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

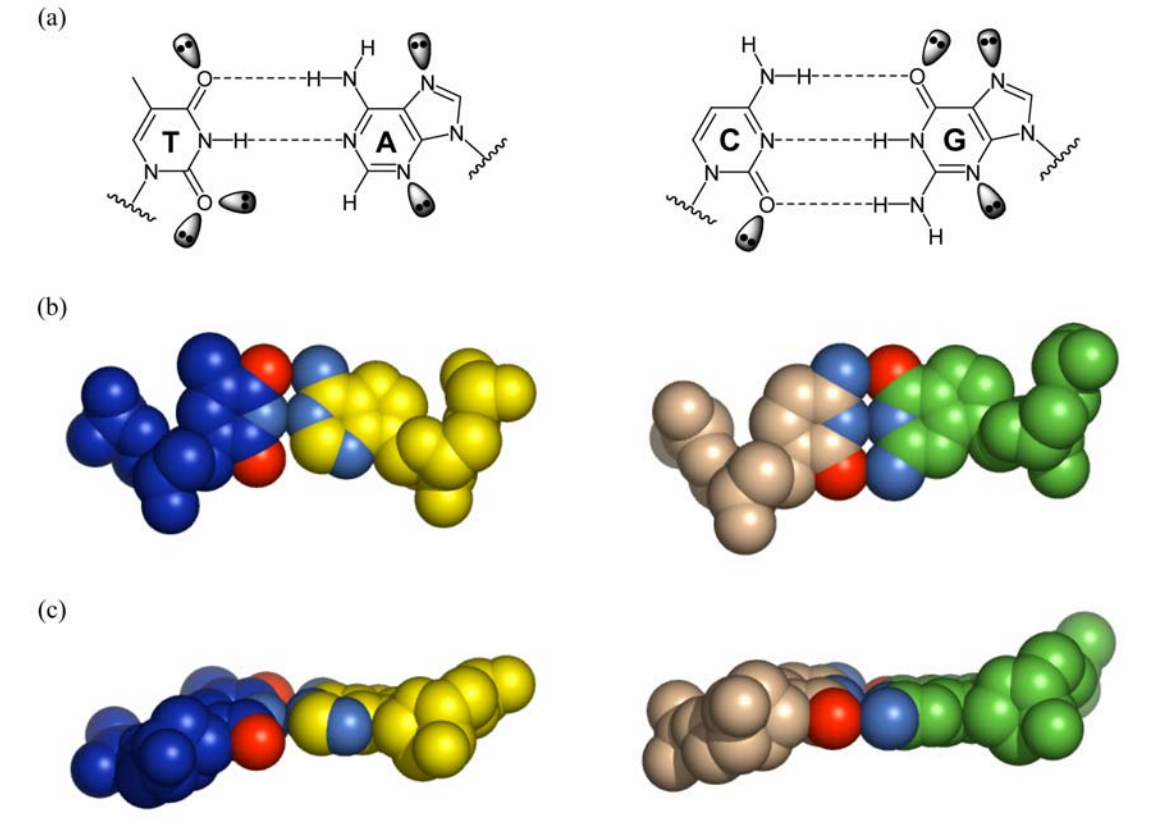


Figure 1.1 DNA base pairs. (a) The chemical structure of the A·T and G·C DNA base pairs. (b) Space-filling models, based on crystal structure data, for the A·T and C·C base pairs, aligned in an analogous fashion to the chemical structures. "T" base is in blue, "A" base is in yellow, "G" base is in green, 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.

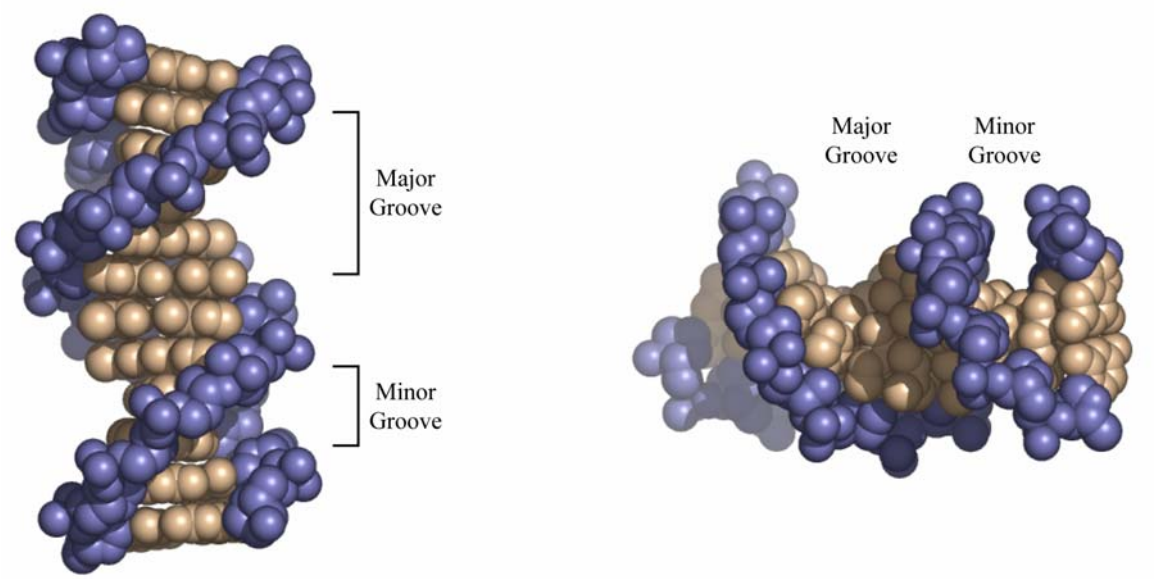


Figure 1.2 Structural features of the DNA double helix. The phosphate backbone of the DNA is in slate, and the nucleic acid bases are in grey. The major and minor grooves are outlined.[7]

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

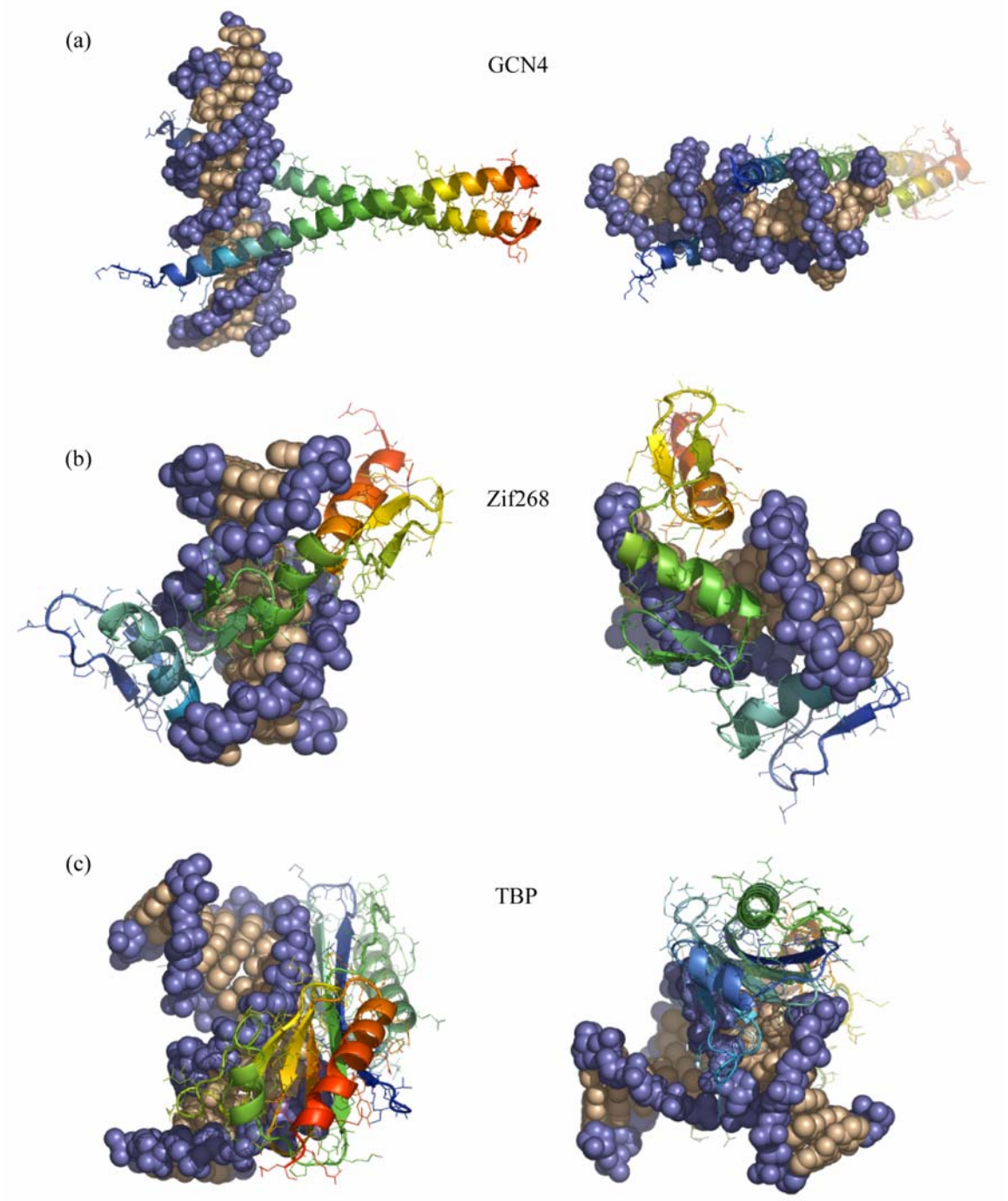


Figure 1.3 X-ray crystal structures of DNA recognition by proteins. (a) GCN4 recognizes DNA through homodimeric recognition of the major groove. (b) Zinc finger Zif268 through recognition of the major groove. (c) TBP recognizes DNA primarily through minor groove contacts, causing a severe bend to the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

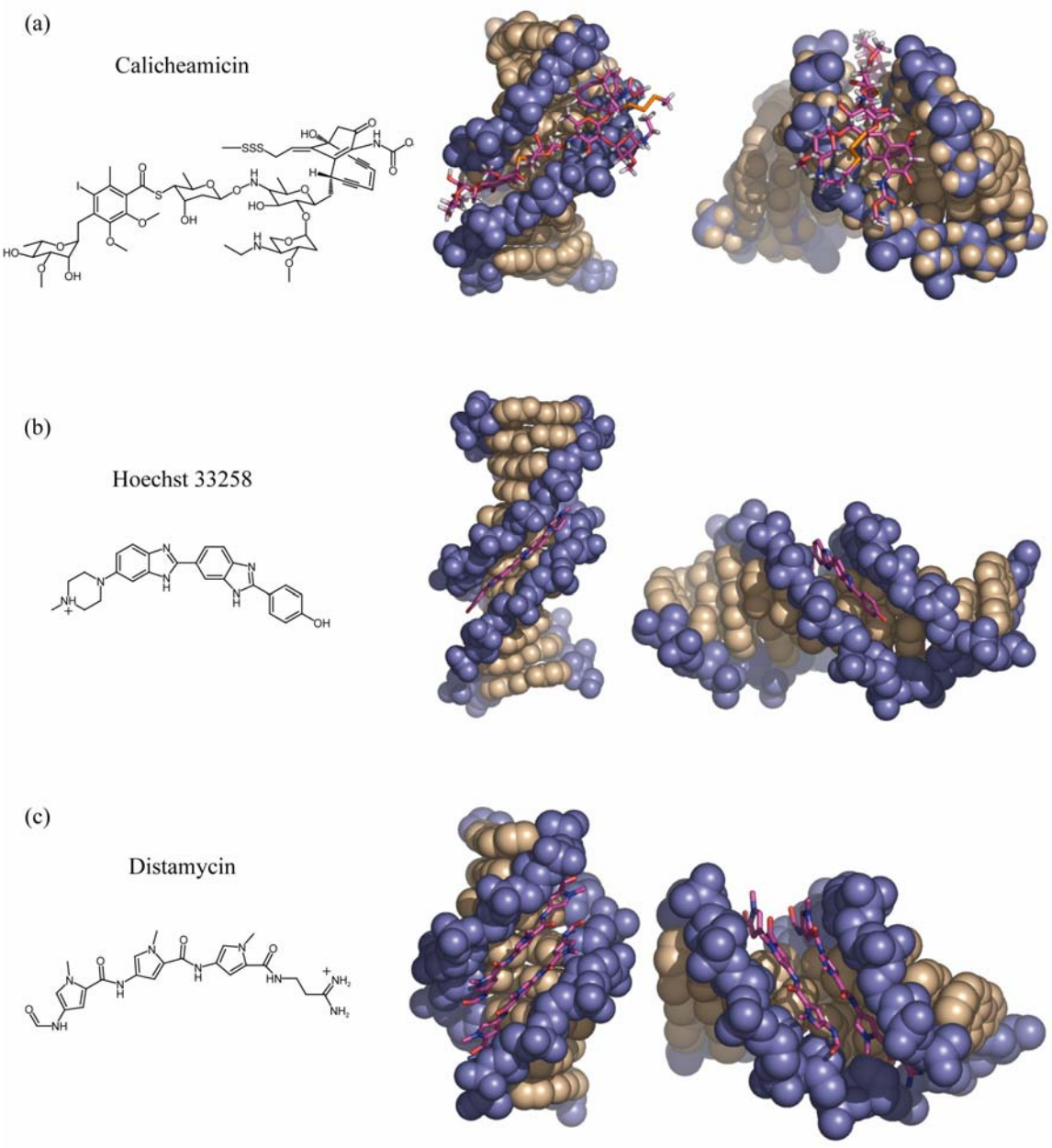


Figure 1.4 X-ray crystal structures of minor-groove binding small molecules. (a) Calicheamicin is an enediyne-containing oligosaccharide that upon recognition of sequences such as TCCT, TCTC, and TTTT, causes a double-strand cleavage. (b) Hoechst 33258 recognizes A·T, T·A domains selectively. (c) Distamycin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

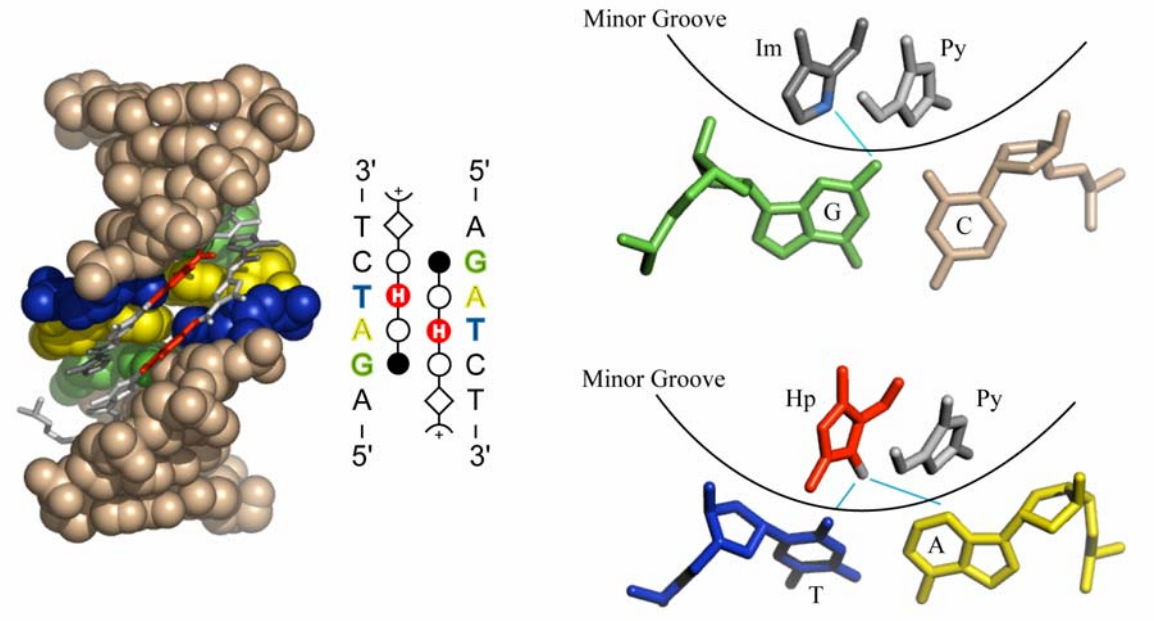


Figure 1.5 Minor groove recognition by polyamides. An X-ray crystal structure of a DNA helix bound by homodimeric polyamides.[26] The polyamides and the DNA base pairs they recognize are represented in the ball-and-stick model. A black circle represents Im, a white circle represents Py, a red circle 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

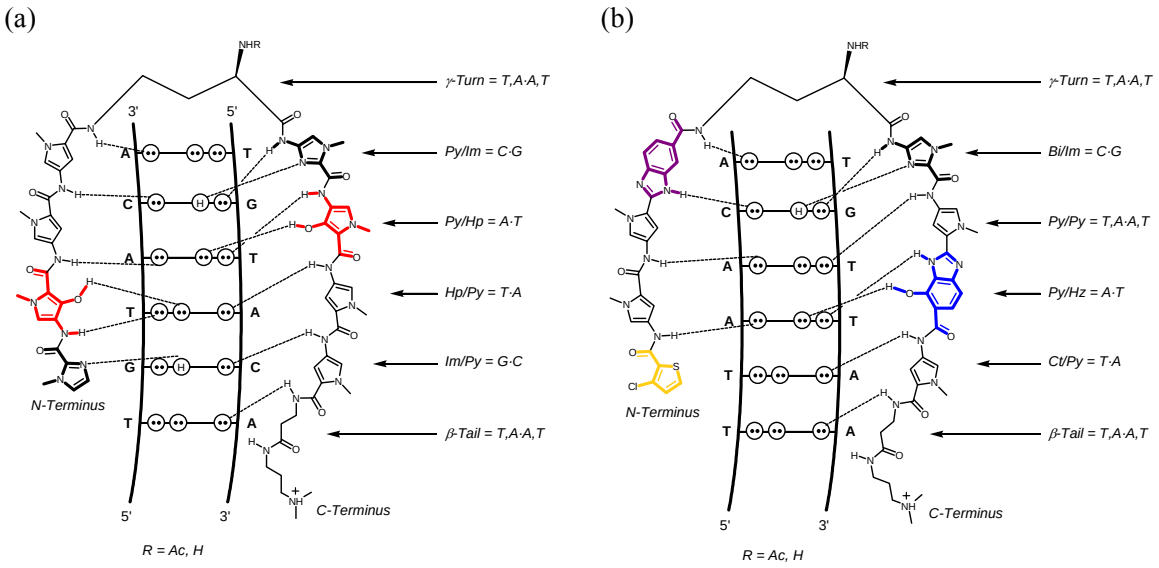


Figure 1.6 Hydrogen-bonding model of the hairpin motif. (a) H-bonding and pairing rules for classic Im/Py/Hp polyamides. (b) H-bonding and pairing rules for hairpin polyamides incorporating new aromatic residues Ct, Bi, and Hz. Im residues are in bold black, Py residues are in regular black, Hp residues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

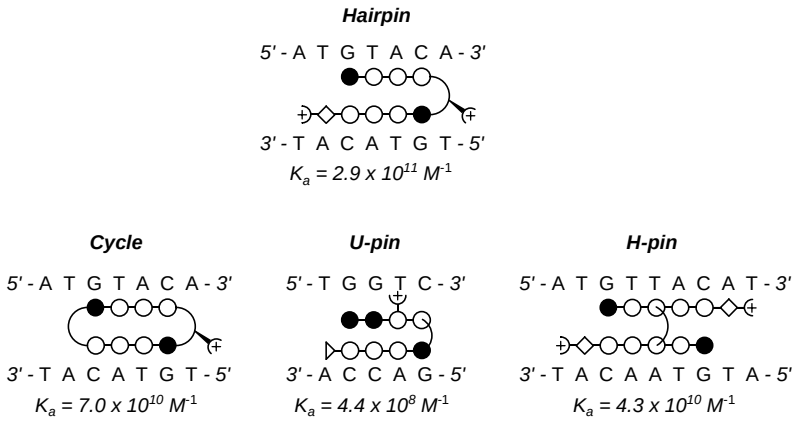


Figure 1.7 Ball-and-stick schematic for covalently linked polyamide motifs.

