Localization of DNA-Binding Polyamides In Living Cells

Thesis by

Timothy Patrick Best

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

California Institute of Technology

Pasadena, California

2005

(Defended 22 July, 2004)

…for my parents…

Acknowledgements

The foundation for any successful effort is its most important part, bearing the weight of the work built upon it. My foundation has been the support of my family: aunts, uncles, and cousins all; my grandmother, Olive Sonner; my sister, Jennifer. Most especially I would like to thank my parents, Ralph and Sharron Best; they have always been, and continue to be, my example. To those people whose words and deeds interested me in science, I offer my thanks: the writings of Carl Sagan; from Corydon—Larry Hauswald and my first teacher of chemistry, Dennis Lopp; from Butler University—Joe Kirsch, Shannon Lieb, Dave Hall, and especially my advisor and friend Anne Wilson.

I would like to thank my research advisor, Peter Dervan, for his enthusiasm and for providing the opportunity to work in a rich academic environment with such stellar co-workers. I would like to thank the members of my committee, Linda Hsieh-Wilson, Carl Parker, and Scott Fraser for their support and guidance. I would also like to thank Dr. Fraser for providing the use of his microscope facilities. I would like to extend my appreciation to Harry Gray for providing an opportunity for weekly escape to the Rathskellar after Friday seminars, and to his wife for excellent pecan pie.

The high quality of my collaborators and co-workers at Caltech has provided an exemplary environment in which to learn about both chemistry and the world: I would like to thank the coffee-time club of postdoctoral fellows Anna Mapp, Inger Kers, Ulf Ellervik, Leonard Prins, Phillip Weyermann, Christoph Briehn, and Dorte Renneberg; French lessons from Pierre Potier; German from Alex Heckel and Hans-Dieter Arndt; my first teacher in the lab and perennial fantasy baseball foe Bobby Arora; and Bogdan

Olenyuk, who has always been a source of good chemical knowledge, good conversation, and good help.

I appreciate the fellowship of many students in the Dervan lab, past and present: Dave Herman, Merideth Howard, Nick Wurtz, John Chevillet, Adam Urbach, Amanda Cashin, Victor Rucker, Jason Belitsky, Shane Foister, Eric Fechter, Michael Marques, Jim Sanchez, Nick Nickols; the Freeballers (Ray Doss, Justin Cohen, Carey Hsu, Ryan Stafford, Adam Poulin-Kerstien); my newest co-workers Jim Puckett, Julie Popowski, and Sherry Tsai; and my parter in crime Ben Edelson. I have also had good co-workers outside of the Caltech community: Zhengxin Wang, Gerd Blobel, and Aseem Ansari.

As the eminent philosophers McCartney and Lennon put it, I've gotten by with a little help from my friends. The Corydon crew: John McCollum, Jared Bachman, Joe Harmon, Jeremy Schoen, Matt Burnham, Jay Hanaver, Micky Emily and Melody Mathes; from Butler: Mike and Kathleen Julius, Sarah Bohl, Pete Wibbenmeyer, Jill Carter, and Sherry Nichols. Caltech has been no exception, as I have found many good people here with whom to share the time: the Wednesday lunch crew—Jeremy Heidel, Steve Spronk, Julie Casperson, and Swaroop Mishra; the house on Mentor—James Peterson and Greg Drummond; the gamers—Micol Christopher, Mat Matuszewski, Meg Wessling, Carlos Mochon, James Chakan, Charlene Ahn, and Geoff Swift; Jenny Roizen, for many interesting conversations and her good company; my roommate and friend Neal Oldham; and Ted Corcovilos, whom I would also like to thank for helping me to edit this text.

To anyone whom I have forgotten and who deserves my thanks: Cheers!

Abstract

Regulation of the processing of genes into nucleic acids and proteins is a substantial goal in medicine. Small molecules that could enter cells, localize to the nucleus, and bind chromosomal DNA sequence-specifically and with high affinity would be important tools for gene regulation. Pyrrole-imidazole polyamides are small molecules that bind the minor groove of DNA in a sequence-specific fashion according to a set of pairing rules, and with affinities rivaling natural transcription factors. Several *in vitro* experiments have shown that by directly competing with transcription factors for binding sites in gene promoter regions, polyamides can act to inhibit transcription of those genes. Polyamides bearing transcription activation domains can bind to promoter regions, recruit the transcriptional machinery to the gene, and activate transcription *in vitro*. Attempts to reproduce these results *in vivo* were largely unsuccessful, perhaps due to poor cellular trafficking properties of polyamides and polyamide-peptide conjugates.

It was found that polyamides bearing the Bodipy fluorophore localize primarily to the cytoplasm of cells, or were excluded from cells altogether. In attempts to overcome this quality, peptides shown to improve cellular trafficking were appended to the polyamides. These peptides were generally not successful at inducing uptake, and were in many cases toxic to the cells. Small molecules were also appended to polyamides, likewise to improve uptake properties, but met with limited success. Surprisingly, the addition of a fluorescein or fluorescein-like fluorophore to polyamides permit them to localize to the nuclei of all cell lines tested, in a molecular content- and shape-dependent manner. This technology has been applied to several *in vivo* experiments, including the inhibition of androgen receptor binding to its cognate element in gene promoter regions.

Table of Contents

List of Figures and Tables

Chapter 1 Page

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 1

Introduction to DNA Recognition by Minor Groove-Binding Polyamides

Background and Significance

DNA is the medium of information storage by which content from the large-scale combinatorial experiment of evolution has been stored and refined over time.[1] This information is retained in lengths of discrete linear polymers of deoxyribonucleic acids, or genes, which serve as instructions for the synthesis of an organism's complement of ribonucleic acids and proteins. Faithful reading, duplication, and utilization of genetic material are the essential steps by which living organisms are created, grow, and multiply. This privileged placement at the root of all processes necessary to the maintenance and propagation of life highlights DNA as an attractive target for a plethora of diagnostic and therapeutic applications. The recently completed initiative to record the complete human genome provides a wealth of knowledge necessary to take advantage of DNA as a target in human medicine. This achievement provides a blueprint for the ground floor of living systems. The human genome describes more than 30,000 putative genes, each a possible therapeutic or diagnostic target, or a subject of basic biochemical research.[2,3] Illumination of the organizational parameters of the genome, the diversity of individual variation in genetic sequence, and the dynamics of genetic processing are among the next growing points of biochemical research. The means by which these central questions will be pursued are diverse; among them are molecules designed to bind predetermined DNA sequences and provide some analyzable effect. Minor groove-binding polyamides, programmed to strongly and selectively bind desired DNA sequences, are powerful tools that may be leveraged in a multitude of diagnostic and therapeutic applications.[4]

Structural Features of DNA

Genomic material is composed of two complementary, antiparallel polydeoxyribonucleotide strands intertwined in a double helical structure. The strands are associated by specific hydrogen bonding interactions between the nucleotide bases, adenine pairing with thymine and guanine pairing with cytosine (Figure 1.1).[5] The sugar phosphate backbones of the paired strands define helical grooves, in whose floors the edges of the heterocyclic bases are exposed (Figure 1.2). The biologically relevant B-form of the DNA double helix is characterized by a wide and shallow major groove, and

