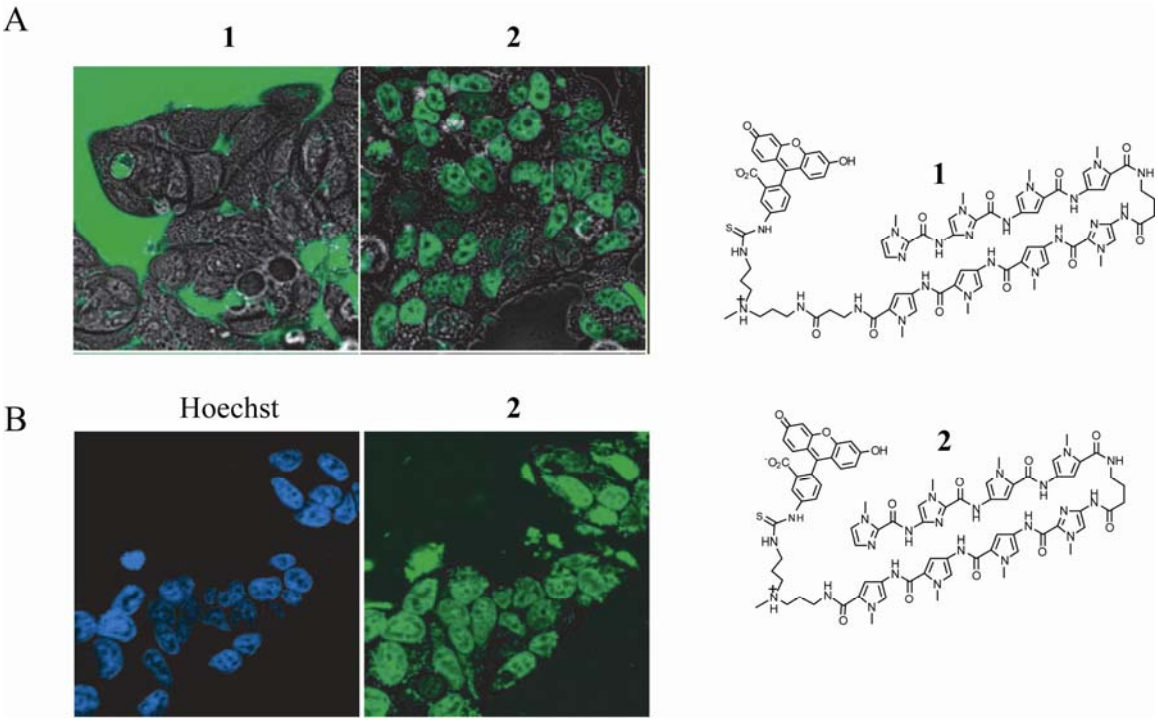

Figure 1.11 Cellular localization of polyamide–fluorescein conjugates **1** and **2** in live MCF7 cells (54). (A) Cells were treated with polyamide-fluorescein conjugate **1** or **2** for 10–14 hours at 5 μM. **1** was excluded from the cells entirely, whereas **2** localized to the nucleus. (B) **2** co-localizes with the DNA-binding dye Hoechst.