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.

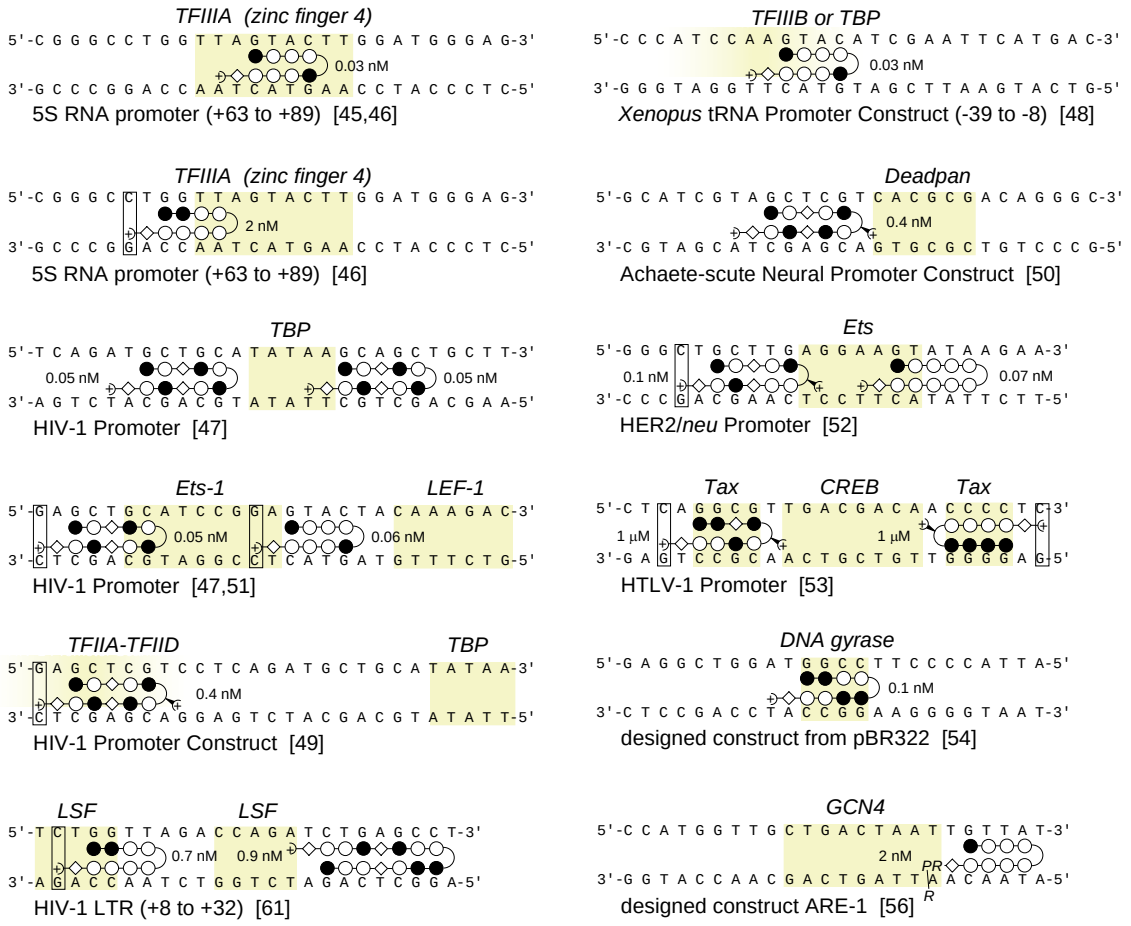


Figure 1.8 Hairpin polyamides have successfully inhibited the binding of several proteins to DNA in solution. Transcription factor binding sites are shaded.

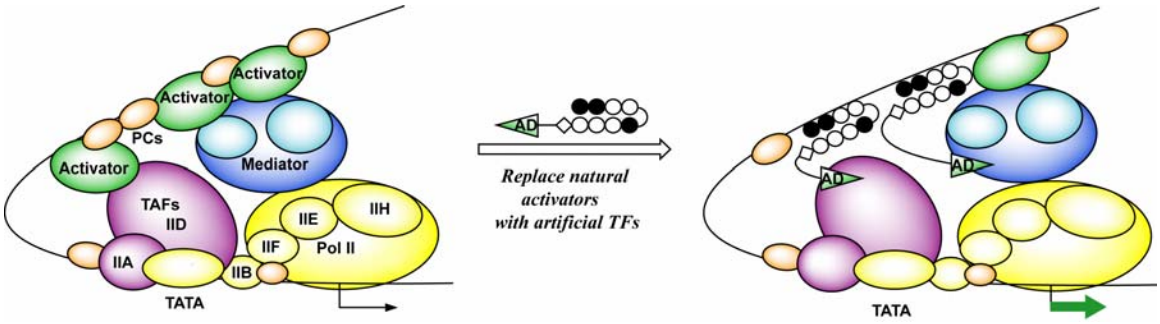


Figure 1.9 Polyamide-peptide conjugates have successfully been used to activate transcription of a target gene *in vitro*. In this experiment, the natural transcription factor (TF) activators (green ovals) have been replaced by polyamide-activator conjugates, binding the appropriate gene promoter sites. 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

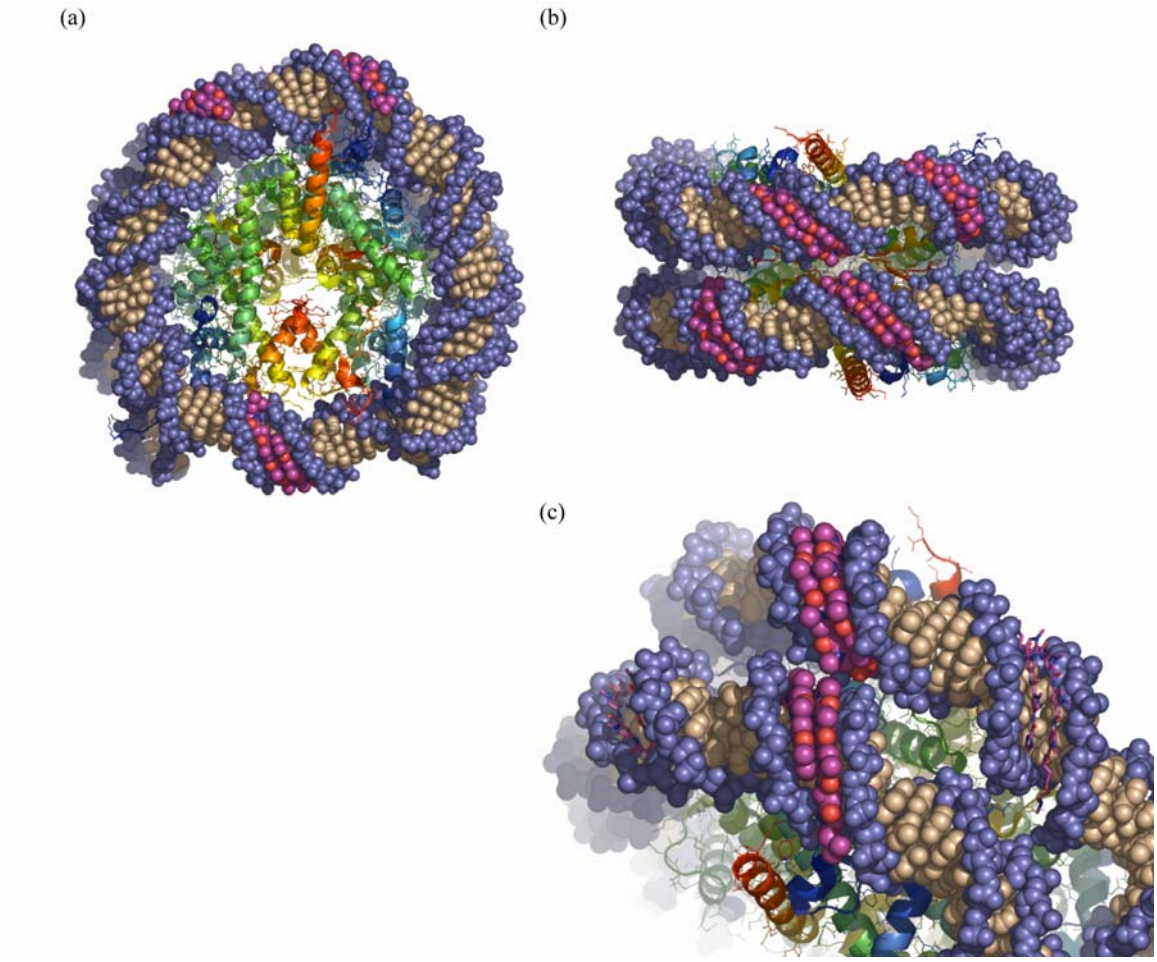


Figure 1.10 X-ray crystal structure of a hairpin polyamide bound to the nucleosome core particle (NCP). Top-down (a) and side-on (b) views of several polyamide molecules bound to the NCP. Histone proteins are depicted as multicolored cartoons. (c) View down the minor "supergroove" in which are bound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

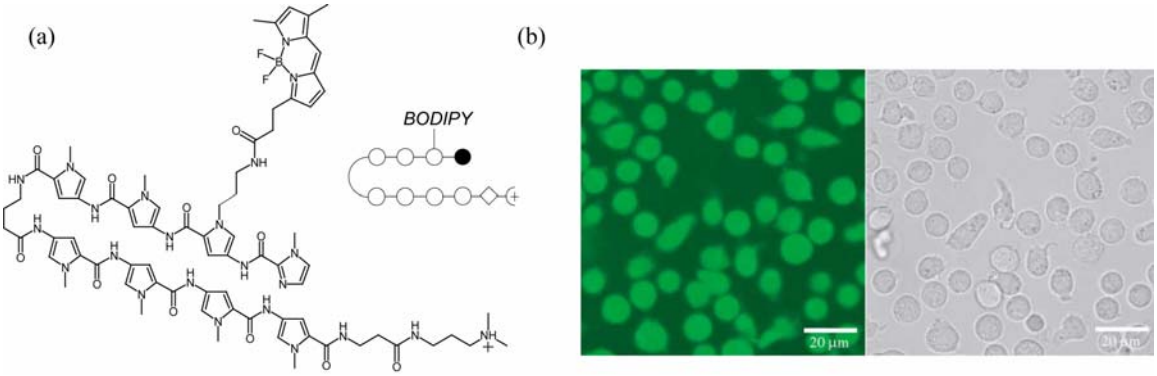


Figure 1.11 Cellular uptake of a fluorescent polyamide. (a) Chemical structure and ball-and-stick representations of a hairpin polyamide labeled with the fluorophore Bodipy. (b) CEM cells stained with Bodipy-polyamide. At left is a fluorescence-only image, and at right is the bright-field image. Polyamide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

Chapter 2

Transcription Activation with Polyamide-Polyproline-Peptide Conjugates

The text of this chapter was taken in part from a manuscript coauthored with Paramjit S. Arora, Aseem Z. Ansari, Professor Mark Ptashne (Sloan-Kettering Institute), and Professor Peter B. Dervan (Caltech).

(Arora, P.S.; Ansari, A.Z.; Best, T.P.; Ptashne, M.; Dervan, P.B. "Design of Artificial Transcriptional Activators with Rigid Poly-*L*-proline Linkers" *J. Am. Chem. Soc*. **2002**, *124*, 13067.)

Abstract

Typical eukaryotic transcriptional activators are composed of distinct functional domains, including a DNA-binding domain and an activating domain. Artificial transcription factors have been designed wherein the DNA-binding domain is a minor groove DNA-binding hairpin polyamide linked by a flexible tether to short activating peptides, typically 16-20 residues in size. In this study, the linker between the polyamide and the peptide was altered in an incremental fashion using rigid oligoproline "molecular rulers" in the 18-45 Å length range. We find that there is an optimal linker length which separates the DNA and the activation region for transcription activation.

Introduction

Transcriptional activators typically bind near a gene and recruit the transcriptional machinery to a nearby promoter, thereby stimulating the expression of the gene. These activators comprise two regions: the DNA binding domain and the activating domain. The former defines the promoter address in the genome where the activating region is to be delivered for recruitment of the transcriptional machinery.[1,2] We have previously reported efforts to replace the natural protein activators with nonnatural components smaller in size than nature's proteins. Artificial transcription activators comprised of three synthetic modules, a hairpin polyamide (PA) DNA-binding domain (DBD) and a short peptide activation domain (AD) (typically 16-20 amino acid residues in size) connected by flexible linkers which vary in length, have been shown to initiate transcription at targeted promoter sites in cell-free systems.[3,4] Modeling studies on these polyamide-peptide conjugates with flexible linkers suggested a distance of 20-40 Å between the DNA and the activating region for transcription activation. We sought to determine whether the length of the linker that projected the activating region away from the DNA would have any effect on the degree of activation (Figure 2.1). In the present study, we replaced the flexible linkers between the DBD and AD with rigid oligoprolines of varying incremental lengths (18-45 Å). Our results suggest that there is an optimal window within which activating regions of the type used here can function efficiently.

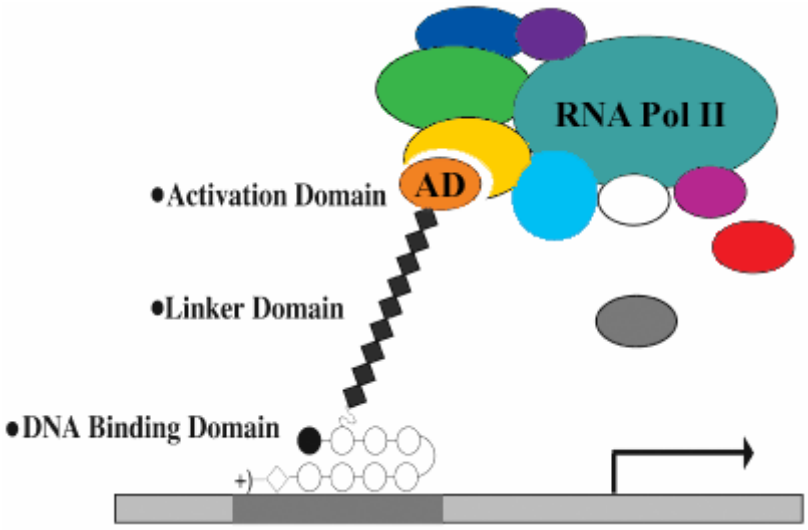
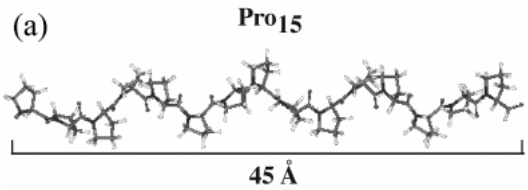


Figure 2.1 Activation of gene transcription by artificial transcription factors. The artificial activator is composed of three separate functional domains. The DNA binding domain consists of minor groove-binding pyrrole/imidazole polyamides. The DNA-binding domain is tethered to the activation domain (AD), a peptide, by a linker domain.

Results

In this study, we replaced the flexible linker between the activating peptide and the hairpin polyamides with 6, 9, 12, and 15 *L*-proline residues (Pro_6-Pro_{15}) (Figure 2.2). Poly-*L*-proline linkers were chosen as the "molecular rulers" for the present study since a stretch of proline residues forms a stable helical structure (the polyproline II helix). Addition of each proline residue increases the length of this helix in a predictable manner, approximately 3 Å per proline residue. Thus, the oligoproline linker projects the activating region peptide away from the DNA-binding polyamide in an incremental manner spanning 18, 27, 36, and 45 Å (Figure 2.2).[5] For the activating region we used two peptides, AH and VP2, that have been previously shown to function efficiently when tethered to a hairpin polyamide (Figure 2.3a).[4] Therefore, three series were synthesized:



(b)

Pro_n	predicted polyproline helix size (Å)
6	18
9	27
12	36
15	45

Figure 2.2 Poly-*L*-proline-based molecular rulers. (a) Structure of a poly-*L*-proline helix. (b) Predicted length of polyproline helices containing the indicated numbers of proline residues.