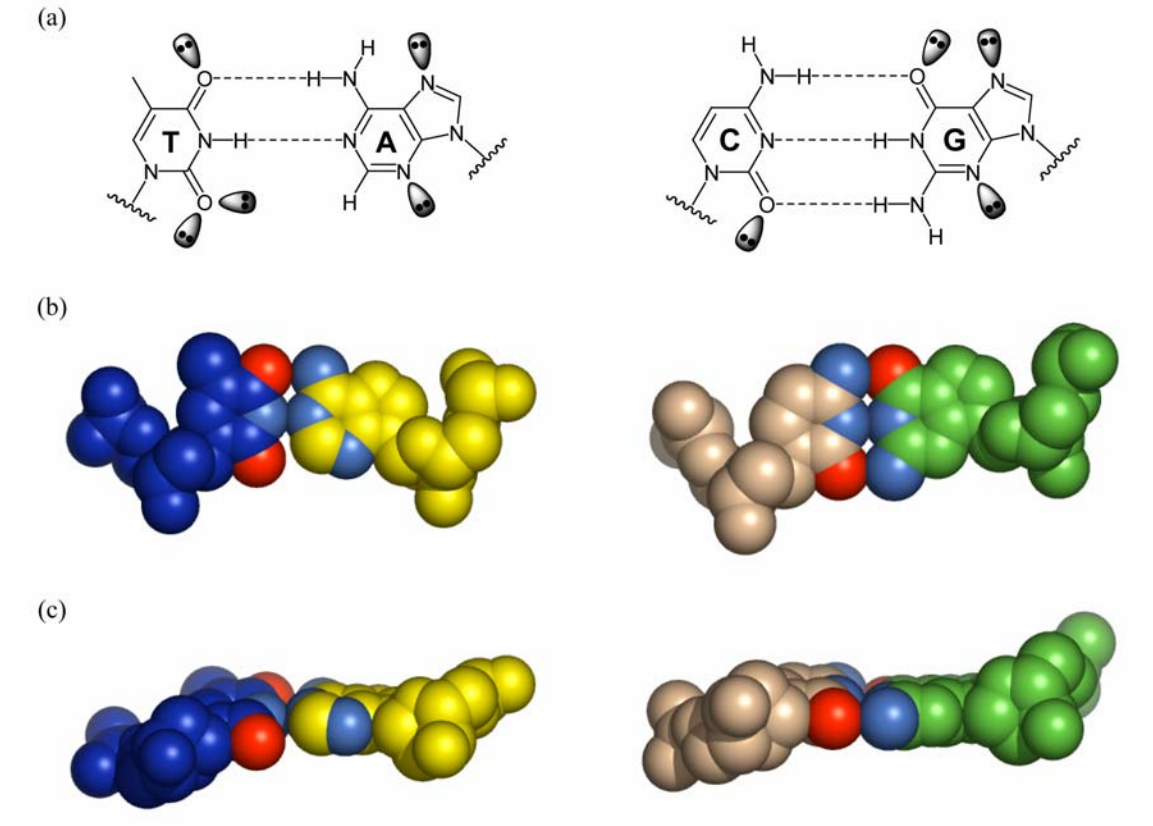

Figure 1.1 DNA base pairs. (a) The chemical structure of the A·T and G·C DNA base pairs. (b) Space-filling models, based on crystal structure data, for the A·T and C·C base pairs, aligned in an analogous fashion to the chemical structures. "T" base is in blue, "A" base is in yellow, "G" base is in green, and "C" base is in wheat. Nitrogen atoms at interface are in marine and oxygen atoms at interface are in red. (c) View of the A·T and G·C base pairs looking down into the minor groove, colored as in (b).[7]

a narrow and deep minor groove.[6] The molecular surfaces and chemical features presented by a given DNA sequence, though essentially identical to first order, at the atomic scale are distinct. This differentiation provides a basis for the sequence-selective recognition of DNA by proteins and small molecules.

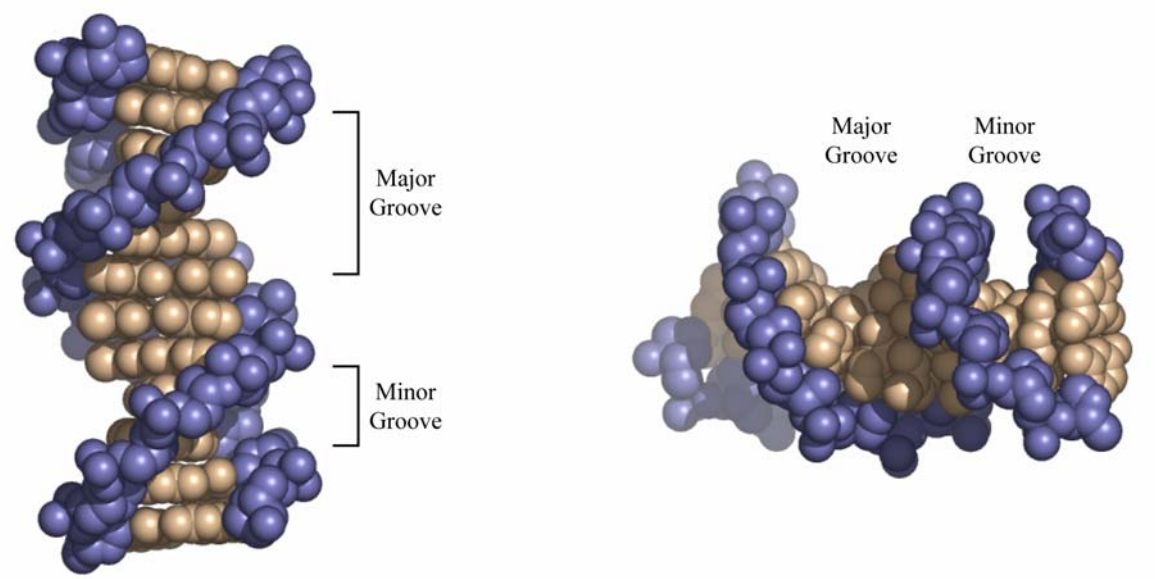

Figure 1.2 Structural features of the DNA double helix. The phosphate backbone of the DNA is in slate, and the nucleic acid bases are in grey. The major and minor grooves are outlined.[7]

DNA Recognition Strategies in Natural Systems

Evolution has resulted in the employment of a diverse selection of structural motifs by proteins recognizing DNA. Combinations of electrostatic interactions with the negatively-charged sugar phosphate backbone and van der Waals interactions with the nucleobases in the floor of the helical grooves provide both the affinity and specificity by which proteins bind DNA (Figure 1.3).[8-12] Specific interactions of proteins with DNA