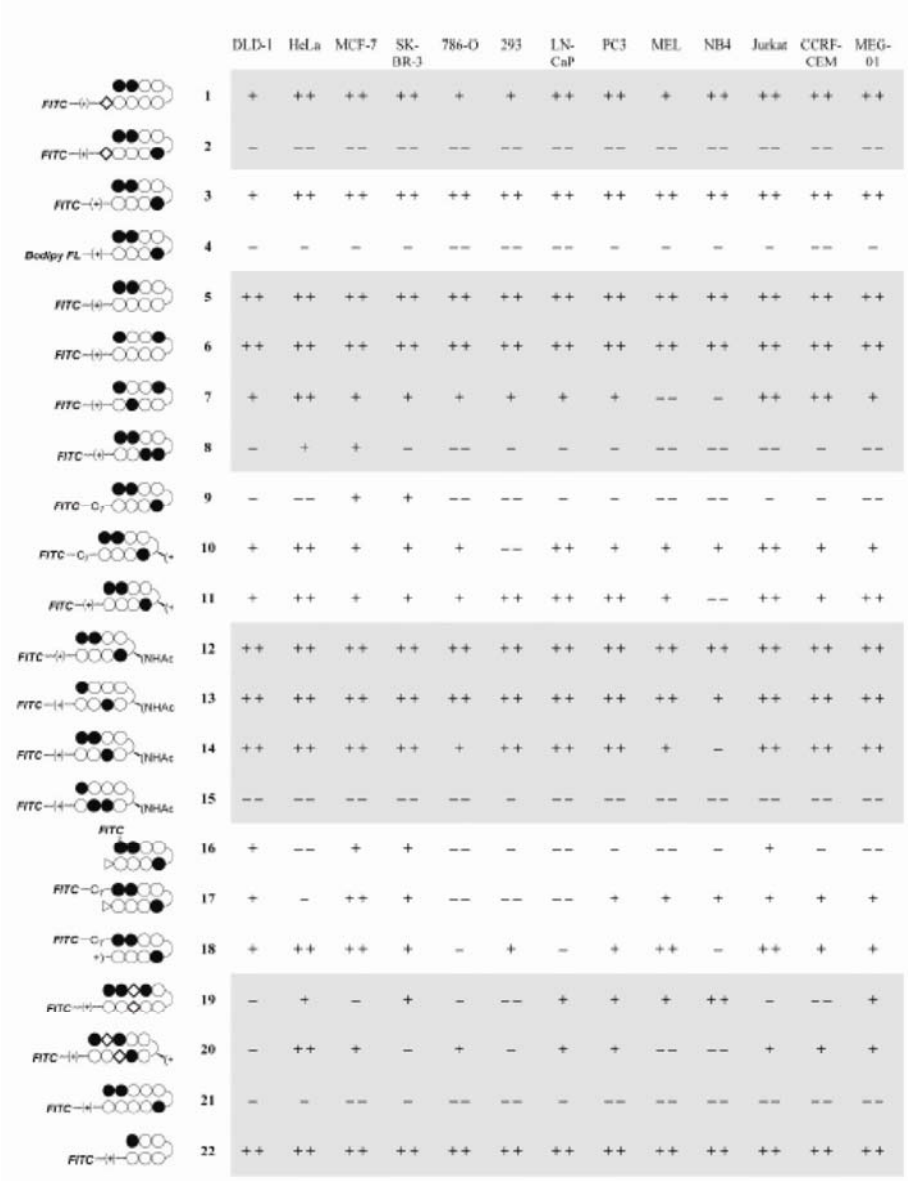

Figure 1.12 Cellular localization of 22 polyamide–dye conjugates in a panel of cultured cells (54). ++, Nuclear staining exceeds that of the medium; +, nuclear staining less than or equal to that of the medium, but still prominent; -, very little nuclear staining, with the most fluorescence seen in the cytoplasm and/or medium; --, no nuclear staining. Py–Im sequence and content, dye choice and position, linker composition, and molecular weight are determinants of nuclear localization

The importance of subtle structural features in determining nuclear uptake of a polyamide is underscored by a comparison of two polyamide-fluorescein conjugates that differ only by the presence or absence of a β-alanine residue at the C-terminus (Figure

1.11) (54). The conjugate containing the β-alanine exhibited no nuclear localization, while the corresponding polyamide lacking the β-alanine exhibited robust nuclear uptake. A full understanding of polyamide uptake and trafficking has not yet been realized. However, hairpin polyamide-dye conjugates that target a broad range of sequences can be shown to enter the nuclei of a large variety of live cells without the use of transfection reagents (Figure 1.12).

1.7 Polyamides in biological systems

The effects of polyamides targeted to sequences of key protein-DNA interactions important for gene transcription *in vitro* has been extensively studied. Transcription of HER2/neu was effectively inhibited in a cell-free system by a polyamide targeted to the Ets-1 site in the promoter sequence (56). Specific polyamides targeted to the terminal inverted repeat sequences of Maloney murine leukemia virus important for DNA integration inhibited viral integration *in vitro* (57). Bacterial gyrase-catalyzed strand cleavage was inhibited *in vitro* by a polyamide targeted to its recognition site (58). In addition to inhibition of transcriptional machinery, conjugates of polyamides to peptide or small molecule activator domains have been shown to activate gene expression in cell free systems (59-62).

In one of the earliest applications of polyamides to a living system, two polyamides were targeted to regions in the HIV promoter adjacent to three transcription factors: TBP, Ets-1, and LEF-1 (63). Although transcription of the HIV genome was not measured directly, these polyamides were shown to inhibit HIV-1 replication in primary human T cells without evidence of obvious cell toxicity (63). Remarkably, the effect on

HIV replication was comparable to that of azidothymidine (AZT). In a different study, polyamides targeted to DNA satellite repeat sequences could mediate chromatin opening at these loci in drosophila cells (64, 65). It was also shown these polyamides, which bind in a 1:1 polyamide:DNA stoichiometry, affected specific gain- and loss-of-function phenotypes in drosophila when fed to developing embryos.