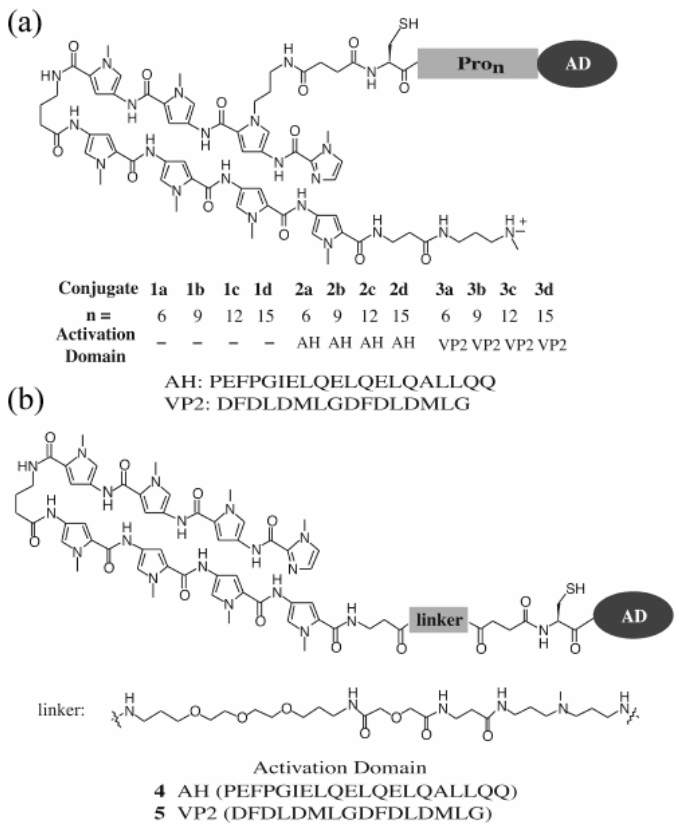


Figure 2.3 Structures of polyamide-activation peptide conjugates **1**-**5** prepared to explore the effect of linker length and flexibility on activation potential. (a) Polyamide-Pro_n-peptide conjugates for which the linker domain consisting of poly-*L*-proline helices serves as a molecular ruler. (b) Polyamide-peptide conjugates bearing a flexible linker.

PAPro$_{6\text{-}15}$-AH (**2a-d**), PA-Pro$_{6\text{-}15}$-VP2 (**3a-d**), and PA-Pro$_{6\text{-}15}$ (**1a-d**) lacking the activation peptides as controls. Conjugates with flexible poly(ethylene glycol) linkers, **4** and **5**, previously shown to activate transcription in vitro were included for comparison (Figure 2.3b).[4]

The eight-ring hairpin polyamides which target 5'-WGWWWW-3' (W = A or T) and two peptides, AH and VP2, were synthesized by solid-phase methods.[6] Each activating peptide contained an N-terminal cysteine for subsequent reaction with a thioester, **6**, via the native ligation reaction to afford the desired conjugate (Figure 2.4).[7] The ability of these conjugates, **1a-d**, **2a-d**, and **3a-d**, to activate transcription *in vitro* was tested using yeast nuclear extract on a template bearing three palindromic binding sites upstream of the minimal core promoter driving the expression of a transcript that

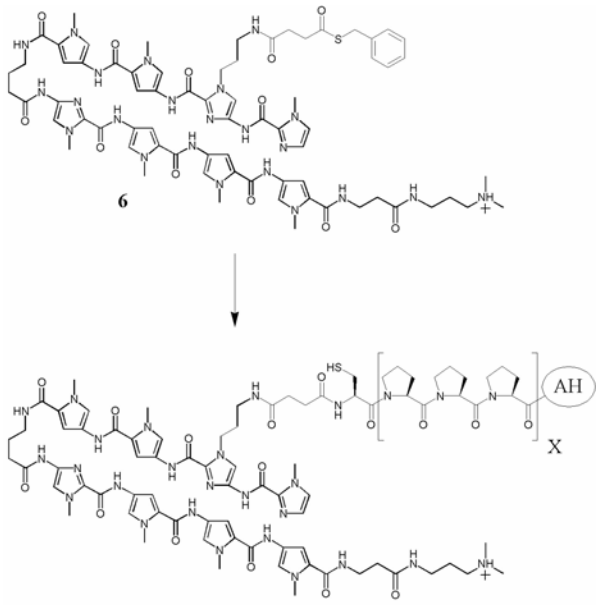


Figure 2.4 Synthesis of polyamide-peptide conjugates. Thiolane **6** was converted to the targeted conjugates via the native ligation reaction.

lacks guanine residues.[3,4] The results show that the linker plays a role in determining the ability of the activating region to stimulate *in vitro* transcription (Figure 2.5). This effect is not a function of the activating peptide, as both activating peptides tested showed similar profiles. Figure 2.5 also compares the two activating peptides attached to the polyamide via different proline linkers to conjugates **4** and **5** bearing the same activating peptides via flexible linkers.

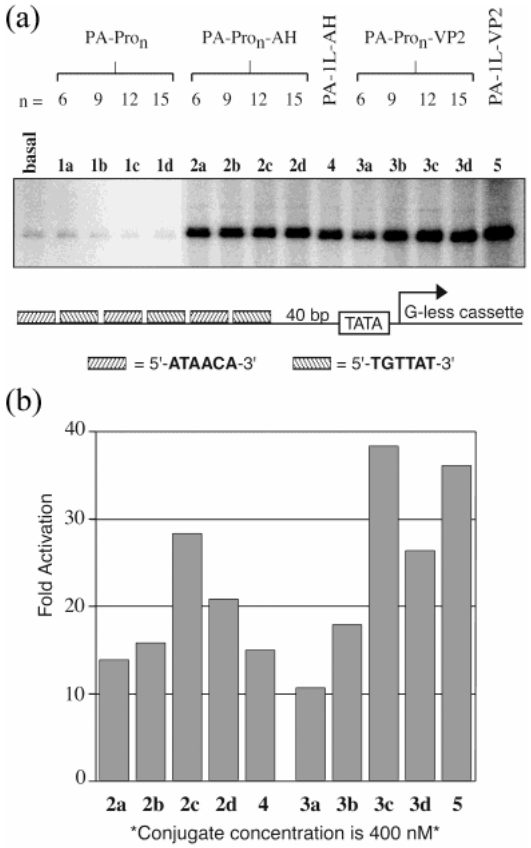


Figure 2.5 *In vitro* transcription reactions with the compounds listed in Figure 2.3; the transcription reaction conditions are described in the Experimental Section. (a) Storage phosphor autoradiogram of the reactions which were performed with a 400 nM concentration of each conjugate. The template configuration showing the sequence and number of the polyamide binding sites is depicted below the gel. (b) Fold activation was determined by comparing the amount of transcription elicited by conjugates **2a-d**, **3a-d**, **4**, and **5** with that of the basal level.

Our results indicate that the strength of the given activating peptide increases with every increment in the spacer length up to 12 residues (36 Å). An increase in the spacer length to 15 residues leads to a decrease in activation. These results suggest an optimal spacing of 36-45 Å between the DNA and the activating regions for efficient transcription. Figure 2.6 shows the relative potency of each compound in comparison to PA-Pro_6-AD (**2a**, AD = AH; **3a**, AD = VP2). Although the absolute fold activation mediated by all conjugates varies slightly between different experiments, the overall relationship between linker length and activation potential remains constant over four independent *in vitro* transcription reactions.

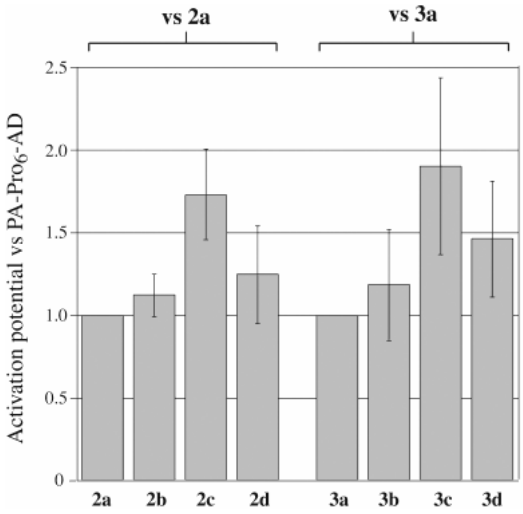


Figure 2.6 Summary of four independent *in vitro* transcription reactions showing the relative potency of each compound in comparison to PA-Pro_6-AD (**2a**, AD = AH; **3a**, AD = VP2).

To determine whether any of the effects observed were due to inefficient or inappropriate binding of the conjugate to its site on the promoter, we measured dissociation constants for conjugates **3a-d** using quantitative DNase I footprinting titrations.[8,9] These studies revealed that conjugates bound with similar affinities and specificities to their sites on the promoter DNA (Figure 2.7 and Table 2.1). In control

experiments polyamide-proline conjugates **1a-d** lacking the activation transcription peptides did not activate.

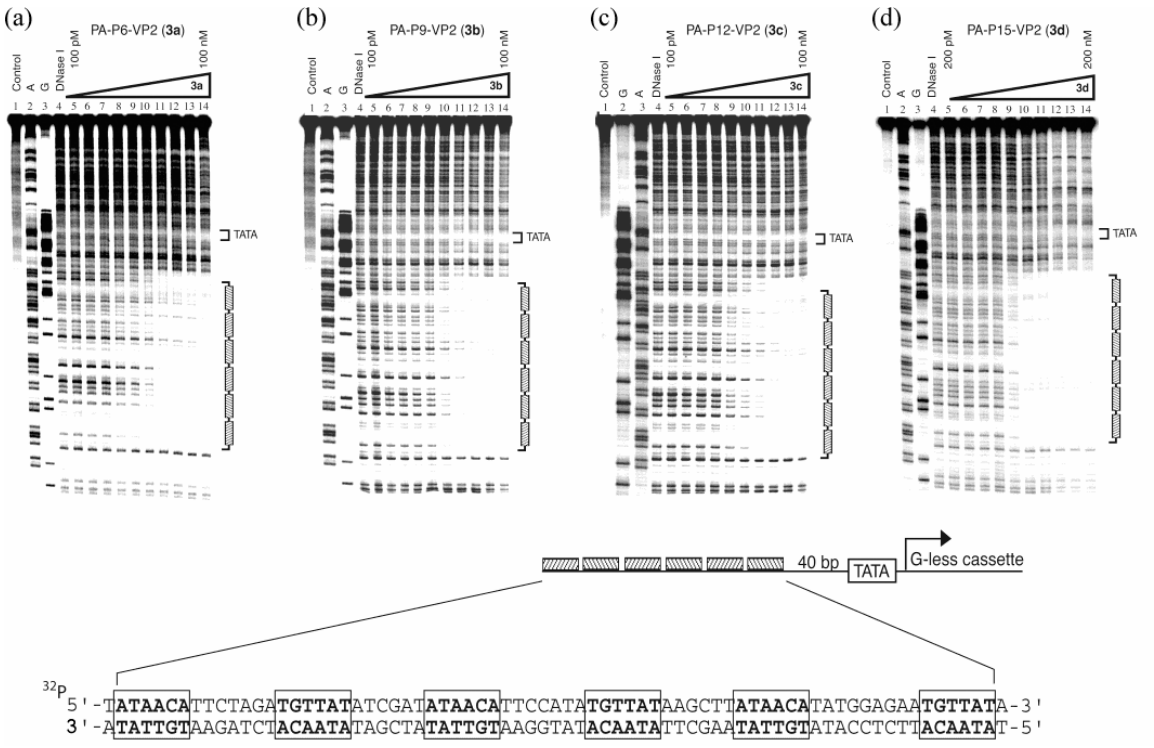


Figure 2.7 Quantitative DNase I footprinting titration of conjugates **3a-d** shows that all conjugates bound to their target sites with similar affinities. (Top) Storage phosphor autoradiogram of a quantitative DNase I footprinting titration of **3a-d** on a 63 bp 5'-^{32}P-labeled PCR fragment containing both the promoter region and 140 bp of the G-less cassette reporter. (a) Lane 1, undigested DNA; lane 2, A reaction; lane 3, G reaction; lane 4, DNase I standard; lanes 5-14, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, and 100 nM **3a**, respectively. (b) Lane 1, undigested DNA; lane 2, A reaction; lane 3, G reaction, lane 4, DNase I standard; lanes 5-14, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, and 100 nM **3b**, respectively. (c) Lane 1, undigested DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-14, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, and 100 nM **3c**, respectively. (d) Lane 1, undigested DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-14, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, and 200 nM **3d**, respectively. (Bottom) Design of the plasmid DNA template used for the DNase I footprinting titration experiments. The promoter region contains six cognate binding sites for the hairpin polyamide upstream of a G-less cassette reporter.

Table 2.1 Equilibrium Dissociation Constants (nM)

conjugate	K_D (nM)	conjugate	K_D (nM)
PA-Pro_6-VP2 (**3a**)	4.8 (±1.4)	PA-Pro_{15}-VP2 (**3d**)	2.3 (±1.3)
PA-Pro_9-VP2 (**3b**)	2.0 (±1.3)	PA-Pro_{12}-AH (**2c**)	7.7 (±1.7)
PA-Pro_{12}-VP2 (**3c**)	3.3 (±1.1)		

The use of poly-*L*-proline linkers as molecular rulers was inspired by the seminal work of Stryer and co-workers, who utilized fluorescence resonance energy transfer (FRET) analysis to show that oligomers composed of up to 12 proline residues can retain a rigid helical structure.[5] To ascertain that an oligomer composed of 15 proline residues does not deviate from the predicted structure, we repeated the FRET analysis on poly-*L*-proline linkers used in our experiments. To perform this analysis, we synthesized poly-*L*-proline linkers bearing Oregon Green (OG) energy donor and tetramethylrhodamine (TAMRA) energy acceptor moieties on each end (conjugates **7a-d**, Figure 2.8). In the Stryer study, naphthyl and dansyl groups were used as the energy donor and the energy acceptor moieties, respectively. The different dyes were used in the present study because the Förster radius (the distance at which energy transfer is 50% efficient) for the dansyl-naphthyl pair (R_o = 27 Å) may be too low to allow for efficient FRET when the energy donor and the acceptor are spaced by the Pro_{15} linker (45 Å). The Förster radius (R_o) for the OG-TAMRA pair is predicted to be 55 Å.[10] FRET measurements indicate that the oligomer composed of 15 proline units does retain a rigid structure. The plot of distance versus FRET efficiency provides a linear relationship between the conjugates, and importantly, the experimentally observed FRET efficiency is in good agreement with the predicted r^{-6} dependence (Figure 2.9).[5] Thus, the FRET studies show that the

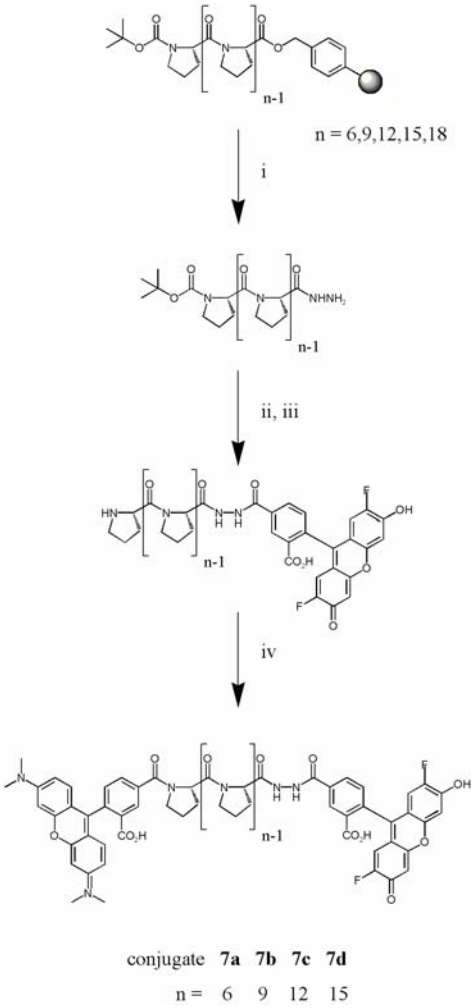


Figure 2.8 Synthesis of polyproline-dye conjugates. (i) 20% hydrazine in ethanol. (ii) Oregon Green-NHS ester, DMF, and DIEA. (iii) 20% TFA in CH_2Cl_2. (iv) TAMRA-NHS ester, DMF, and DIEA.

dependence of transcription activation on the number of *L*-prolines is not due to a change in the linker structure.