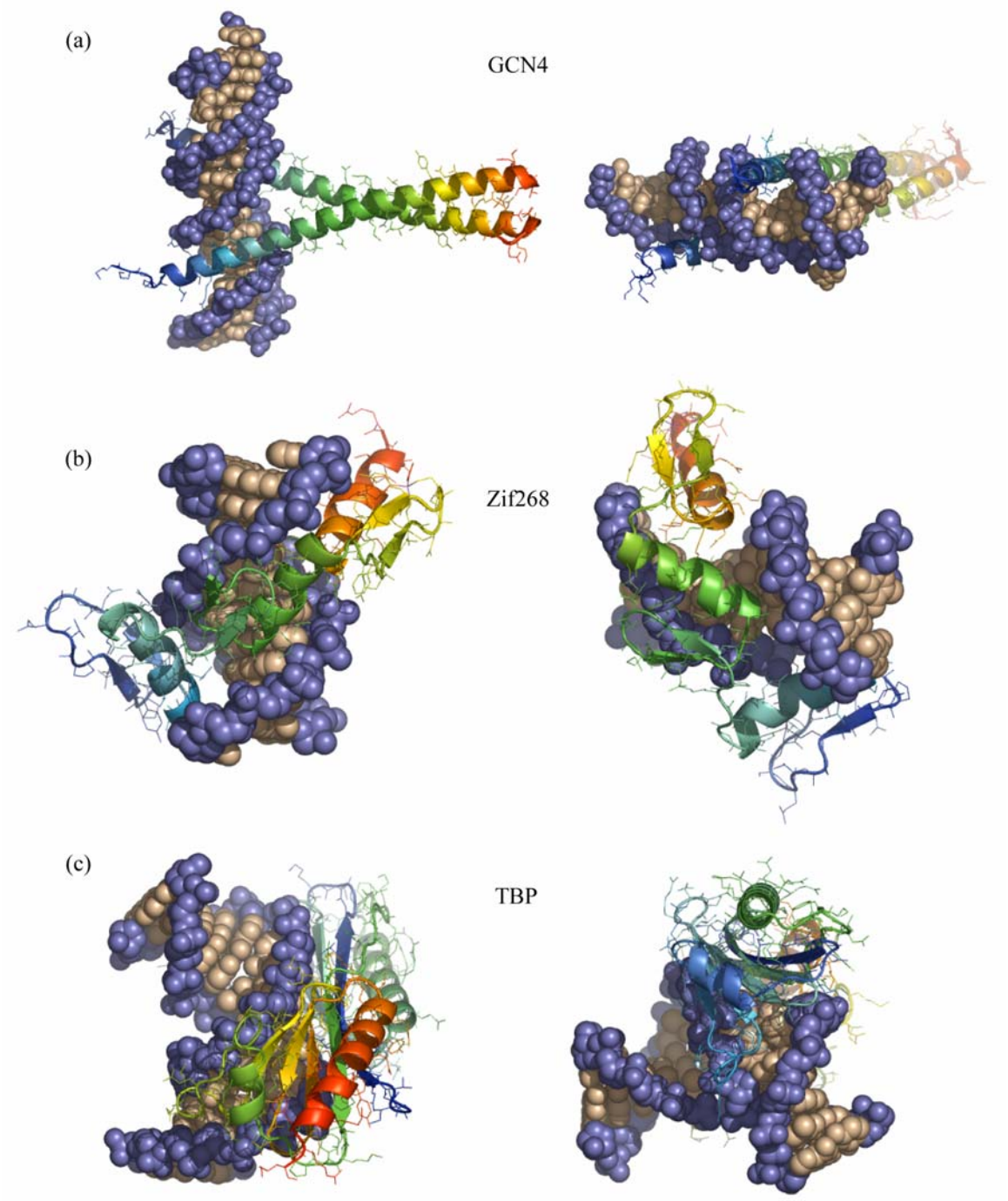

Figure 1.3 X-ray crystal structures of DNA recognition by proteins. (a) GCN4 recognizes DNA through homodimeric recognition of the major groove. (b) Zinc finger Zif268 through recognition of the major groove. (c) TBP recognizes DNA primarily through minor groove contacts, causing a severe bend to the DNA helix.

often require multiple proteins, or multiple copies of the same protein, to achieve the programmed effect. These interactions are dynamically rich phenomena, often proceeding through structural shifting of either or both the protein components and the DNA helix.

A majority of DNA-binding proteins rely on major groove contacts to provide sequence specificity, though a few, such as TATA-binding protein, utilize minor groove contacts. The engineering of proteins with novel DNA-binding properties is an ongoing effort in biochemical research. There have been some successes using phage display techniques to select zinc finger proteins that bind predetermined sequences. However, attempts to modify the DNA-binding qualities of these proteins in a rational manner have met with little success. This has shown that, to date, there is no known general recognition code linking target DNA base pair sequence with protein amino acid sequence.[13]

Though the complexity of protein structure makes the rational design of DNA-binding polypeptides quite challenging, other paradigms exist in nature for the recognition of DNA sequences. A number of small molecules have been discovered that bind DNA with high affinity and with some sequence specificity.[14-16] Many of these compounds utilize the narrow, deep minor groove as their recognition domain on the DNA double helix. These natural products include calicheamicin, Hoechst 33258, distamycin, and the closely related netropsin (Figure 1.4). Among these, the polypyrrole compounds netropsin and distamycin are particularly attractive as lead compounds due to their relative chemical simplicity, small size, and, most especially, their modular nature. The synthetic manipulation of the polypyrrole lead molecules into a sophisticated class of