1.8 Summary

Cells integrate stimuli from environmental and internal cues through networks of proteins that affect programs of gene expression. The expression of genes is controlled by transcription activators and repressors. Molecules that can modulate the activator/repressor-DNA interface would have broad applications in human medicine. DNA binding polyamides, developed in large part in the laboratories of Professor Peter Dervan, are currently the only class of programmable small molecule capable of sequence specific DNA binding. In order to affect transcription of endogenous genes by targeting key transcription factor-DNA interactions, polyamides must permeate cells, localize in the nucleus, access chromatin and bind DNA sequences with affinities and specificities sufficient to disrupt protein-DNA binding. Numerous polyamide-dye conjugates localize in the nuclei of cultured cells. *In vitro* studies suggest that, in principle, the binding of polyamides to DNA is not abrogated by the presence of a nucleosome. Multiple studies have shown that polyamides are capable of displacing DNA binding proteins from their target sites in a sequence specific manner. Early experiments in viral systems showed that polyamides can be active in cell culture. The remainder of this thesis describes work aimed at modulating selected transcriptional

pathways in cultured human cells using polyamides that disrupt the binding of endogenous cellular transcriptional activators with a particular focus on hypoxia inducible factor 1 and androgen receptor.

References

1. Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001) *Nature* **409,** 860-921.
2. Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., Mural, R. J., Sutton, G. G., Smith, H. O., Yandell, M., Evans, C. A., Holt, R. A., et al. (2001) *Science* **291,** 1304.
3. Brivanlou, A. H. and Darnell, J. E. (2002) *Science* **295,** 813-818.
4. Wellbrock, C., Karasarides, M. and Marais, R. (2004) *Nat. Rev. Mol. Cell Biol.* **5,** 875-885.
5. Tsai, M. J. and Omalley, B. W. (1994) *Annu. Rev. Biochem.* **63,** 451-486.
6. Ivan, M., Kondo, K., Yang, H. F., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J. M., Lane, W. S. and Kaelin, W. G. (2001) *Science* **292,** 464-468.
7. Semenza, C. L. (2001) *Pediatr. Res.* **49,** 614-617.
8. Darnell, J. E. (2002) *Nat. Rev. Cancer* **2,** 740-749.
9. Dervan, P. B. (2001) *Bioorg. Med. Chem.* **9,** 2215-2235.
10. Roeder, R. G. (2005) *FEBS Lett.* **579,** 909-915.
11. Mitchell, P. J. and Tjian, R. (1989) *Science* **245,** 371-378.
12. Dynan, W. S. and Tjian, R. (1985) *Nature* **316,** 774-778.
13. Korzus, E., Torchia, J., Rose, D. W., Xu, L., Kurokawa, R., McInerney, E. M., Mullen, T. M., Glass, C. K. and Rosenfeld, M. G. (1998) *Science* **279,** 703-707.
14. Kamei, Y., Xu, L., Heinzel, T., Torchia, J., Kurokawa, R., Gloss, B., Lin, S. C., Heyman, R. A., Rose, D. W., Glass, C. K., et al. (1996) *Cell* **85,** 403-414.

15. Malik, S. and Roeder, R. G. (2005) *Trends Biochem.Sci.* **30,** 256-263.

16. Dickerson, R. E., Drew, H. R., Conner, B. N., Wing, R. M., Fratini, A. V. and Kopka, M. L. (1982) *Science* **216,** 475-485.

17. Nikolov, D. B., Chen, H., Halay, E. D., Hoffmann, A., Roeder, R. G. and Burley, S. K. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93,** 4862-4867.

18. Leonard, G. A. and Hunter, W. N. (1993) *J. Mol. Biol.* **234,** 198-208.

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

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

21. Pabo, C. O. and Sauer, R. T. (1992) *Annu. Rev. Biochem.* **61,** 1053-1095.

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

23. Wu, H., Yang, W. P. and Barbas, C. F. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 344-348.

24. Mitra, S. N., Wahl, M. C. and Sundaralingam, M. (1999) *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **55,** 602-609.