Discussion

In this study we examine the role of the distance separating the DBD and AD in our artificial transcription factors. This study was prompted by previous observations in

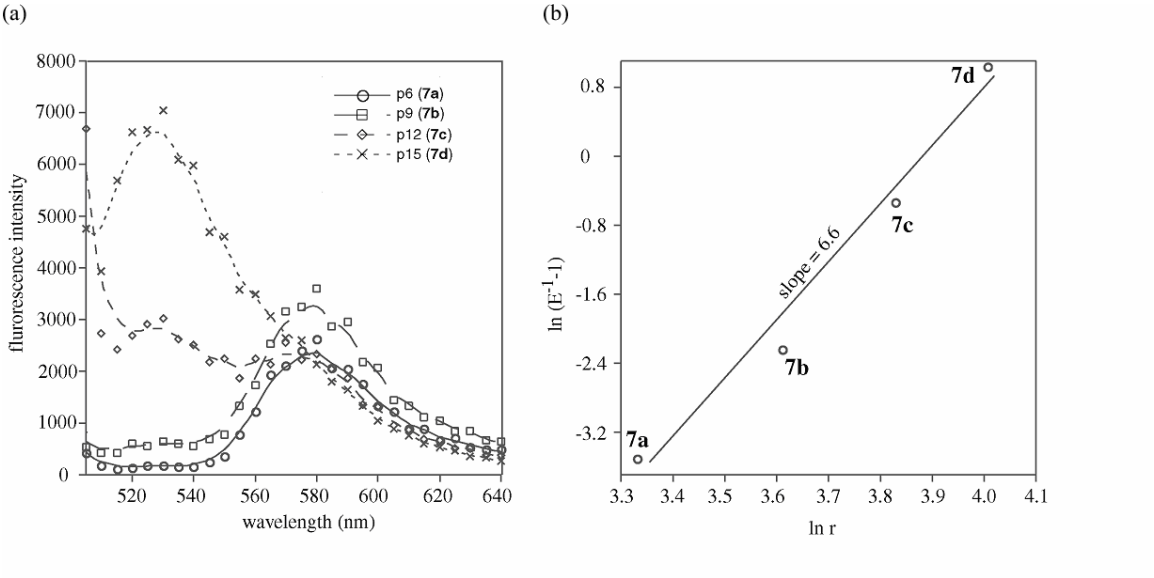


Figure 2.9 FRET data for polyproline helices. (a) The emission spectra of OG-Pro_n-TAMRA conjugates. The conjugates were excited at 495 nm. (b) The dependence of the efficiency of energy transfer on distance is given in this plot of ln (E^{-1} – 1) vs. ln r. The slope is 6.6, in good agreement with the predicted r^{-6} dependence.

which we found that activating regions activated transcription to different degrees on the basis of the point of attachment as well as the nature of the linker separating them from the DNA binding polyamide.[4] This dependence on spatial presentation led us to test the role of linker length in the degree of activation elicited by two acidic activating peptides. The proline linker domain is conjugated to an internal pyrrole of the polyamide, rather than to the C-terminus, which allows the linkers to project out of the minor groove away from the DNA helix. Our data suggest that for the two acidic activating regions (AH and VP2) the strength of activation increases as the linker reaches a length of 36 Å. The increase, while modest, is reproducible and displays a clear trend (Figures 2.5 and 2.6). The effects are muted in part due to three features that were incorporated into the experimental design: the use of multiple DNA binding sites to elicit robust activation, the presence of a flexible hinge tethering the proline linker to the polyamide, and the

unstructured nature of the activating peptides, which themselves may sample a significant amount of solvent space when fully extended.

We did not find a significant contribution of poly-*L*-proline linkers themselves to the level of activation in the absence of tethered activating regions. It has been reported that proline-rich (rather than poly-*L*-proline) activating domains can function to activate transcription even in yeast extracts, though this activation is not as robust as that elicited by acidic activators.[11] Presumably the proline oligomer is not a sufficiently strong activator; its effects are therefore not detected in our studies. An alternative possibility is that the ability of proline-rich activating domains to elicit transcription may be more sensitive to their spatial location.

As the principle of recruitment implies, for a DNA-tethered activator to function efficiently, it must sample sufficient nuclear solvent space to bind and recruit various complexes that participate in transcriptional initiation to a given promoter. However, beyond a certain distance the activating region would not function to recruit, as the local concentration of the machinery at a given promoter would not be tremendously enhanced by binding a distant activating region. It has been shown that eukaryotic activating regions that are not tethered to DNA—even when presented at exceedingly high concentrations such that it would bind to its targets in the transcriptional machinery—do not improve the level of transcription from a given promoter. In fact, as expected, it "squelches" the ability of a DNA-bound activator from functioning—presumably by binding the targets in the machinery.[2,3,12] Thus, for efficient recruitment in the context of our *in vitro* studies with artificial activators on nonchromatinized templates using yeast

extracts, we find that a spacing of 36-45 Å serves as an optimal distance between the DNA and the activating regions to elicit transcriptional activation.

This work is a step forward toward the goal of engineering and integrating at the molecular level the components for functional artificial transcription factors. The field is at an early stage, and one can imagine other components replacing the design reported here, such as triple-helix-forming oligonucleotides or PNAs tethered to activating peptides.[13,14] Long-term goals for the field would be the replacement of the activating peptide with nonpeptide constructs and the activation (or repression) of endogenous genes in cell culture experiments.

Experimental Section

Synthesis of Polyamide-Peptide Conjugates 1-3

Polyamide thioester **6** was transformed into conjugates **1a-d**, **2a-d**, and **3a-d** by previously reported methods.[7] The identities of all conjugates were verified by MALDI-TOF mass spectrometry. Characterization: **1a** (PA-Pro_6), MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2050.6, obsd 2050.0; **1b** (PA-Pro_9), MALDI-TOF $[M+H]^+$ (average mass) calcd 2341.9, obsd 2341.4; **1c** (PA-Pro_{12}), MALDI-TOF $[M+H]^+$ (average mass) calcd 2633.3, obsd 2633.1; **1d** (PA-Pro_{15}), MALDI-TOF $[M+H]^+$ (average mass) calcd 2924.6, obsd 2925.1; **2a** (PAPro_6-AH), MALDI-TOF $[M+H]^+$ (average mass) calcd 4354.5, obsd 4354.8; **2b** (PA-Pro_9-AH), MALDI-TOF $[M+H]^+$ (average mass) calcd 4646.8, obsd 4647.5; **2c** (PA-Pro_{12}-AH), MALDI-TOF $[M-H]^-$ (average mass) calcd 4937.5, obsd 4937.5; **2d** (PA-Pro_{15}-AH), MALDI-TOF $[M+H]^+$ (average mass) calcd 5228.7, obsd 5229.2; **3a** (PA-Pro_6-VP2), MALDI-TOF $[M+H]^+$ (average mass) calcd

3862.3, obsd 3862.7; **3b** (PA-Pro_9-VP2), MALDI-TOF $[M+H]^+$ (average mass) calcd 4153.7, obsd 4154.8; **3c** (PA-Pro_{12}-VP2), MALDI-TOF $[M-H]^{+-}$ (average mass) calcd 4445.0, obsd 4445.3; **3d** (PA-Pro_{15}-VP2), MALDI-TOF $[M+H]^+$ (average mass) calcd 4736.3, obsd 4736.8.

Synthesis of Poly-*L*-proline -Dye Conjugates 7a-d

Poly-*L*-proline peptides of lengths *n* proline residues, where *n* = 6, 9, 12, 15, were individually synthesized by standard *t*-Boc solid phase methods, retaining the n-terminal *t*-Boc protecting group. All peptides were cleaved from resin by treatment with 20% (v/v) anhydrous hydrazine in absolute ethanol for 18 hrs at room temperature. Peptides were purified by reverse-phase HPLC, and lyophilized to provide white powders. Purified peptides (2.5 μmoles) were treated with the 5-*N*-hydrosuccinimidyl ester of Oregon Green 488 (OG, Molecular Probes) (5.0 μmoles), DMF (100 μL), and *N*,*N*-diisopropylethylamine (20 μL), and purified by reverse-phase HPLC to yield C-terminal dye intermediates. The N-terminal *t*-Boc group was removed by treatment with 20% TFA in CH_2Cl_2 and the resulting molecules were treated with the 5-*N*-hydrosuccinimidyl ester of tetramethylrhodamine (TAMRA, Molecular Probes) (5.0 μmoles), DMF (100 μL) and *N*,*N*-diisopropylethylamine (20 μL). Purification by reverse-phase HPLC afforded dye conjugates **7a-e** (typical isolated yields: 20-40%). The identity of all conjugates was verified by MALDI-TOF mass spectrometry. Characterization: **7a** (TAMRA-Pro_6-OG): MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1422.5, obsd 1422.6; **7b** (TAMRA-Pro_9-OG): MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1712.7, obsd 1712.5; **7c** (TAMRA-Pro_{12}-OG): MALDI-TOF $[M+H]^+$ (monoisotopic

mass) calcd 2003.8, obsd 2003.9; **7d** (TAMRA-Pro_{15}-OG): MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2295.0, obsd 2295.2.

Fluorescence Resonance Energy Transfer Experiments

Fluorescence emission spectra were collected on an ISS-K2 fluorometer. Solutions (25 nM) of conjugates **7a-d**, Oregon Green, and TAMRA were prepared in 1:1 methanol:bicarbonate buffer (150 mM, pH 8.3). The emission spectra of these solutions were collected from 500-640 nm, with an excitation wavelength of 495 nm.

DNase I Footprinting Titration Experiments

A 363 bp 5' ^{32}P-labeled PCR fragment was generated from template plasmid pAZA812 in accordance with standard protocols and isolated by nondenaturing gel electrophoresis. All DNase I footprinting reactions were carried out in a volume of 400 μL. A polyamide stock solution or water (for reference lanes) was added to TKMC buffer, with final concentrations of 50 mM Tris-HCl, 50 mM KCl, 50 mM $MgCl_2$, and 25 mM $CaCl_2$, pH 7.0, and 15 kcpm 5'-radiolabeled DNA. The solutions were equilibrated for 12-18 h at 22°C. Cleavage was initiated by the addition of 10 μL of a DNase I stock solution and was allowed to proceed for 7 min at 22°C. The reactions were stopped by adding 50 μL of a solution containing 2.25 M NaCl, 150 mM EDTA, 0.6 mg/mL glycogen, and 30 μM base pair calf thymus DNA and then ethanolprecipitated. The cleavage products were resuspended in 100 mM Trisborate-EDTA/80% formamide loading buffer, denatured at 85°C for 10 min, and immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V for 2 h and 15 min.

The gels were dried under vacuum at 80 °C and quantitated using storage phosphor technology.

In Vitro Transcription Assays

Template plasmid pAZA812 was constructed by cloning a 78 bp oligomer bearing three cognate palindromic sequences for conjugates **1a-d**, **2a-d**, and **3a-d** into a *Bgl*2 site 30 bp upstream of the TATA box of pMLΔ53. This plasmid has the AdML TATA box 30 bp upstream of a 277 bp G-less cassette. For each reaction, 20 ng of plasmid (30 fmol of palindromic sites) was preincubated with a 400 nM concentration of the compound for 75 min prior to the addition of 90 ng of yeast nuclear extract in a 25 μL reaction volume. The reactions were performed as previously described and resolved on 8% 30:1 polyacrylamide gels containing 8 M urea. The gels were dried and exposed to photostimulatable phosphorimaging plates (Fuji Photo Film Co.). The data were visualized using a Fuji phosphorimager followed by quantitation using MacBAS software (Fuji Photo Film Co.).

Acknowledgement

We are grateful to the National Institutes of Health for research support, the American Cancer Society for a postdoctoral fellowship to P.S.A., and the Helen Hay Whitney Foundation for a fellowship to A.Z.A.

References

55) Ptashne, M.; Gann, A. *Nature* **1997**, *386*, 569.

56) Ptashne, M.; Gann, A. *Genes and Signals*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2001.

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

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

59) Stryer, L.; Haugland, R.P.; *Biochemistry* **1967**, *6*, 719.

60) Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1996**, *118*, 6141.

61) Mapp, A.K.; Dervan, P.B. *Tetrahedron Lett*. **2000**, *41*, 9451.

62) Senear, D.F.; Brenowitz, M.; Shea, M.A.; Ackers, G.K. *Biochemistry* **1986**, *25*, 7344.

63) Brenowitz, M; Senear, D.F.; Shea, M.A.; Ackers, G.K. *Methods Enzymol*. **1986**, *130*, 132.

64) Haugland, R.P. *Handbook of Fluorescent Probes and Research Chemicals*; 6th ed.; Molecular Probes Inc.: Eugene, OR, 1996.

65) Kim, T.K.; Roeder, R.G. *J. Biol. Chem*. **1993**, *268*, 20866.

66) Gill, G.; Ptashne, M. *Nature* **1988**, *334*, 721.

67) (a) Kuznetsova, S.; Ait-Si-Ali, S.; Nagibneva, I.; Troalan, F.; Le Villain, J.-P.; Harel-Bellan, A.; Svinarchuk, F. *Nucleic Acids Res*. **1999**, *27*, 3995. (b) Stanojevic, D.; Young, R.A. *Biochemistry* **2002**, *41*, 7209.

68) (a) Denison, C.; Kodadek, T. *Chem. Biol*. **1998**, *5*, 129. (b) Liu, B.; Han, Y.; Corey, D.R.; Kodadek, T. *J. Am. Chem. Soc*. **2002**, *124*, 1838.

Chapter 3

Polyamide-Peptide and Polyamide-Small Molecule Conjugates for Cellular Uptake Studies

Chapter 3A

Polyamide-Peptide Conjugates for Cellular Uptake Studies

Abstract

In order that polyamides or polyamide-peptide conjugates be useful as diagnostic or therapeutic tools in living organisms, they must be able to cross the cellular membrane of cells within target tissues and transit to the nuclei of those cells. In proof-of-principle *in vitro* experiments polyamides were shown to inhibit the transcription of target genes, presumably by competing against natural transcription factors for promoter binding sites. In other *in vitro* experiments polyamide-peptide conjugates were shown to activate transcription of target genes, presumably by recruitment of the RNA polymerase II holoenzyme to promoter regions. Attempts to duplicate these successes in living model systems met with limited success.

It was postulated that the difficulty in transferring successful *in vitro* results to living systems lay in the inability of polyamides to cross cell membranes. In order to overcome this challenge, polyamides were conjugated to peptides shown to increase cellular uptake and to small molecules that might be recognized by cellular receptors, facilitating polyamide influx into the cell. Polyamides were also added to cells in growth medium containing drugs shown to induce cellular uptake of certain compounds. Though meeting with little initial success, these attempts increased the body of knowledge of how polyamide conjugates interact with DNA, how they interact with cells, and laid the groundwork for more successful experiments documented in Chapter 4.

Introduction

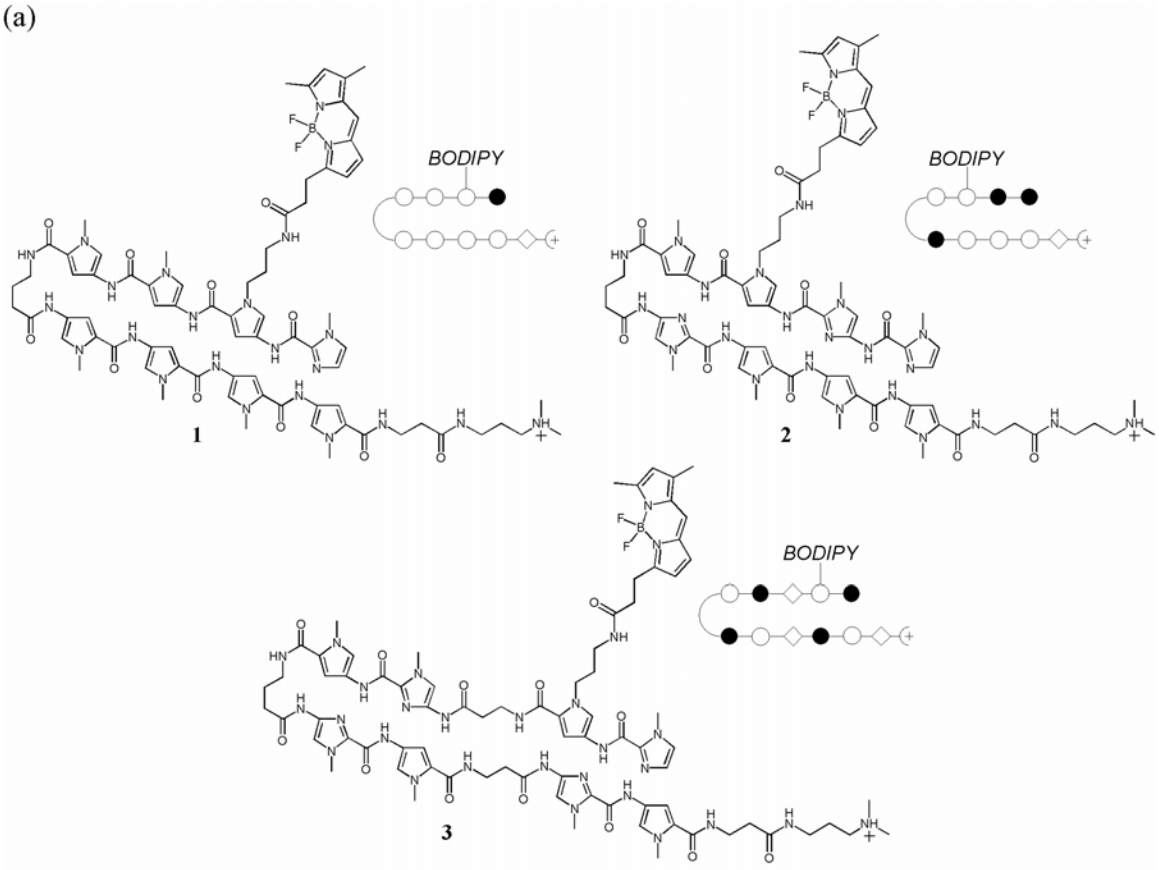
Polyamides are inherently interesting as probes of DNA recognition, due to the modularity of their design and their resulting ability to target nucleic acid sequences in a programmed manner. These molecular tools have been used effectively in several *in vitro* systems to regulate gene processing. As inhibitors of transcription, polyamides have been quite effective in a number of model systems (Figure 1.8). Once conjugated to activation peptides, polyamides have also functioned as transcriptional activators (Chapter 2). Ultimately it is desirable that polyamides have the capacity to affect the transcriptional machinery in cell cultures and living organisms. Studies with the human immunodeficiency virus (HIV) promoter and inhibition of viral replication provided early examples of the successful application of polyamides as potent regulators in primary human lymphocytes[1], as well as in live-cell models.[2] Polyamides were also effective in causing gain- or loss-of-function phenotypes when fed to *Drosophila melanogaster*.[3,4]

Attempts to inhibit the transcription of endogenous genes in cell lines other than insect or T-lymphocytes have met with little success. For example, polyamides targeted to the Ets binding site of the HER2/neu promoter were shown to inhibit binding of the transcription factor ESX at nanomolar concentrations and to inhibit transcription of the HER2/neu gene in cell-free experiments.[5] However, those polyamides display no activity in HER2-overexpressing SKBR-3 cells.[6] In an attempt to determine the reason why no inhibition of transcription was observed in the cell-based assay, fluorescent conjugates of polyamides attached to the Bodipy FL fluorophore were synthesized. These compounds were then applied to SKBR-3 cells and imaged by laser scanning confocal microscopy.