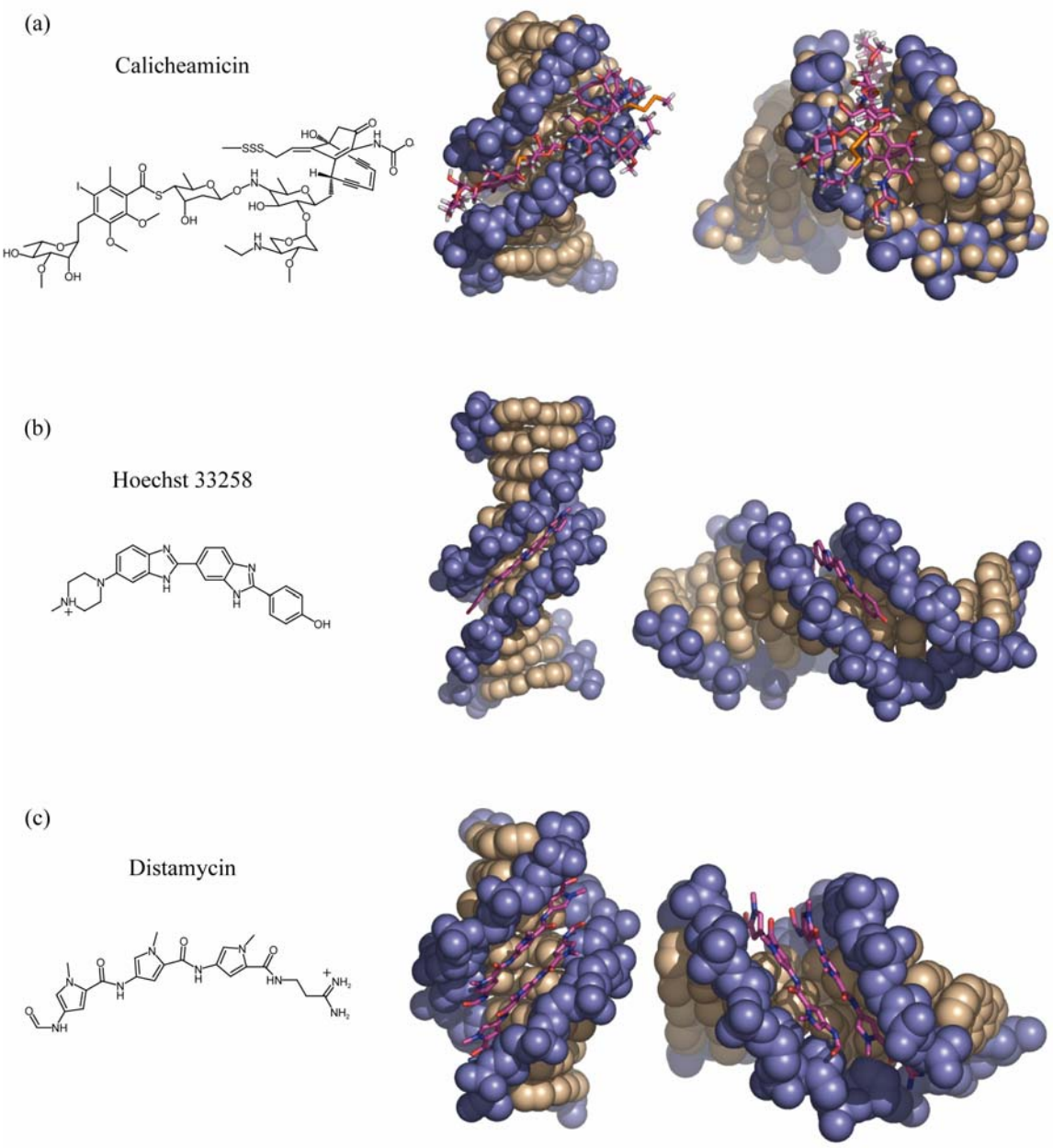

Figure 1.4 X-ray crystal structures of minor-groove binding small molecules. (a) Calicheamicin is an enediyne-containing oligosaccharide that upon recognition of sequences such as TCCT, TCTC, and TTTT, causes a double-strand cleavage. (b) Hoechst 33258 recognizes A·T, T·A domains selectively. (c) Distamycin also recognizes A·T, T·A tracts selectively, and may bind as either a monomer, or as antiparallel dimers (as shown) within the minor groove.

heterocyclic oligomers has led to the investigation of the static and dynamic parameters governing their interaction with the DNA minor groove in great detail.[17]

The laboratories of Prof. Peter B. Dervan have utilized the distamycin scaffold to synthesize a new class of small molecules which are able to bind the minor groove of DNA with high affinity and specificity.[4,18-20] This set of molecules is composed of *N*-methylimidazole (Im), *N*-methylpyrrole (Py), *N*-methyl-3-hydroxypyrrole (Hp), 3-chloro-thiophene-2-carboxamide (Ct), benzimidazole (Bi), and hydroxybenzimidazole (Hz) amino acid residues linked into crescent-shaped oligomers. As with distamycin, polyamides may bind the DNA minor groove as monomeric units or as antiparallel dimers, in a concentration-dependent manner. DNA association is driven by a combination of van der Waals interactions and specific hydrogen bonds. Side-by-side pairings of aromatic residues stack the polyamide heterocyclic rings against each other and the walls of the minor groove, positioning the polyamide backbone and aromatic 3-substituents to contact the edges of nucleotide bases in the minor groove floor. This arrangement allows polyamides to exploit the specific pattern of hydrogen bond donors and acceptors present along the edges of the nucleic acid bases, as well as the subtle variations in molecular shape in the floor of the minor groove, to recognize and bind to specific Watson-Crick base pairs (Figure 1.5).[21]

Exhaustive physical studies of the interaction of minor groove-binding polyamides with DNA using a variety of techniques, including DNase I footprinting, X-ray crystallography, multidimensional NMR spectroscopy, and fluorescence asssays, have yielded a set of guidelines. These *pairing rules* dictate that specific unsymmetrical pairings of Im with Hp, Hz, and Py residues, and of Ct with Py, in a cofacial

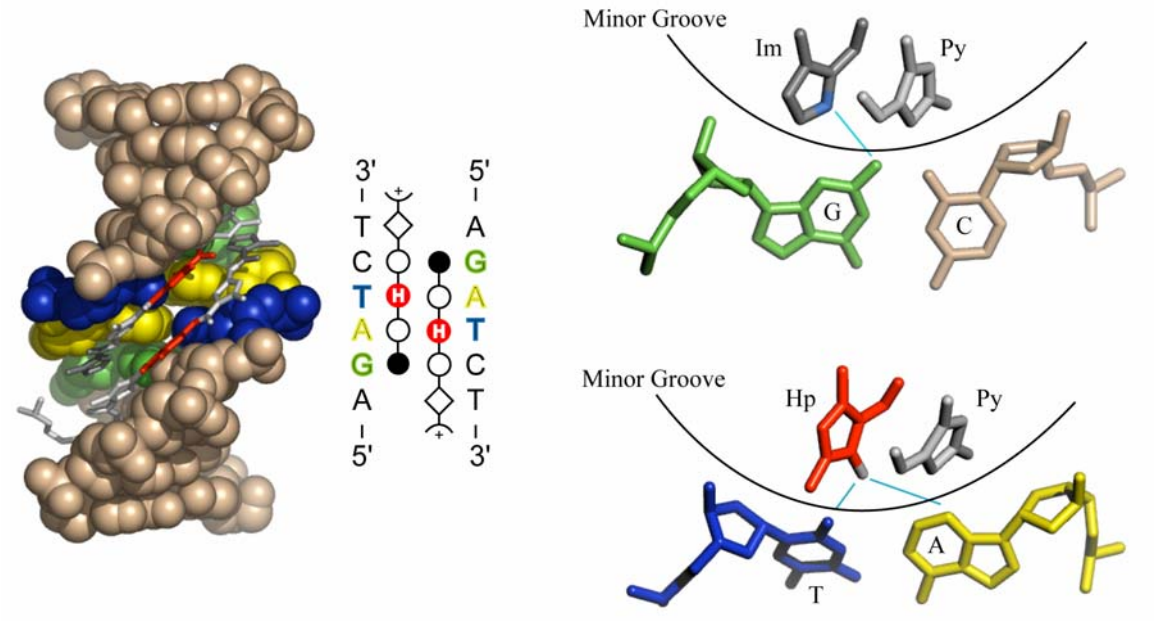

Figure 1.5 Minor groove recognition by polyamides. An X-ray crystal structure of a DNA helix bound by homodimeric polyamides.[26] The polyamides and the DNA base pairs they recognize are represented in the ball-and-stick model. A black circle represents Im, a white circle represents Py, a red circle containing an "H" represents Hp, a diamond represents β-alanine, and an arc with a "+" represents Dp. The recognition of minor groove contacts in a T·A base pair by Hp/Py and the recognition of a G·C base pair by Im/Py are shown at right, from the same crystal structure. DNA bases in the polyamide binding site are colored as in Figure 1.1.

arrangement, underlie the binding characteristics of minor groove-binding polyamides.[22] A pairing of Im with Py (symbolized Im/Py) targets a G·C base pair, while Py/Im targets C·G.[23] The physical basis for this specific interaction is primarily due to a linear hydrogen bond formed between the N3 of the Im residue and the exocyclic amine of guanine.[24] A pairing of Py with itself (Py/Py) is degenerate for both T·A and A·T. In order to specify between adenine and thymine DNA residues, either the Hp or Hz residues may be employed. Hp/Py and Hz/Py target T·A, while Py/Hp and Py/Hz target A·T. The selectivity of Hp is derived from both hydrogen bonding interactions with the thymine O2, and shape-selective recognition of the asymmetric cleft between the T·A

versus A·T base pair.[21,25,26] The selectivity of Hz is thought to be due to the same structural and chemical reasons, though no X-ray crystal structures or NMR structures have been determined thus far.[27] A pairing of Ct with Py at the N-terminal cap position has been shown to target T·A selectively over A·T, G·C, or C·G, it is thought through projection of the 3-chloro substituent into the floor of the minor groove.[28]