25. Kumar, R. A., Ikemoto, N. and Patel, D. J. (1997) *J. Mol. Biol.* **265,** 187-201.

26. Quintana, J. R., Lipanov, A. A. and Dickerson, R. E. (1991) *Biochemistry* **30,** 10294-10306.

27. Arcamone, F., Nicolell.V, Penco, S., Orezzi, P. and Pirelli, A. (1964) *Nature* **203,** 1064.

28. Mrksich, M., Wade, W. S., Dwyer, T. J., Geierstanger, B. H., Wemmer, D. E. and Dervan, P. B. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 7586-7590.

29. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E. and Dervan, P. B. (1998) *Nature* **391,** 468-471.

30. Briehn, C. A., Weyermann, P. and Dervan, P. B. (2003) *Chem.-Eur. J.* **9,** 2110-2122.

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

32. Marques, M. A., Doss, R. M., Foister, S. and Dervan, P. B. (2004) *J. Am. Chem. Soc.* **126,** 10339-10349.

33. Dervan, P. B., Poulin-Kerstien, A. T., Fechter, E. J. and Edelson, B. S. (2005) in *DNA Binders and Related Subjects*, **253**, pp. 1-31.

34. Viger, A. and Dervan, P. B. (2006) *Bioorg. Med. Chem.* **14,** 8539-8549.

35. Chenoweth, D. M., Poposki, J. A., Marques, M. A. and Dervan, P. B. (2007) *Bioorg. Med. Chem.* **15,** 759-770.

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

37. Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B. and Rees, D. C. (1998) *Science* **282,** 111-115.

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

39. Foister, S., Marques, M. A., Doss, R. M. and Dervan, P. B. (2003) *Bioorg. Med. Chem.* **11,** 4333-4340.

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

41. deClairac, R. P. L., Geierstanger, B. H., Mrksich, M., Dervan, P. B. and Wemmer, D. E. (1997) *J. Am. Chem. Soc.* **119,** 7909-7916.

42. Herman, D. M., Baird, E. E. and Dervan, P. B. (1998) *J. Am. Chem. Soc.* **120,** 1382-1391.

43. Trauger, J. W., Baird, E. E. and Dervan, P. B. (1996) *Nature* **382,** 559-561.

44. Herman, D. M., Turner, J. M., Baird, E. E. and Dervan, P. B. (1999) *J. Am. Chem. Soc.* **121,** 1121-1129.

45. Herman, D. M., Baird, E. E. and Dervan, P. B. (1999) *Chem.-Eur. J.* **5,** 975-983.

46. Weyermann, P. and Dervan, P. B. (2002) *J. Am. Chem. Soc.* **124,** 6872-6878.

47. Heckel, A. and Dervan, P. B. (2003) *Chem.-Eur. J.* **9,** 3353-3366.

48. Olenyuk, B., Jitianu, C. and Dervan, P. B. (2003) *J. Am. Chem. Soc.* **125,** 4741-4751.

49. Oakley, M. G., Mrksich, M. and Dervan, P. B. (1992) *Biochemistry* **31,** 10969-10975.

50. Fechter, E. J. and Dervan, P. B. (2003) *J. Am. Chem. Soc.* **125,** 8476-8485.

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

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

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

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

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

56. Chiang, S. Y., Burli, R. W., Benz, C. C., Gawron, L., Scott, G. K., Dervan, P. B. and Beerman, T. A. (2000) *J. Biol. Chem.* **275,** 24246-24254.

57. Yang, F., Belitsky, J. M., Villanueva, R. A., Dervan, P. B. and Roth, M. J. (2003) *Biochemistry* **42,** 6249-6258.

58. Simon, H., Kittler, L., Baird, E., Dervan, P. and Zimmer, C. (2000) *FEBS Lett.* **471,** 173-176.

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

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

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

62. Kwonj, Y., Arndt, H. D., Qian, M., Choi, Y., Kawazoe, Y., Dervan, P. B. and Uesugi, M. (2004) *J. Am. Chem. Soc.* **126,** 15940-15941.

63. Dickinson, L. A., Gulizia, R. J., Trauger, J. W., Baird, E. E., Mosier, D. E., Gottesfeld, J. M. and Dervan, P. B. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 12890-12895.

64. Janssen, S., Durussel, T. and Laemmli, U. K. (2000) *Mol. Cell.* **6,** 999-1011.

65. Janssen, S., Cuvier, O., Muller, M. and Laemmli, U. K. (2000) *Mol. Cell.* **6,** 1013-1024.