The polyamide conjugates were not seen to localize in the nucleus of the living cells, rather they were sequestered in the cytoplasm.[6]

In order to more generally determine the cellular localization characteristics of polyamides, a small series of polyamide-Bodipy conjugates were synthesized, applied to a variety of cell lines, and imaged by confocal microscopy (Figure 3.1).[7] With the

(a)



(b)

Conjugate	SKBR-3: human breast cancer	CEM: human cultured T-cells	T-cells: human primary T-cells	NB4: human leukemia	293: human kidney fibroblast	Sf9: fall armyworm (insect)	Kc: Drosophila (insect)
1	Cytoplasm	Nuclear	Nuclear	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
2	Cytoplasm	Nuclear	Nuclear	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
3	Cytoplasm	Nuclear	Nuclear	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm

Figure 3.1 Localization of polyamide–Bodipy conjugates in live cells as determined by confocal microscopy. (a) Chemical structures and ball-and-stick models of polyamide-Bodipy conjugates. (b) The designation 'nucleus' indicates observation of fluorescence in the interior of the nucleus. The designation 'cytoplasm' indicates cellular, non-nuclear fluorescence. Cells were imaged directly following 20 h incubation with 5 μM **1**–**3** under normal growth conditions for each cell-line.[7]

exception of human T-cells and T-cell-derived cell lines, polyamide conjugates **1-3** localized in the cytoplasm of all cells tested (Figure 3.2). This result suggested that transcriptional regulation experiments occurring in T-cells or T-cell-derived lines would not be negatively influenced by the cellular uptake of polyamides, while those in other cell lines were prohibited from showing positive results due to poor polyamide uptake.

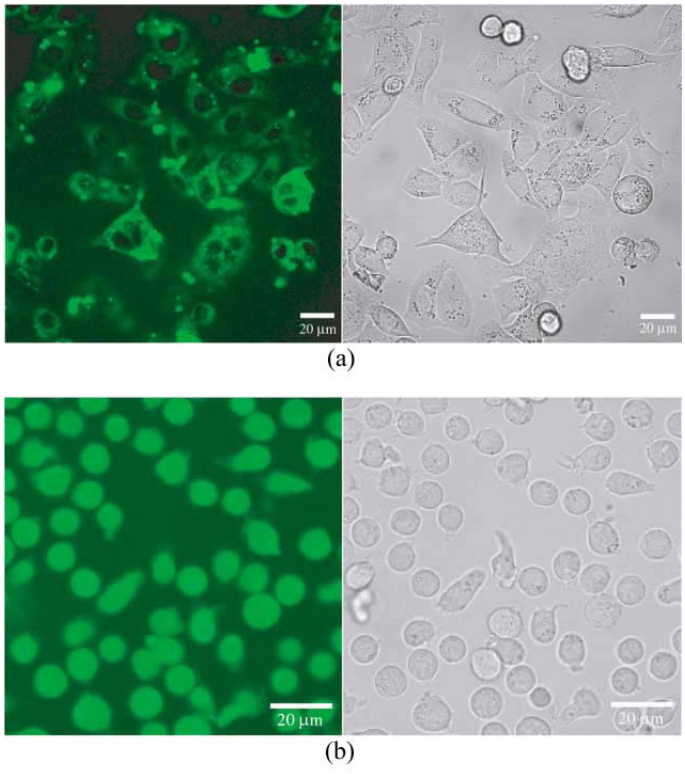


Figure 3.2 Localization of **1** in live human cell lines. (a) Conjugate **1** in live SKBR-3 cells, showing cytoplasmic staining. The fluorescent image is on the left, bright field image on the right. Bar: 20 µm. (b) Conjugate **1** in live CEM cells, showing nuclear staining. The fluorescent image is on the left, bright field image on the right. Bar: 20 µm.[7]

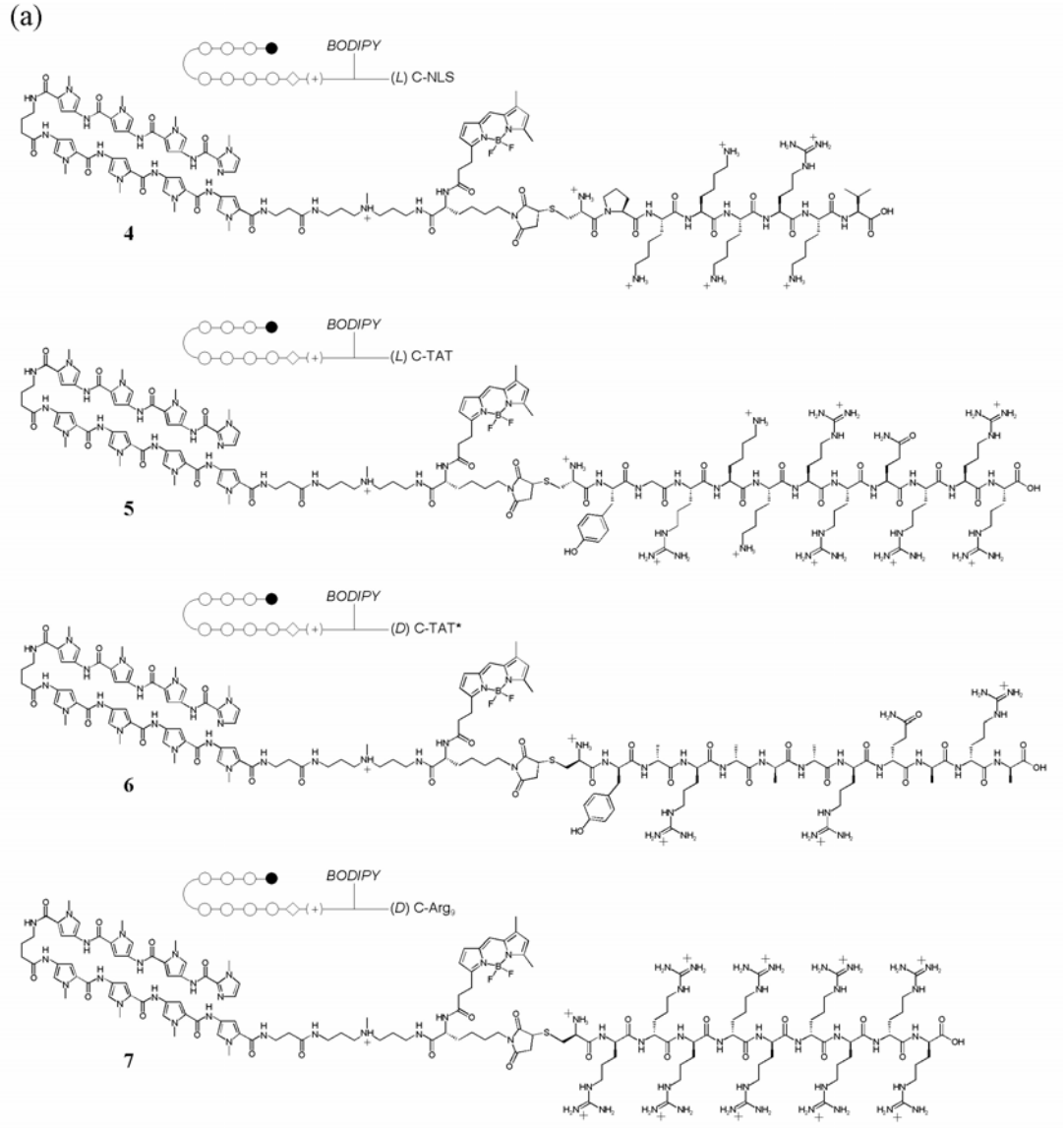
In order to overcome cytoplasmic sequestration of polyamide-fluorophore conjugates, polyamides bearing both a Bodipy fluorophore and one of several carrier peptides were synthesized. Carrier peptides have been used to transport small molecules, proteins, and oligonucleotides into both cells and nuclei.[8] Two of the most widely-

studied carrier peptides are TAT,[9-12] a basic domain derived from the HIV-1 Tat protein,[13] and the simian vacuolating virus-40 (SV-40) nuclear localization sequence (NLS) peptide,[14-17] derived from a basic region from the SV40 T-antigen protein.[18] The NLS is known to be recognized by a specific membrane protein receptor, karyopherin α.[18,19] The mechanism of TAT transduction is unknown, appearing not to involve receptor-, transporter-, endosome-, or adsorptive-endocytosis-mediated processes.[8,20] Since TAT-mediated uptake bypasses known endosomal routes, TAT is known to be a nuclear localization signal,[21] and since polyamide-Bodipy conjugates localize to the cytoplasm in discrete spots, possibly endosomes, TAT-polyamide-Bodipy conjugates should be expected to transit to the nucleus.

Several synthetic derivatives of TAT showing increased ability to facilitate uptake have been developed. Polyarginine peptide Arg_9 was created by Wender and co-workers as an analog of TAT. It was shown that (*D*) Arg_9 outperformed TAT in terms of uptake into Jurkat cells.[22] TAT* was developed by Dowdy and co-workers through alanine scanning to be a helical mimetic of TAT bearing a minimum number of arginine residues.[23] TAT* peptides composed of both *L*- and *D*-amino acids have been shown to be active uptake carriers. Peptides composed of *D*-amino acids are expected to have greater stability to proteolytic degradation than their natural counterparts of *L*-chirality.

Carrier peptide-polyamide-Bodipy conjugates **4**-**7** (Figure 3.3a) were synthesized and tested for their ability to stain the nuclei of several cell lines. TAT-polyamide-Bodipy and (*D*) Arg_9-polyamide-Bodipy conjugates **5** and **7** were observed to stain the nuclei of most cell lines tested, while NLS- and (*D*) TAT*-conjugates **4** and **6** generally were found to localize in the cytoplasm or extracellular medium (Figure 3.3b). It was

(a)



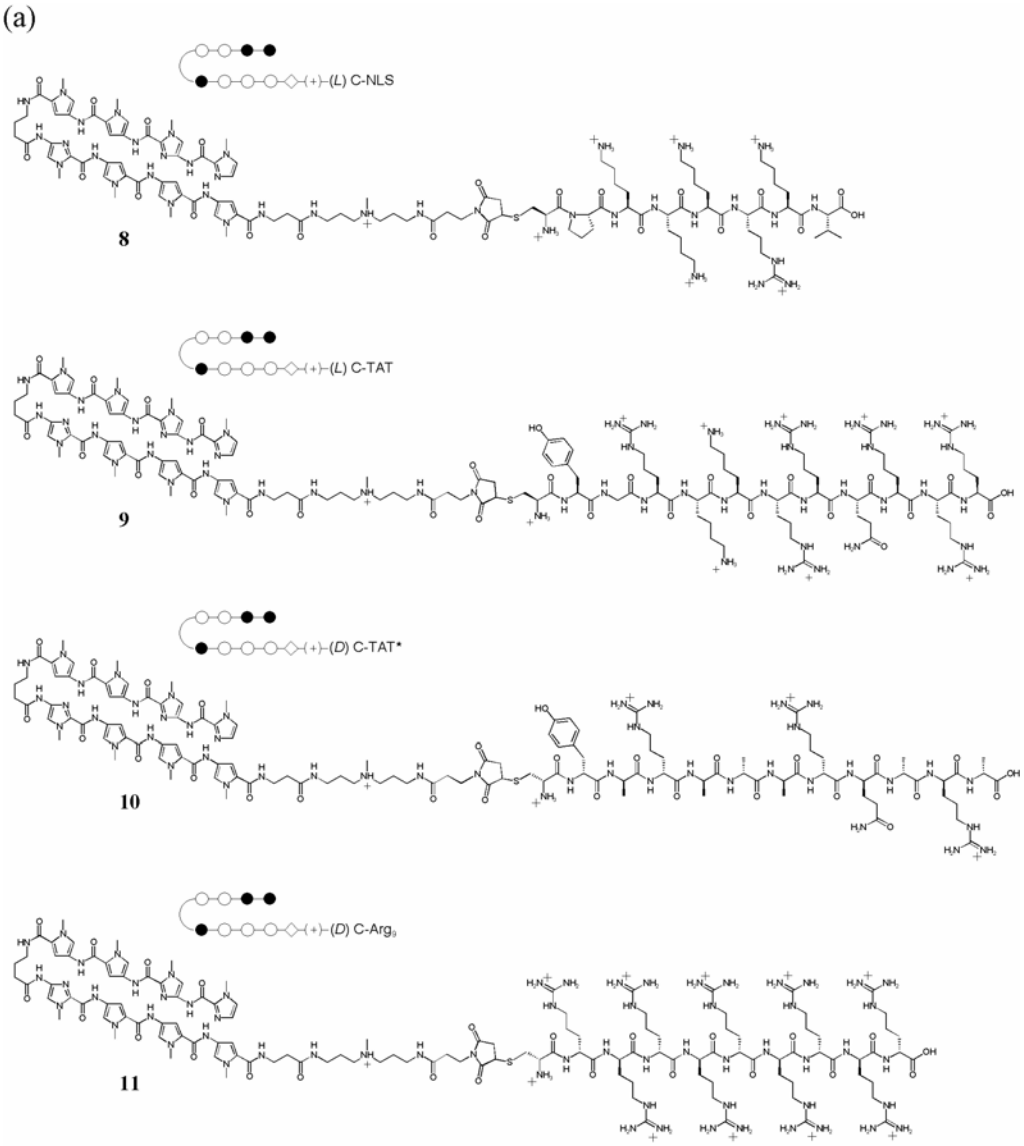
(b)

Conjugate	SKBR-3: human breast cancer	CEM: human cultured T-cells	NB4: human leukemia	293: human kidney fibroblast	Sf9: fall armyworm (insect)	Kc: Drosophila (insect)
4	Cytoplasm	Non-cellular	Non-cellular	Non-cellular	Cytoplasm	Nuclear
5	Cytoplasm	Nuclear	Nuclear	Nuclear	Cytoplasm	Nuclear
6	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Nuclear	Nuclear
7	Nuclear	Nuclear	Nuclear	Nuclear	Nuclear	Nuclear

Figure 3.3 Localization of peptide-polyamide–Bodipy conjugates in live cells as determined by confocal microscopy. (a) Chemical structures and ball-and-stick models of peptide-polyamide-Bodipy conjugates. (b) The designation 'nuclear' indicates observation of fluorescence in the interior of the nucleus. The designation 'cytoplasm' indicates cellular, non-nuclear fluorescence. The designation 'non-cellular' indicates extracellular fluorescence. Cells were imaged directly following 20 h incubation with 5 μM **4**–**7** under normal growth conditions for each cell line.[25]

also found that **5** and **7** tended to be toxic to most cell lines at the concentration used in the assay (5 μM). In order to determine the DNA-binding qualities of polyamide-peptide

(a)



(b)

Conjugate	Match site affinity K_a (M^{-1})	Specificity over coating
8	5×10^8	20-fold
9	$\sim 1 \times 10^9$	None
10	6×10^8	20-fold
11	$\sim 1 \times 10^9$	None

Figure 3.4 DNA-binding characteristics of tail-linked polyamide-peptide conjugates.

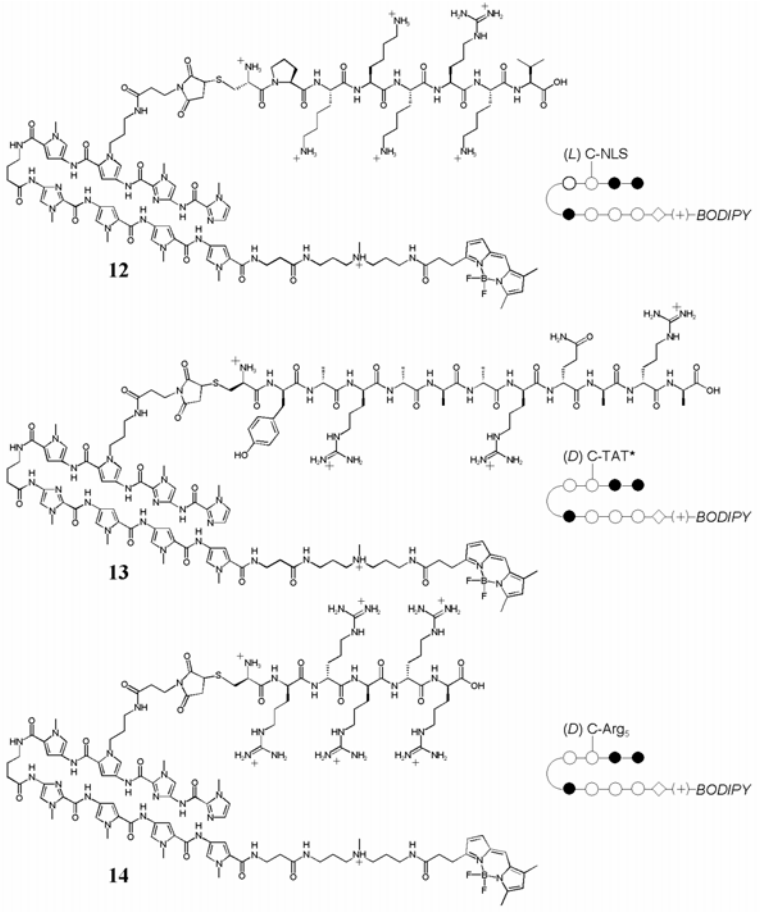
conjugates, compounds **8**-**11** (Figure 3.4a) were synthesized and tested by DNase I footprinting titration.[24] All compounds bound DNA at concentrations at or below 50 nM, with varying degrees of specificity over non-match binding sites. TAT- and (*D*) Arg_9-polyamide conjugates **9** and **11** coated DNA, showing no specificity for any binding site over any other. NLS- and (*D*) TAT*-polyamide conjugates **8** and **10** bound DNA with a specificity for their match sites over all other sites (coating) of 20-fold (Figure 3.4b).[25]

In order to improve the DNA-binding characteristics of polyamide-peptide conjugates, the peptide moiety was moved from the tail to the *N*-methyl position of an internal pyrrole. This arrangement removes the cationic peptides from the DNA minor groove, with the goal of providing compounds with increased affinity and specificity, hopefully retaining or improving their uptake characteristics. Peptide-polyamide-Bodipy conjugates **12**-**14** were synthesized and assayed for their uptake characteristics (Figure 3.5). These compounds had generally poorer uptake characteristics than the tail-linked cognates. Polyamide-peptide conjugates **15**-**17** were then synthesized and assayed for their DNA-binding qualities.