The Hairpin Motif

The ability of linear polyamides to dimerize within the DNA minor groove suggested the possibility of creating a covalent linkage between the antiparallel strands, reducing the entropic cost of association with DNA. This hairpin motif, connecting the N-terminus of one polyamide strand with the C-terminus of the second strand through an aliphatic γ-aminobutyric acid residue (γ), provides ligands with affinities and specificities

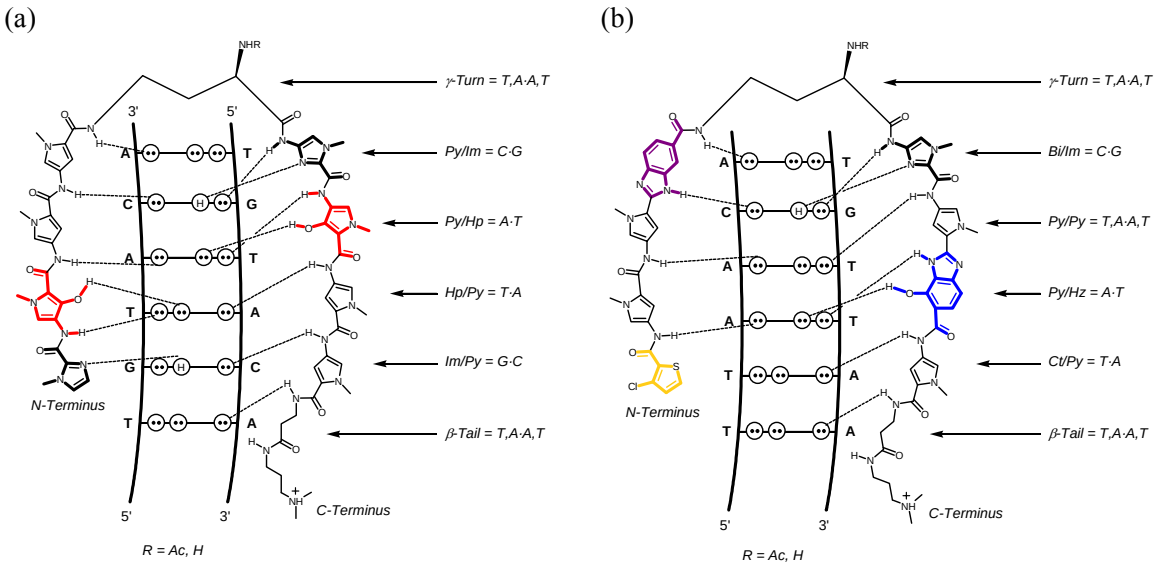

Figure 1.6 Hydrogen-bonding model of the hairpin motif. (a) H-bonding and pairing rules for classic Im/Py/Hp polyamides. (b) H-bonding and pairing rules for hairpin polyamides incorporating new aromatic residues Ct, Bi, and Hz. Im residues are in bold black, Py residues are in regular black, Hp residues are in red, Bi residue is in violet, Hz residue is in blue, Ct residue is in yellow.

rivaling DNA-binding proteins.[29,30] Hairpin polyamides utilize the same hydrogen bonding and van der Waals interactions to govern DNA binding, exhibiting the same orientational preference as unlinked dimers, aligning N→C with respect to the 5'→3' direction of the adjacent DNA strand (Figure 1.6).[31,32]

Exceptional cases have been observed in which some hairpins in some DNA contexts will bind in a reverse fashion, N→C with respect to the 3'→5' direction of the adjacent DNA strand. There are also examples of hairpins binding DNA in an unfolded conformation as a single extended strand. Both of these issues are resolved upon introduction of a chiral amine moiety on the α carbon of the γ-turn residue.[33] Both turn residues exhibit selectivity for A·T and T·A base pairs over G·C and C·G base pairs, presumably due to steric clashes with the guanine exocyclic amine.[30,33]

Similarly, the aliphatic β-alanine (β) and *N*,*N*-dimethylaminopropylamine (Dp) residues often found at the C-terminal tails of polyamides are selective for A·T and T·A base pairs over G·C and C·G. Though imparting additional selectivity to polyamide binding, the tail residues are not critical to DNA binding, and can be replaced or removed.[34] Often, the β-alanine residue is useful as an internal residue within the polyamide strand, paired against either Im, Py, or other β residues.[35,36] This utility arises due to the supercurvature of polyamide strands with reference to the curvature of the DNA minor groove. Though the rise per residue of polyamides correlates closely with the pitch of the B-form DNA helix, beyond five consecutive aromatic residues the shape of the polyamide is no longer complementary to the DNA minor groove.[37] Internal β-alanine residues, which are inherently more flexible than aromatic residues, allow the

relaxation of the curvature of polyamide strands, allowing DNA sequences longer than five base pairs to be recognized effectively.

Despite the successes of the hairpin polyamide motif, the set of sequences that the hairpin motif can target is limited by the T,A selectivity of the aliphatic turn moiety. In order to enlarge the set of targetable DNA sequences, other polyamide motifs have been developed, employing a variety of strategies to covalently-link individual polyamide strands. These new motifs include cycles[38], H-pins[39], and U-pins[40] (Figure 1.7).

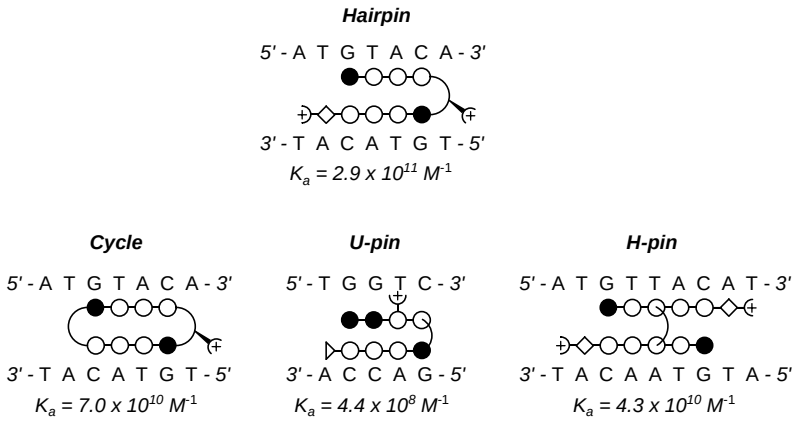

Figure 1.7 Ball-and-stick schematic for covalently linked polyamide motifs.

Gene Regulation with Polyamides

By utilizing the pairing rules, minor groove-binding polyamides may be synthesized to bind predetermined DNA sequences, including those within the promoter regions of genes. Proper placement of polyamides in the promoter sequence may allow them to interfere with the association of transcription factors, leading to inhibition of transcription. It has been demonstrated that polyamides are able to inhibit the binding of certain zinc finger proteins (including Zif268, Figure 1.3b) that bind DNA without any

minor groove contacts, presumably by an allosteric mechanism.[41] However, minor groove-binding ligands can also co-occupy DNA while certain proteins, such as GCN4 (Figure 1.3a) occupy the major groove.[42] GCN4 binding has been inhibited successfully by hairpin polyamides bearing "positive patches" targeting protein-phosphate contacts,[43,44] and by polyamide-intercalator conjugates.[45]

Better targets for polyamides have been minor groove-binding transcription factors, such as TBP (Figure 1.3c) and LEF-1. Several transcription factors and promoters have been successfully targeted by hairpin polyamides, including the TFIIIA zinc finger[46,47] and the HIV-1 promoter,[48,49] resulting in inhibition of RNA polymerase II transcription of targeted genes *in vitro* (Figure 1.8).

It is also possible to selectively activate transcription by synthesizing a polyamide, designed to bind in the promoter region of a gene of interest, to which has been appended a moiety that recruits the transcriptional machinery to the promoter (Figure 1.9). Polyamides bearing a viral peptide known to activate transcription have been used to increase transcription of diagnostic genes more than 30-fold over basal levels.[50-52] Since polyamides can target a wide variety of sequences with high affinity and specificity, this approach has great potential as a general method of gene-specific activation by small molecules. Though the principle was proven, little was determined about the physical parameters of the DNA/polyamide/transcriptional machinery interaction—especially the optimal distance between the DNA and the activation domain.

The successful use of polyamides in model systems has shown their promise as diagnostic and therapeutic agents in more complex environments. For a polyamide to have a DNA-mediated effect in an organism, it must be able to enter the organism's cells,

traffic to the nucleus, and bind chromosomal DNA. The fundamental repeating unit of chromatin is the nucleosome, which consists of the nucleosome core particle (NCP) and 20-80 base pairs of linker DNA. The NCP is built from two superhelical turns of DNA (147 base pairs) wrapped around a core of eight histone proteins to form a disc-shaped tightly-packed unit.

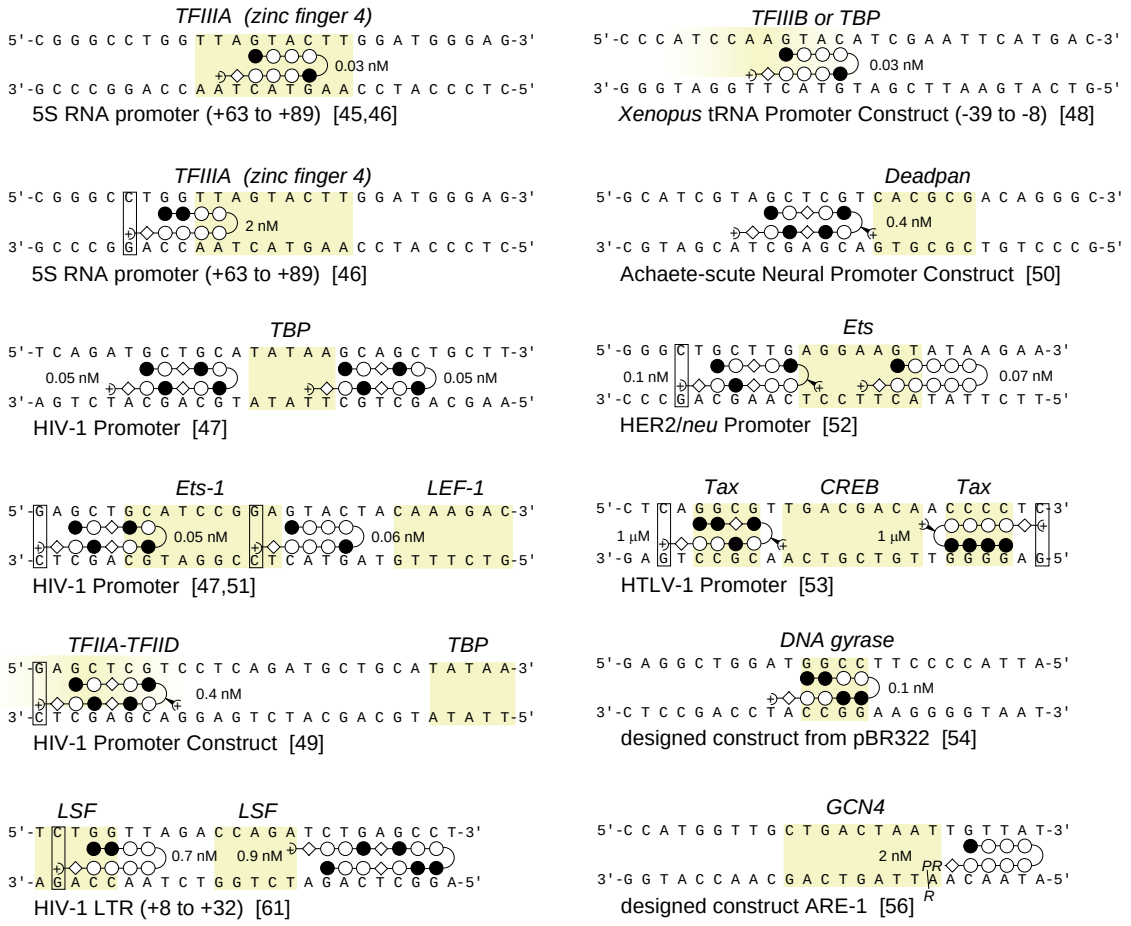

Figure 1.8 Hairpin polyamides have successfully inhibited the binding of several proteins to DNA in solution. Transcription factor binding sites are shaded.

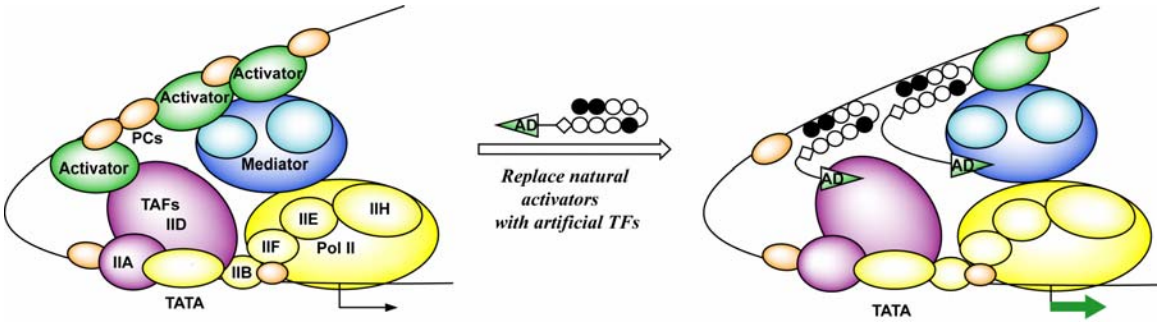

Figure 1.9 Polyamide-peptide conjugates have successfully been used to activate transcription of a target gene *in vitro*. In this experiment, the natural transcription factor (TF) activators (green ovals) have been replaced by polyamide-activator conjugates, binding the appropriate gene promoter sites. These artificial TF's interact with other transcription factors (purple ovals) and parts of the transcriptional machinery (yellow ovals), sometimes via mediating proteins (blue ovals) to recruit the transcriptional machinery to the target gene, initiating transcription.

The question of whether polyamides can bind chromosomal DNA is important because the vast majority of the DNA in a cell is locked in nucleosomes. In order to determine the binding characteristics of polyamides on the NCP, several X-ray crystal structures were completed of nucleosome core particles into which had been bound polyamides (Figure 1.10).[53] The structures show that not only can polyamides bind nucleosomal DNA, they have some access to the internal regions of the nucleosome core particle, where they can co-localize with the histone proteins. The only sites which were found to be inaccessible to polyamide binding were those blocked by the presence of the tails of the histone proteins in the minor groove.