Compounds bearing up to five cationic charges exhibited specificity for binding the DNA match site, at generally higher affinities than the tail-linked counterparts (Figure 3.6). The good uptake characteristics of tail-linked conjugates and the good DNA-binding characteristics of *N*-methyl-linked compounds suggest that molecular shape is an important variable regarding polyamide-peptide conjugates. It was desirable to design a system that incorporated shape characteristics that would promote both good uptake and good DNA affinity and specificity. Such a situation might be achieved by compounds

bearing the peptide moiety on the terminal imidazole residue, rather than an internal residue (Figure 3.7).

(a)

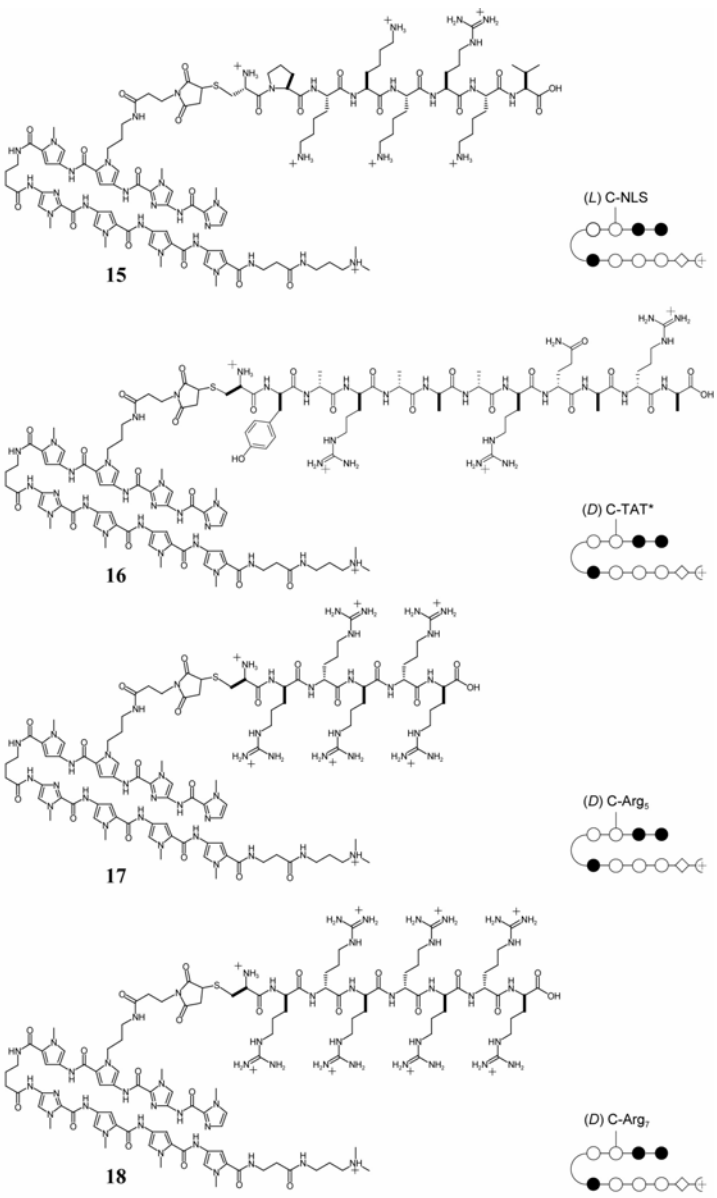


(b)

Conjugate	SKBR-3: human breast cancer	CEM: human cultured T-cells	NB4: human leukemia	293: human kidney fibroblast	Sf9: fall armyworm (insect)	Kc: Drosophila (insect)
12	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
13	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
14	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm

Figure 3.5 Localization of *N*-methyl-linked peptide-polyamide–Bodipy conjugates.[25]

(a)



(b)

Conjugate	Match site affinity K_a (M^{-1})	Specificity over coating
15	1×10^{10}	~100-fold
16	$\sim 3 \times 10^{9}$	~100-fold
17	4×10^{9}	25-fold
18	$\sim 5 \times 10^{8}$	2-5 -fold

Figure 3.6 DNA-binding characteristics of *N*-methyl-linked polyamide-peptide conjugates.[25]

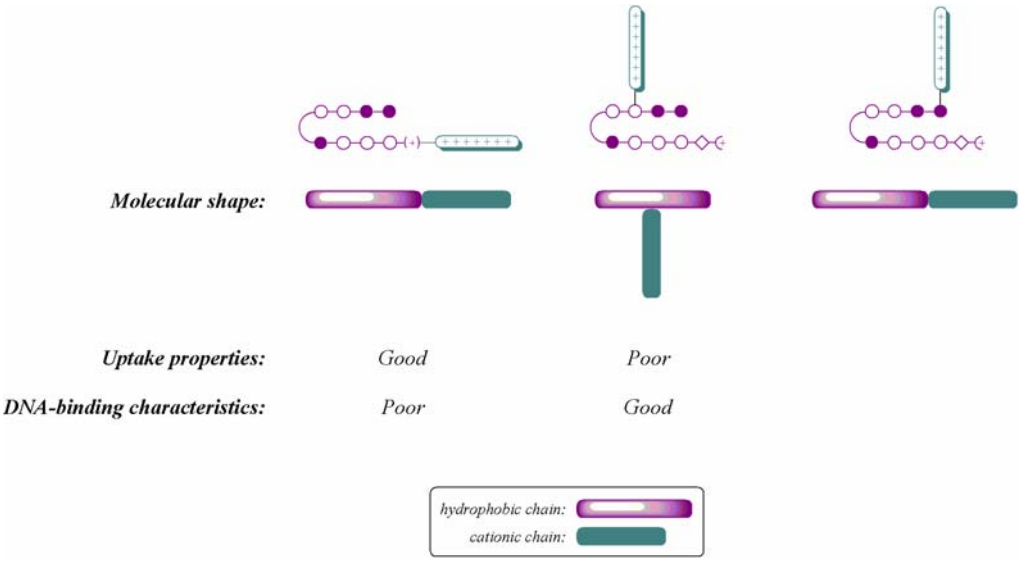


Figure 3.7 Effect of molecular shape on uptake and DNA-binding properties. Attachment of carrier peptides to the terminal imidazole's *N*-methyl position should combine the good uptake properties of tail-linked conjugates and the good DNA-binding properties of *N*-methyl-linked conjugates.

Results

N-terminal imidazole-linked peptide-polyamide-Bodipy conjugates **19**-**24** (Figure 3.8) were synthesized as shown in Figure 3.9. For the conjugation of the peptides to the polyamide, a maleimide-thiol Michael addition reaction was utilized because this reaction proceeds rapidly and selectively in aqueous buffers a physiological conditions. Compounds were added to a panel of fourteen cell lines, consisting of ten human cell lines (PC3, CEM, K562, NB4, MEG, HeLa, Jurkat, SKBR-3, MCF-7, and 786O), two mouse lines (3T3 and MEL), and two insect lines (KC and SF-9). These cell lines represent a wide variety of human cancers, including hard-to-treat varieties such as kidney (786O) and ovarian (HeLa) cancer. Treated cells were imaged by confocal microscopy and assayed for the localization of fluorescence (Figure 3.10).

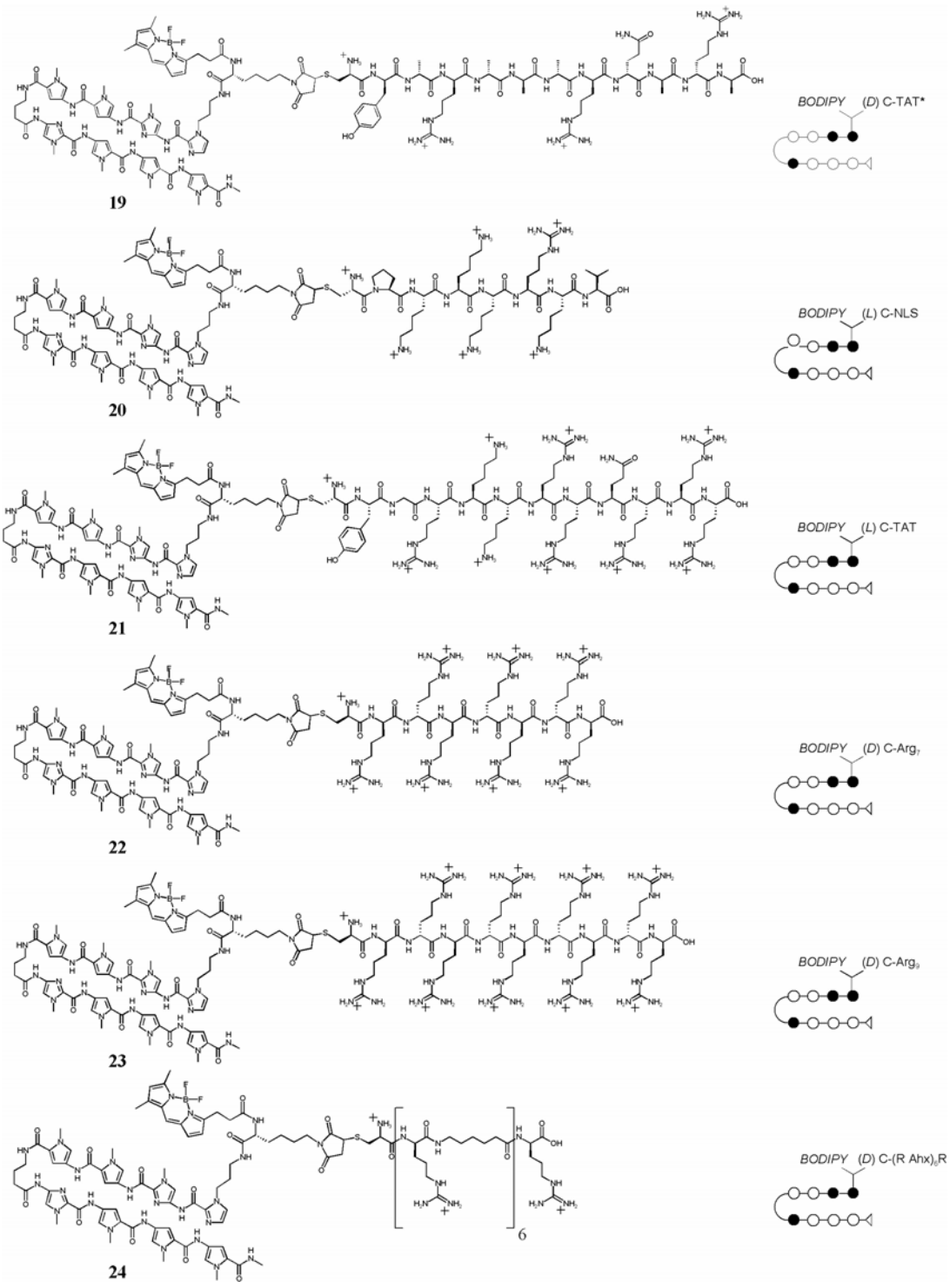


Figure 3.8 N-terminal imidazole-linked peptide-polyamide-Bodipy conjugates for nuclear uptake studies.

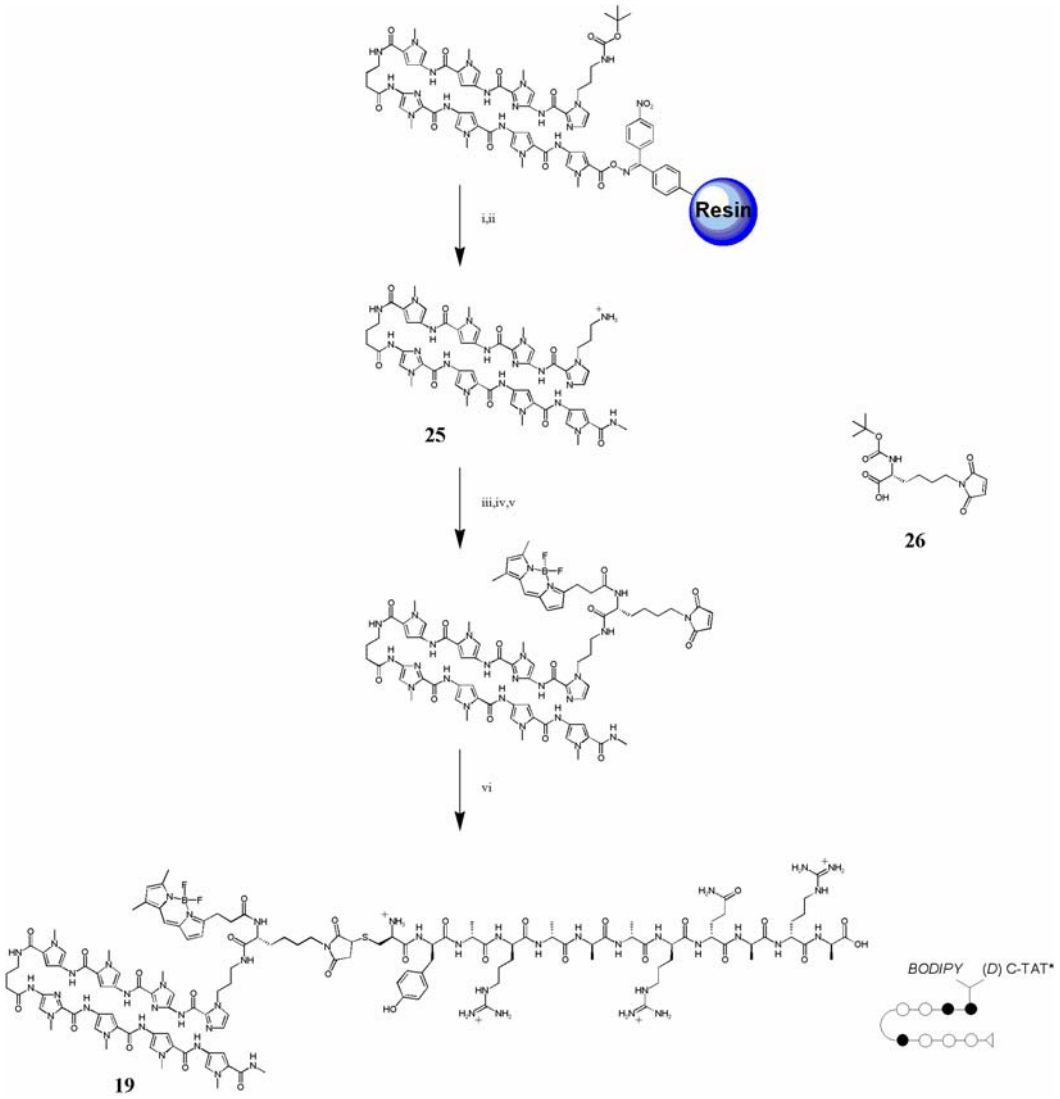


Figure 3.9 Synthesis of a representative compound in the N-terminal-linked series. (i) 1:1 mixture of methylamine (2M in THF):CH_2Cl_2, 6 hrs at 37ºC. (ii) 20% TFA in CH_2Cl_2 (v/v), 30 min at room temperature. (iii) **26**, HOBt, DCC, DIEA, DMF. (iv) 20% TFA in CH_2Cl_2 (v/v), 30 min at room temperature. (v) Bodipy-FL, DCC, HOBt, DIEA, DMF. (vi) TCEP, (*D*) Cys-TAT*, 6M Gn·HCl, pH = 7.0, DMF, 30 min at room temperature.

(a)

Conjugate	PC3	CEM	K562	NB4	MEG	HeLa	Jurkat	SKBR-3	MCF-7	786O	3T3	MEL	Kc	Sf9
19	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt
20	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt
21	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Toxic	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt
22	Cyt	Toxic	Toxic	Toxic	Toxic	Cyt	Toxic	Cyt	Cyt	Cyt	Cyt	Toxic	Cyt	Cyt
23	Toxic	Toxic	Toxic	Toxic	Toxic	Toxic	Toxic	Toxic	Toxic	Toxic	Cyt	Toxic	Cyt	Cyt
24	Cyt	Toxic	Cyt	Nuclear	Toxic	Cyt	Toxic	Toxic	Cyt	Toxic	Cyt	Cyt	Cyt	Cyt

(b)

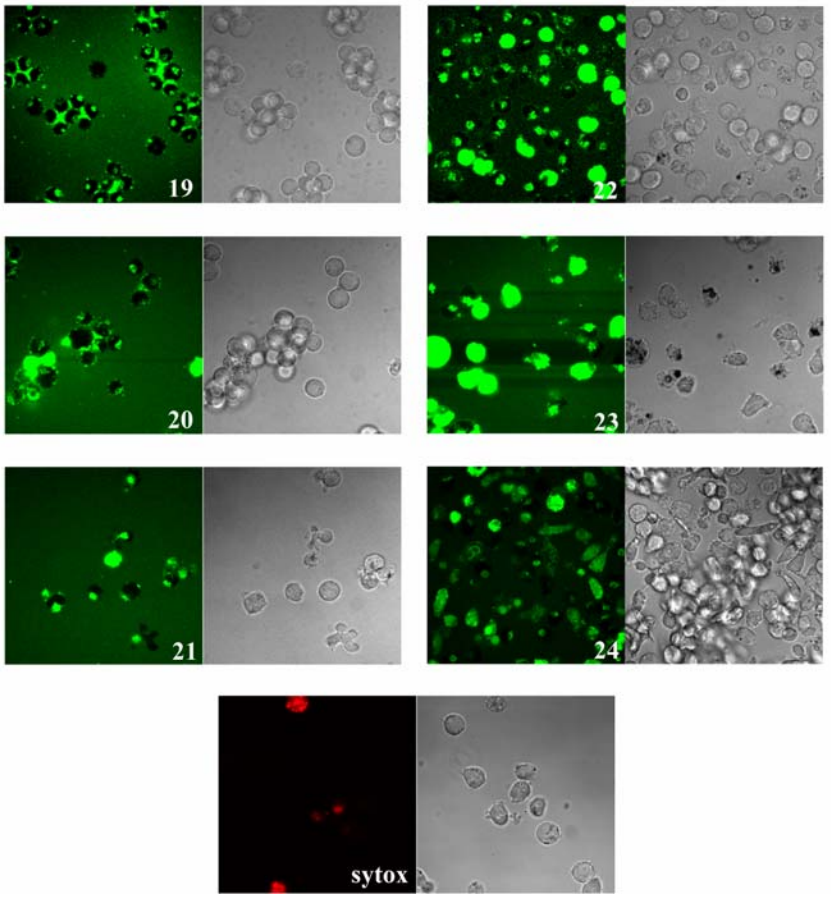


Figure 3.10 Localization of **19-24** in live cells as determined by confocal microscopy. (a) The designation 'nuc' indicates observation of fluorescence in the interior of the nucleus. The designation 'cyt' indicates cellular, non-nuclear fluorescence. The designation 'toxic' indicates compound induces toxicity at the concentration studied. Cells were imaged directly following 20 h incubation with 5 μM **19**–**24** under normal growth conditions for each cell-line. (b) Images of **19-24** and dead-cell stain Sytox orange in NB4 cells.