In order to access chromosomal DNA, polyamides must be able to localize to the nucleus of cells. The cellular trafficking characteristics of polyamides have been studied by attachment of a fluorescent dye to a small selection of polyamides, exposing the resulting conjugates to various mammalian and insect cell lines, and viewing their

subcellular localization by confocal laser scanning microscopy.[54] These results showed that in a limited number of cell lines, primarily lines derived from mammalian T-cells, fluorescent polyamide conjugates could enter the cells and traffic to the nucleus (Figure 1.11). However, in most cell lines, these compounds either did not enter the cells, or entered the cells but were sequestered in the cytoplasm.

This raises the issue of how might the uptake profile of polyamides and polyamide conjugates be altered in order that they be useful in the context of living

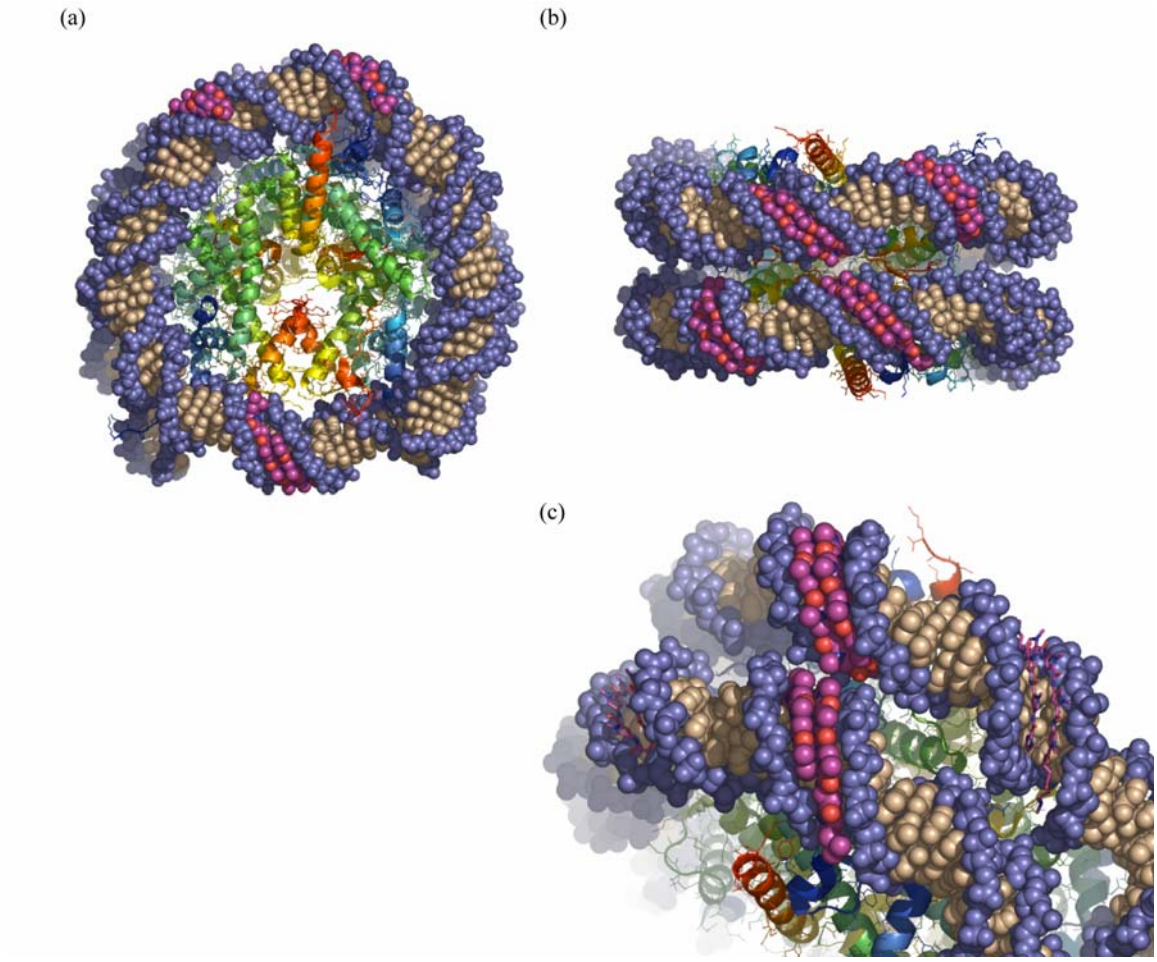

Figure 1.10 X-ray crystal structure of a hairpin polyamide bound to the nucleosome core particle (NCP). Top-down (a) and side-on (b) views of several polyamide molecules bound to the NCP. Histone proteins are depicted as multicolored cartoons. (c) View down the minor "supergroove" in which are bound two hairpin polyamides.

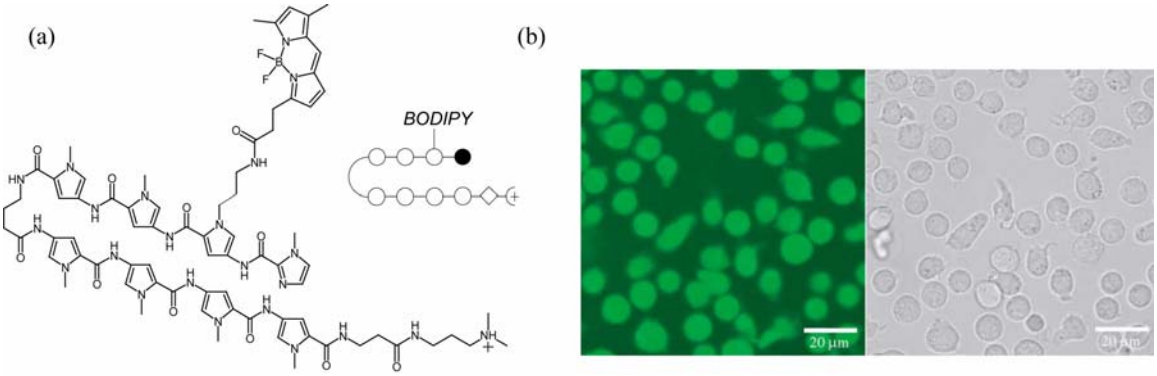

Figure 1.11 Cellular uptake of a fluorescent polyamide. (a) Chemical structure and ball-and-stick representations of a hairpin polyamide labeled with the fluorophore Bodipy. (b) CEM cells stained with Bodipy-polyamide. At left is a fluorescence-only image, and at right is the bright-field image. Polyamide is located in the nucleus.

organisms. Several methods of altering the subcellular localization of molecules have been developed for medicinal chemistry and diagnostic applications. The application of many of these methods to polyamides is the subject of the majority of this work. By creating polyamides or polyamide conjugates that are able to localize to the nucleus of living cells, technology advances one step closer to *in vivo* regulation of gene transcription by small molecules.

Scope of This Work

This thesis describes work examining DNA-binding polyamides in biological systems. Chapter 2 is conceptually related to the transcription activation studies outlined as in Figure 1.9. In previous work, polyamide-activation peptide conjugates had been synthesized with either short or long flexible linkers separating the two active domains. These compounds were efficient activators of transcription, though it was uncertain what effect the spacing between the activating domain and the DNA had on RNA Pol II recruitment and transcription activation. In order to answer this question, a series of polyamide-polyproline-activation peptide compounds were synthesized and analyzed for their ability to activate transcription. The polyproline domain is a rigid linker allowing discrete spacing studies to be undertaken. These experiments, though successful *in vitro*, were not successful *in vivo*. It was hypothesized that this result was due to poor cellular uptake properties.

Chapter 3 details some attempts that were made to overcome poor cellular uptake and nuclear localization of polyamides. Chapter 3A recounts the addition of peptide domains shown in other systems to increase the uptake of conjugated small molecules and oligonucleotides to polyamides, along with the fluorophore Bodipy. Chapter 3B records the attachment of small molecules to polyamides in attempts to increase their cellular and nuclear localization. The subcellular localization characteristics of these compounds were determined with confocal laser scanning microscopy. Though largely unsuccessful, these experiments provided useful experience in microscopy and cell culture methods that were necessary to any successful uptake results.

Chapter 4 recounts the modification of polyamides with fluorescein and other structurally similar fluorophores, and the successful use of these compounds to stain nuclei of living cells. Chapter 4A records early efforts to determine the uptake characteristics of hairpin polyamides conjugated to fluorescein isothiocyanate. Chapter 4B presents a much wider view of the uptake profiles of several polyamide motifs conjugated to a variety of fluorophores. Chapter 4C is a first-pass overview of the DNA-binding characteristics of these fluorophore conjugates, showing that the fluorescent compounds, themselves, are appropriate in many cases for use in transcription inhibition experiments *in vivo*. Chapter 5 is a record of one such application, using polyamides and polyamide conjugates to inhibit the interaction of the androgen receptor with its cognate DNA-binding site.

References

1) Brown, T.A. *Genomes*; Wiley-Liss: New York, 1999.

2) Consortium, I.H.G.S. *Nature* **2001**, *409*, 860.

3) Venter, J.C. et. al. *Science* **2001**, *291*, 1304.

4) Gottesfeld, J.M.; Neely, L.; Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Nature* **1997**, *387*, 202.

5) Dickerson, R.E.; Drew, H.R.; Conner, B.N.; Wing, M.; Fratini, A.V.; Kopka, M.L. *Science* **1982**, *216*, 475.

6) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag; New York, 1984.

7) Leonard, G.A.; Hunter, W.N. *J. Mol. Biol*. **1993**, *234*, 198.

8) Pabo, C.O.; Sauer, R.T. *Annu. Rev. Biochem*. **1992**, *61*, 1053

9) Ellenberger, T.E.; Brandl, C.J.; Struhl, K.; Harrison, S.C. *Cell* **1992**, *71*, 1223.

10) Nikolov, D.B.; Chen, H.; Halay, E.D.; Hoffman, A.; Roeder, R.G.; Burley, S.K. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4862.

11) Pavletich, N.P.; Pabo, C.O. *Science* **1991**, *252*, 809.

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

13) Jantz, D.; Amann, B.T.; Gatto, G.J., Jr.; Berg, J.M. *Chem. Rev*. **2004**, *104*, 789.

14) Kumar, R.A.; Ikemoto, N.; Patel, D.J. *J. Mol. Biol*. **1997**, *265*, 187.

15) Quintana, J.R.; Lipanov, A.A.; Dickerson, R.E. *Biochemistry* **1991**, *30*, 10294.

16) Mitra, S.N.; Wahl, M.C.; Sundaralingam, M. *Acta Crystallogr. D Biol. Crystallogr*. **1999**, *55*, 602.

17) Bailly, C.; Chaires, J.B. *Bioconj. Chem*. **1998**, *9*, 513.

18) Dervan, P.B.; Burli, R.W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688.

19) Dervan, P.B. *Bioorg. Med. Chem*. **2001**, *9*, 2215.

20) Dervan, P.B.; Edelson, B.S. *Curr. Opin. Struct. Biol*. **2003**, *13*, 284.

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

22) Wade, W.S.; Mrksich, M.; Dervan, P.B. *J. Am Chem. Soc* **1992**, *114*, 8783.

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

24) Kielkopf, C.L.; Baird, E.E.; Dervan, P.B. Rees, D.C. *Nat. Struct. Biol*. **1998**, *5*, 104.

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

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

27) Renneberg, D.; Dervan, P.B. *J. Am. Chem. Soc*. **2003**, *125*, 5707.

28) Foister, S.; Marques, M.A.; Doss, R.M.; Dervan, P.B. *Bioorg. Med. Chem. Lett.* **2003**, *11*, 4333.

29) Mrksich, M.; Parks, M.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1994**, *116*, 7983.

30) Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Nature* **1996**, *382*, 559.

31) deClairac, R.P.L.; Geirstanger, B.H.; Mrksich, M.; Dervan, P.B.; Wemmer, D.E. *J. Am. Chem. Soc*. **1997**, *119*, 7909.

32) White, S.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc.* **1997**, *119*, 8756.

33) Herman, D.M.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1998**, *120*, 1382.

34) Swalley, S.E.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1999**, *121*, 1113.

35) Turner, J.M.; Swalley, S.E.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1998**, *120*, 6219.

36) Wang, C.C.C.; Ellervik, U.; Dervan, P.B. *Bioorg. Med. Chem*. **2001**, *9*, 653.

37) Kelly, J.J.; Baird, E.E.; Dervan, P.B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6981.

38) Herman, D.M.; Turner, J.M.; Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1999**, *121*, 1121.

39) Olenyuk, B.; Jitianu, C.; Dervan, P.B. *J. Am. Chem. Soc*. **2003**, *125*, 4741.

40) Heckel, A.; Dervan, P.B. *Chem. Eur. J*. **2003**, *9*, 3353.

41) Nguyen-Hackley, D.H.; Ramm, E.; Taylor, C.M.; Joung, J.K.; Dervan, P.B.; Pabo, C.O. *Biochemistry* **2004**, *43*, 3880.

42) Oakley, M.G.; Mrksich, M.; Dervan, P.B. *Biochemistry* **1992**, *31*, 10969.

43) Bremer, R.E.; Baird, E.E.; Dervan, P.B. *Chem. Biol*. **1998**, *5*, 119.

44) Bremer, R.E.; Wurtz, N.R.; Szewczyk, J.W.; Dervan, P.B. *Bioorg. Med. Chem*. **2001**, *9*, 2093.

45) Fechter, E.J.; Dervan, P.B. *J. Am. Chem. Soc.* **2003**, *125*, 8476.

46) Gottesfeld, J.M.; Neely, L.; Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Nature* **1997**, *387*, 202.

47) Neely, L.; Trauger, J.W.; Baird, E.E.; Dervan, P.B.; Gottesfeld, J.M. *J. Mol. Biol*. **1997**, *274*, 439.

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

49) Dickinson, L.A.; Trauger, J.W.; Baird, E.E.; Dervan, P.B.; Graves, B.J.; Gottesfeld, J.M. *J. Biol. Chem*. **1999**, *274*, 12765.

50) Ansari, A.Z.; Mapp, A.K. Curr. Opin. Chem. Biol. 2002, 6, 765.

51) Mapp, A.K.; Ansari, A.Z.; Ptashne, M.; Dervan, P.B. Proc. Natl. Acad. Sci. U.S.A. 2000, 97, 3930.

52) Ansari, A.Z.; Mapp, A.K.; Nguyen, D.H.; Dervan, P.B.; Ptashne, M. Chem. Biol. 2001, 8, 583.

53) 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.

54) Belitsky, J.M.; Leslie, S.J.; Arora, P.S.; Beerman, T.A.; Dervan, P.B. *Bioorg. Med. Chem*. **2002**, *10*, 3313.