In general, compounds **19**-**24** showed poor cellular uptake characteristics similar to those found in internal *N*-methyl-linked conjugates **15**-**18**. Conjugates **19**-**21** were found to be cytoplasmic (or in the case of Jurkat, toxic) in all cell lines studied. Polyarginine conjugates **22**-**24** showed either cytoplasmic staining or toxicity, with Arg_9 conjugate **23** being the most toxic compound in each human cell line studied. The sole entry showing somewhat promising localization was compound **24** in NB4 cells. The $(R\text{-}Ahx)_6R$ peptide was developed by Wender and co-workers as a superior carrier peptide in jurkat cells.[26] In this assay it was generally too toxic to be useful as a carrier at the concentrations studied.

Though the uptake characteristics of the N-terminal imidazole-linked peptide conjugates were not promising, the DNA-binding characteristics of this system were nonetheless interesting to determine. To that end, compounds **27**-**32** (Figure 3.11) were synthesized and assayed by DNase I footprinting on the plasmid pDEH9, which bears a match 5'-TGGTCA-3' site and discrete mismatch binding sites 5'-TGGCCA-3' and 5'-TGGGCA-3' for the polyamide unit used in this study (Figure 3.12a). Analysis of the footprinting gels (Figure 3.12b) shows that the DNA-binding characteristics of N-terminal-linked conjugates are intermediate between tail linkage and internal *N*-methyl linkage. (*D*) TAT*- and NLS-conjugates **27** and **28** show the best specificity for their DNA match site over nonspecific binding (coating). Surprisingly, the $(R\text{-}Ahx)_6R$-conjugate **32** shows some specificity for its match site over coating (10-20 fold) even though the R_7-conjugate **30**, bearing the same overall molecular charge, shows negligible specificity over coating. The poor DNA-binding qualities of the polyarginine conjugates, coupled with their toxicity to cells, makes them unsuitable for conjugation to polyamides

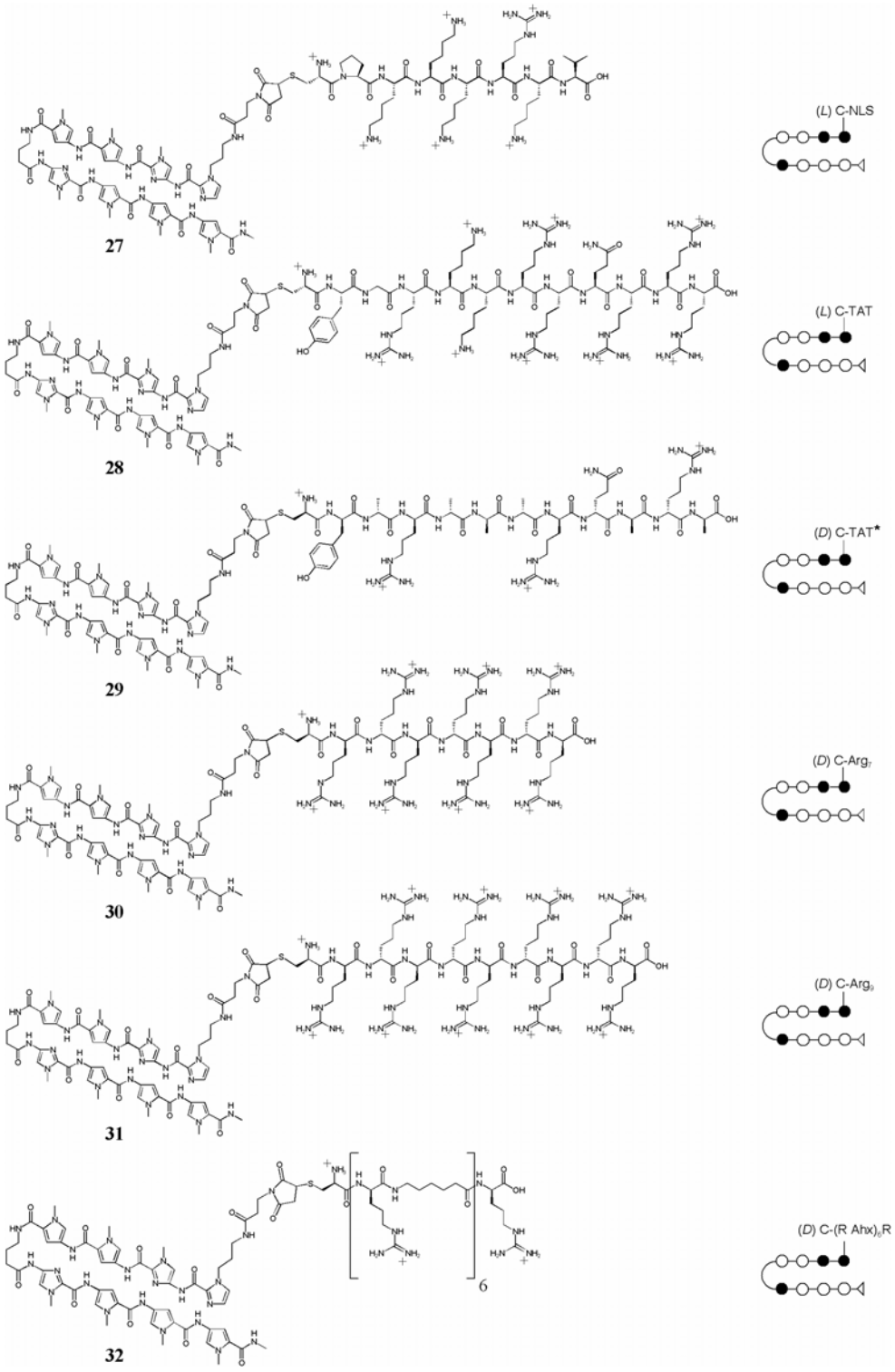


Figure 3.11 N-terminal imidazole-linked polyamide-peptide conjugates for footprinting studies.

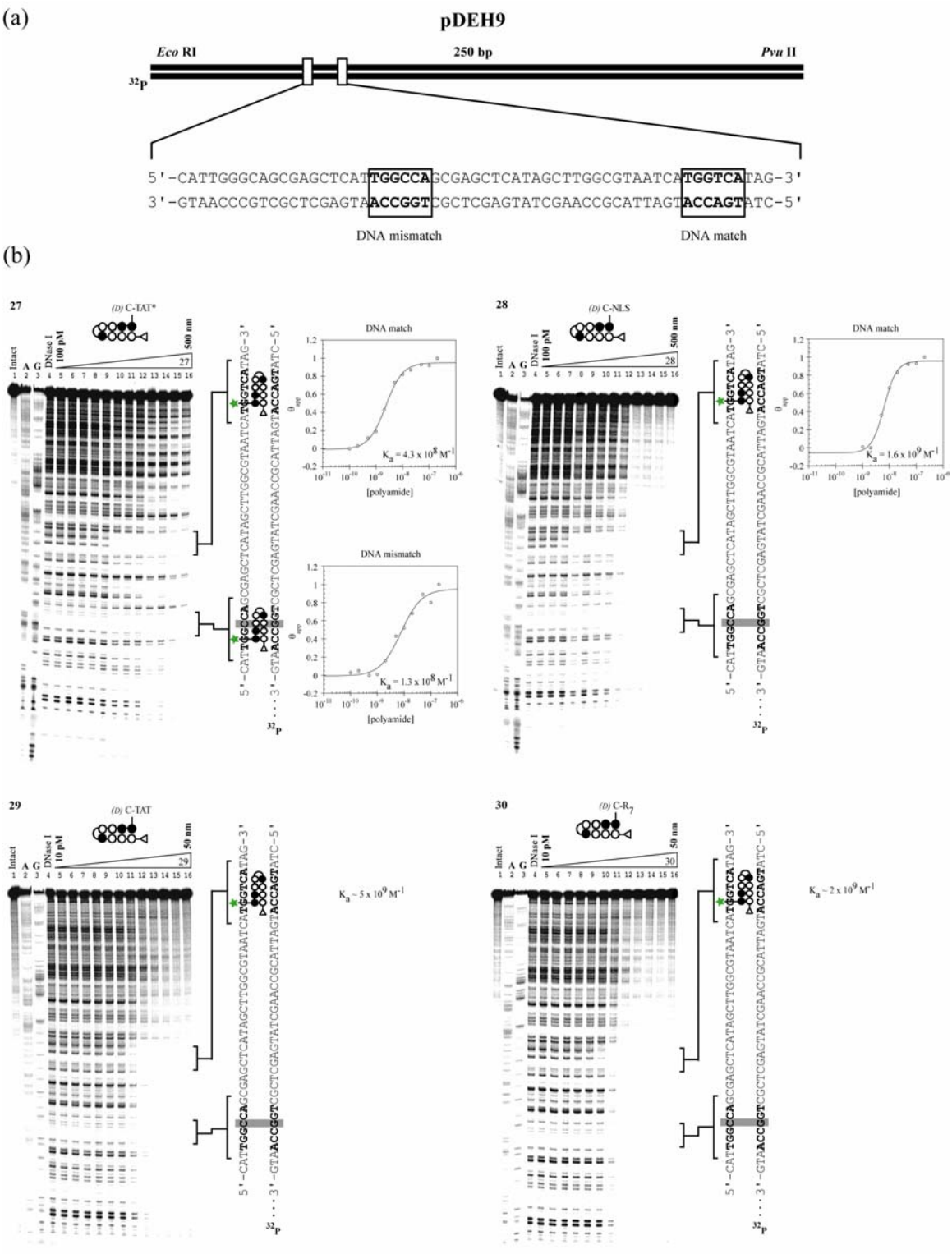
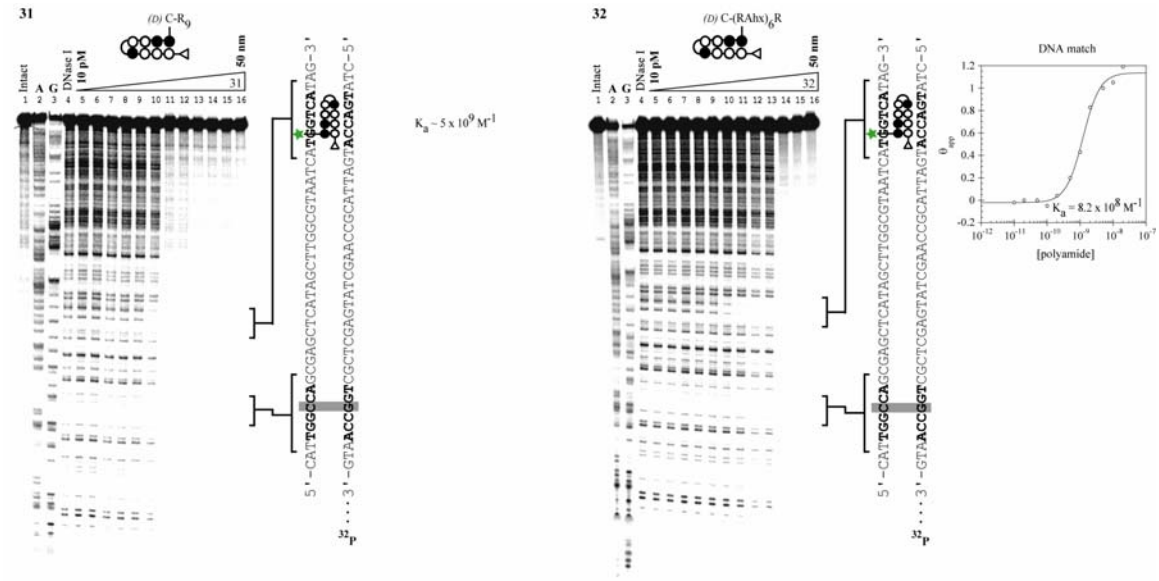


Figure 3.12 DNA-binding characteristics of N-terminal imidazole-linked polyamide-peptide conjugates. (a) Plasmid pDEH9 illustrating match and mismatch sites for the polyamide used in this study. (b) DNase I footprinting gels for **27**-**32**, showing isotherms where measurable.



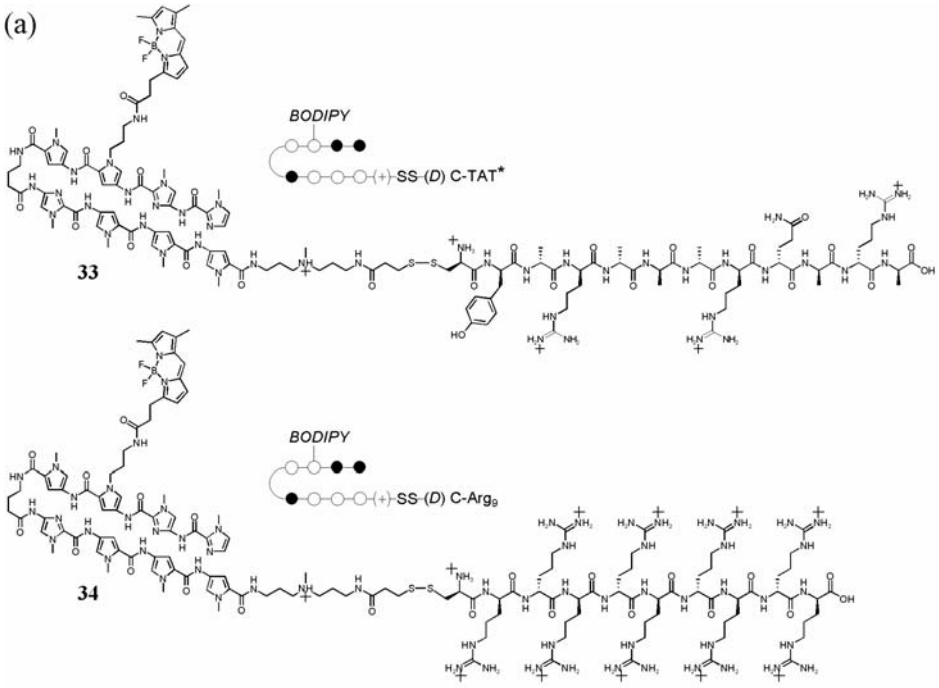
(c)

Conjugate	Match site affinity K_a (M^{-1})	Specificity over coating	Conjugate	Match site affinity K_a (M^{-1})	Specificity over coating
8 + –(*L*) C-NLS	5×10^8	20-fold			
9 + –(*L*) C-TAT	~1×10^9	None	**27** (*L*) C-NLS	2×10^9	25-fold
10 (+)–(*D*) C-TAT*	6×10^8	20-fold	**28** (*L*) C-TAT	~5×10^9	None
11 (+)–(*D*) $C\text{-}Arg_9$	~1×10^9	None	**29** (*D*) C-TAT*	4×10^8	~100-fold
15 (*L*) C-NLS	1×10^{10}	~100-fold	**30** (*D*) $C\text{-}Arg_7$	~3×10^9	2-5 -fold
16 (*D*) C-TAT*	~3×10^9	~100-fold	**31** (*D*) $C\text{-}Arg_9$	~5×10^9	None
17 (*D*) $C\text{-}Arg_5$	4×10^9	25-fold	**32** (*D*) C-(R Ahx)$_6$R	8×10^8	10-20 -fold
18 (*D*) $C\text{-}Arg_7$	~5×10^8	2-5 -fold			

Figure 3.12 DNA-binding characteristics of N-terminal imidazole-linked polyamide-peptide conjugates. (c) Comparison of DNA-binding characteristics of carrier peptide-polyamide conjugates in three morphologies. In general **27**-**32** possess intermediate binding characteristics between tail-linked **8**-**11** and internally-linked **15**-**18**.

for use as DNA-binding ligands in living systems.

One possible way to overcome the poor DNA-binding characteristics of carrier peptide-polyamide conjugates is by employing a pro-drug approach, attaching a carrier moiety to a polyamide via a linking domain that will be cleaved upon entering the cellular environment. One such linking strategy is the disulfide bond, which is cleaved upon entering the reductive cytoplasmic environment to the constituent thiols.[8] This strategy was implemented by the synthesis of compounds **33** and **34**, containing a Bodipy-polyamide moiety linked to either (*D*) TAT* or Arg_9 through a disulfide bond



(b)

Conjugate	PC3	CEM	NB4	MEG	HeLa	Jurkat	MCF-7	786O	MEL
33	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt
34	Cyt	Cyt	Toxic	Toxic	Cyt	Cyt	Cyt	Cyt	Toxic

Figure 3.13 Disulfide-linked polyamide-peptide prodrugs. (a) Chemical structures and ball-and-stick models of **33** and **34**. (b) Uptake profile of disulfide conjugates in nine mammalian cell lines.

(Figure 3.13a). These conjugates were synthesized according to the scheme outlined in Figure 3.14. Upon exposure to a panel of mammalian cell lines, **33** and **34** were assayed for cellular uptake by confocal microscopy (Figure 3.13b). Unfortunately, neither conjugate was successful in staining the nucleus of any cell line. Arg_9 remained a toxic constituent in some cell lines. In general, carrier peptide conjugates do not seem to be a fruitful path forward in promoting uptake of polyamides. It is likely that small-molecule conjugates will be a more successful approach to the problem of utilizing polyamides in living systems.

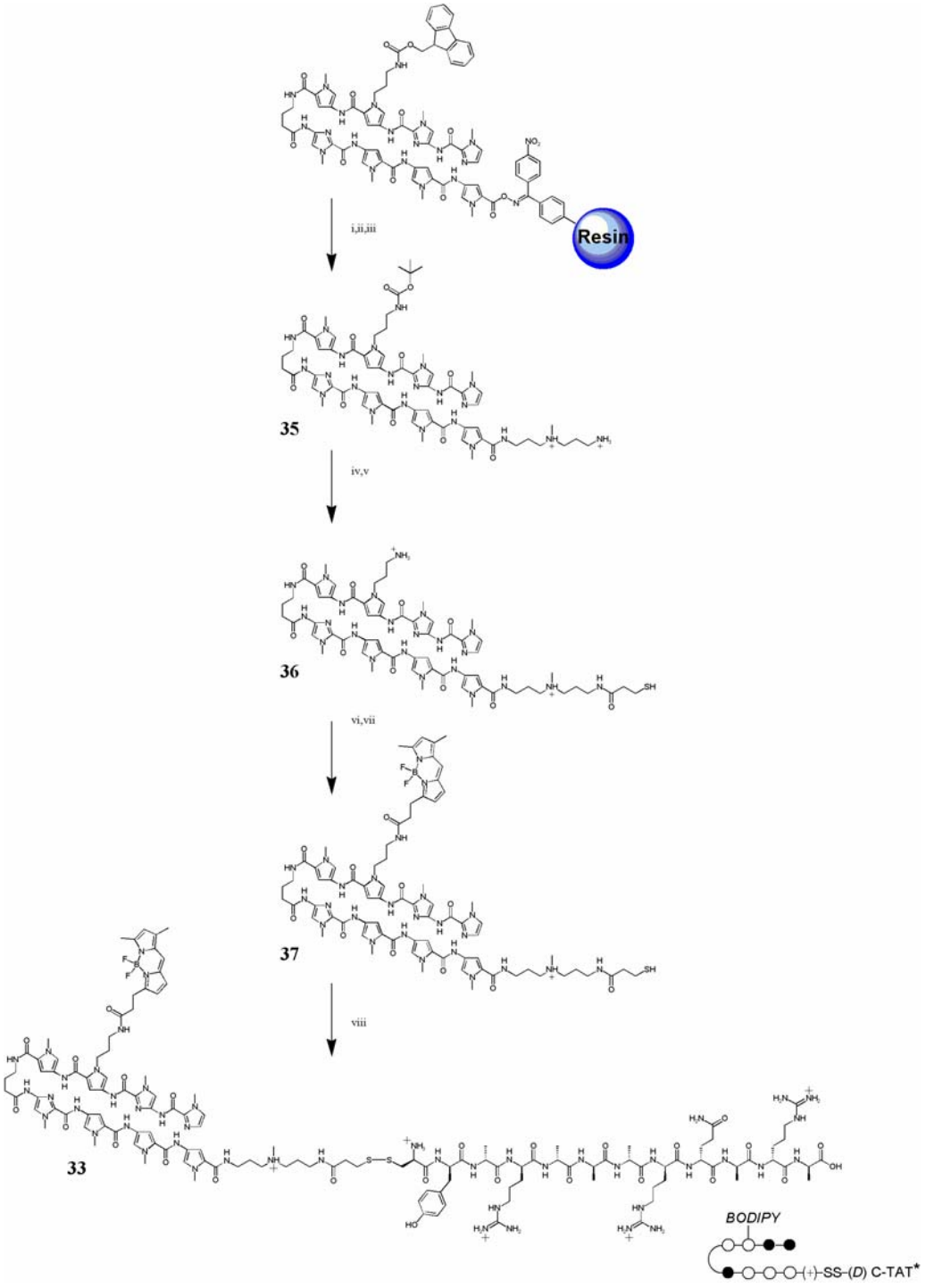


Figure 3.14 Synthesis of disulfide conjugates. (i) 20% piperidine/DMF (v/v), 30 min at room temperature. (ii) Boc_2O, DIEA, DMF. (iii) 3,3'-diamino-*N*-methyldipropylamine, 6 hr at 37ºC. (iv) *S*-trityl-3-mercaptopropionic acid, HOBt, DCC, DIEA, DMF. (v) 50% TFA in CH_2Cl_2 (v/v), 1 hr at room temperature. (vi) Bodipy-FL, HOBt, DCC, DIEA, DMF. (vii) 3-dimethylaminopropylamine, 2 hr at 37ºC. (viii), 2-Aldrithiol, (*D*) Cys-TAT*, DMF, 4 hrs at room temperature.

Chapter 3B

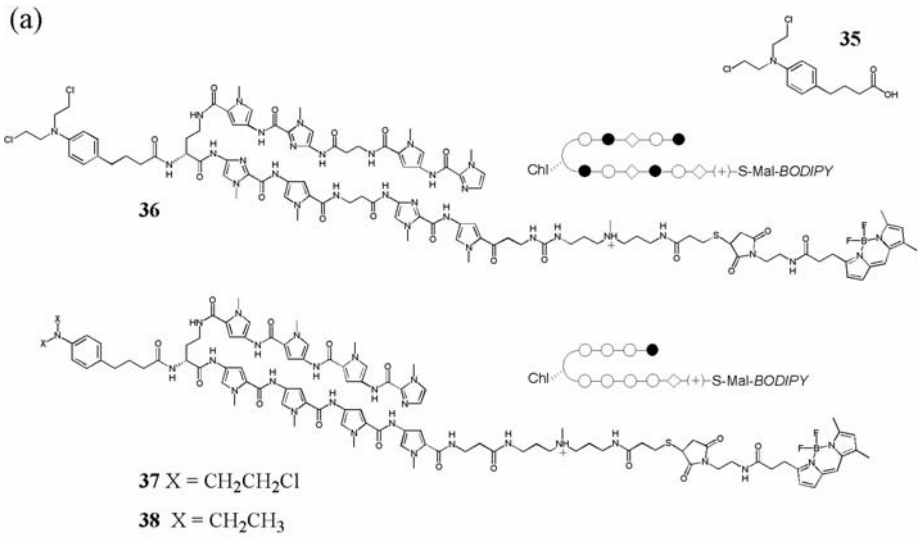
Polyamide-Small Molecule Conjugates for Cellular Uptake Studies

Introduction

Continuing experiments aimed at creating polyamides or polyamide conjugates that cross the outer membranes of living cells and transit to the nucleus resulted in the creation of several small molecule-polyamide conjugates. The first series of such compounds studied incorporated the DNA-alkylating agent chlorambucil (CHL, **35**). The laboratories of both Terry Beerman and Joel Gottesfeld have observed effects such as toxicity and DNA-alkylation in a variety of cell lines with polyamide-CHL conjugates.[29] As they were producing an effect in living cells, it was interesting to test Bodipy-polyamide-CHL conjugates for cellular localization. Compounds **36**-**38** were synthesized and tested in a wide panel of cell lines, localizing to the cytoplasm of all cell lines save NB4, CEM, and human primary CD4+ T-cells (Figure 3.15). These cell lines exhibited nuclear staining upon treatment with **36** and **37**, without accompanying toxicity at the concentration studied (5 μM).

The successful staining of NB4 nuclei suggested that further attempts to optimize the uptake characteristics of Bodipy-polyamide-CHL conjugates might meet with some success. Many small molecules serving as nutrients or signal transduction messengers are taken up into cells by specific transmembrane protein receptors. It has been shown that the cellular uptake of several types of biopolymers, including DNA, peptide nucleic acids, and proteins, can be increased when attached to receptor-specific small molecules.[30-32] The folate receptor is a frequent target of these compounds, since its process of uptake by receptor-mediated endocytosis is well understood and the receptor is

(a)



(b)

Conjugate	PC3	CEM	NB4	SKBR-3	MCF-7	MEL	293	LnCap	Kc	Sf9
36	Cyt	Nuclear	Nuclear	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt
37	Cyt	Nuclear	Nuclear	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt
38	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt	Cyt

Figure 3.15 Bodipy-polyamide-chlorambucil conjugates for cell uptake studies. (a) Chemical structures and ball-and-stick models of chlorambucil, **35**, and Bodipy-polyamide-CHL conjugates **36**[29] and **37**-**38**[25]. (b) The designation 'nuclear' indicates observation of fluorescence in the interior of the nucleus. The designation 'cyt' indicates cellular, non-nuclear fluorescence.

overexpressed in several cancers.[33] In order to exploit cell receptors as possible paths of cellular influx, Bodipy-polyamide conjugates with folic acid and cholic acid were synthesized and tested in cell uptake assays (Figure 3.16). These compounds were ineffective at increasing cellular uptake in the standard panel of cell lines. However, other small-molecule conjugates might have increased uptake if the operative receptor-targeting moiety is attached to the polyamide at the optimal linkage site, using an optimal linking domain.

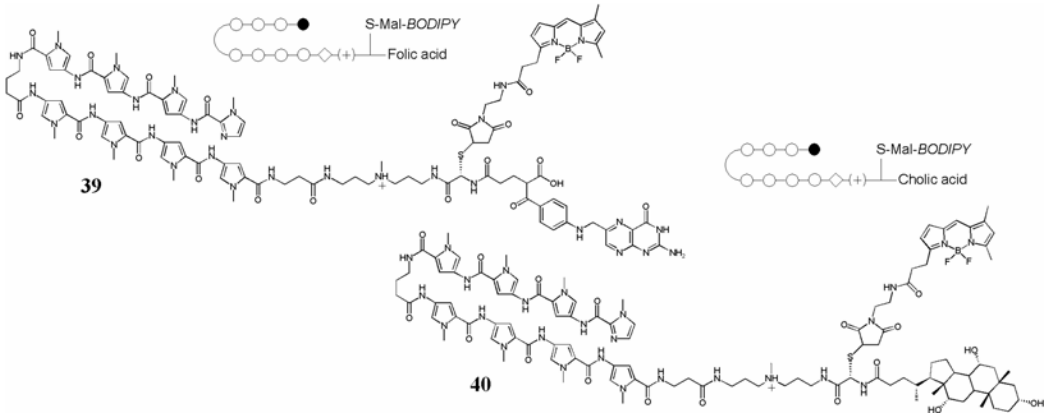
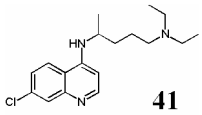


Figure 3.16 Bodipy-polyamide-small molecule conjugates for cell uptake studies.[25]

Most fluorescent conjugates studied show localization in the cytoplasm of living cells in a punctate pattern. This suggests that polyamides and polyamide conjugates are, by and large, able to cross the cellular membrane, but are then tied up in cytoplasmic vesicles such as endosomes or lysosomes. Another possible method to increase the uptake of polyamides and polyamide conjugates in living cells is co-treatment of the cells with drugs. A known endosomal disruption agent, chloroquine (**41**), was studied as one possible formulation agent for polyamide-Bodipy conjugate **2**.[34] At 10 μM added chloroquine, **2** was found to stain the nuclei of PC3 and MEL cells. At higher chloroquine concentrations (~100 μM) the cells showed toxicity, illustrating that any added agent must be both potent and specific in its effect, in order that the triggering of stress mechanisms or apoptosis be avoided.

The success of chloroquine in increasing the uptake of a polyamide conjugate suggests that other drugs, such as those known to inhibit the multidrug response (MDR) might be effective formulating agents.[35]



41

Results

Polyamide-chlorambucil conjugates

Two Bodipy-polyamide-CHL conjugates were designed as alternate morphologies of **36-38** to hopefully increase their cellular uptake and nuclear staining properties beyond merely NB4, CEM, and CD4+ T-cells. Compounds **42** and **43** were synthesized as in Figure 3.17. The necessity of having three differentially-protected amine moieties on **43** led to the use of pyrrole monomer **46**, possessing an *N*-propylamine masked as an azide. After addition of the CHL moiety, the azide was cleanly and specifically converted to an amine via the Staudinger reaction, leaving the CHL moiety intact. This strategy of amine protection/deprotection may prove useful in the future in the synthesis of complex polyamide conjugates. Unfortunately, upon exposure to the standard panel of cell lines, both **42** and **43** were found to localize to the cytoplasm in all cases.

Polyamide-dihydrotestosterone conjugate

Dihydrotestosterone (DHT), a steroid hormone, has been used to target prostate cancer cells for uptake.[30] Based on this, and pursuant to an ongoing collaboration with ZhengxinWang's lab at the University of Texas Southwestern Medical Center (see chapter 5), compound **50** was synthesized (Figure 3.18). This fluorescent compound incorporates a hydrophilic polyethylene glycol (PEG) linking domain between the polyamide and steroid moieties. It was hoped that the increased hydrophilicity of the compound, due to the linking domain, and the receptor-targeting ability of the DHT domain would afford increased uptake in prostate cancer cell lines PC3 and LNCaP. Indeed, a comparison of the uptake of **2** with **50** shows marked increase in cytoplasmic

staining in PC3 and LNCaP cells. However, no nuclear staining is detected, suggesting that the novel domains do act to increase the transit of the polyamide across the cellular membrane, but are unable to overcome the sequestration of compound in cytoplasmic vesicles.

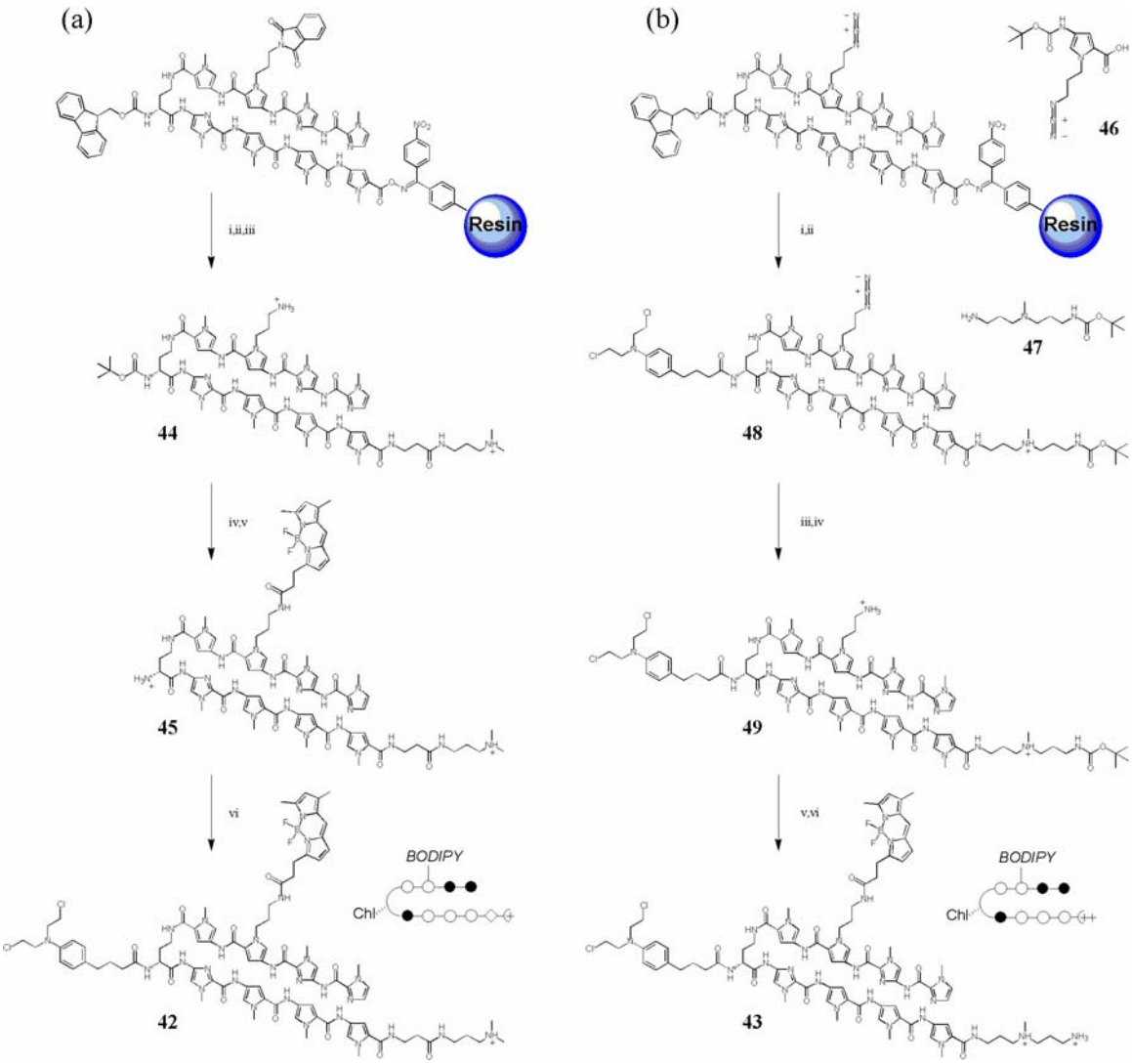


Figure 3.17 Synthesis of CHL conjugates. (a): (i) 20% piperidine/DMF (v/v), 30 min at room temperature. (ii) Boc_2O, DIEA, DMF. (iii) 3-dimethylaminopropylamine, 6 hrs at 37ºC. (iv) Bodipy-FL, HOBt, DCC, DIEA, DMF. (v) 20% TFA in CH_2Cl_2 (v/v), 30 min at room temperature. (vi) **35**, HOBt, DCC, DIEA, DMF. (b): (i) **47**, 18 hrs at 37ºC. (ii) **35**, HOBt, DCC, DIEA, DMF. (iii) THF, PPh_3, 3 hrs at room temperature. (iv) $NaOH_{(aq)}$ 1 eq, 2 hrs at 70ºC. (v) Bodipy-FL, HOBt, DCC, DIEA, DMF. (vi) $BF_3{\cdot}(Et_2O)_2$, CH_2Cl_2, 30 min at room temperature.

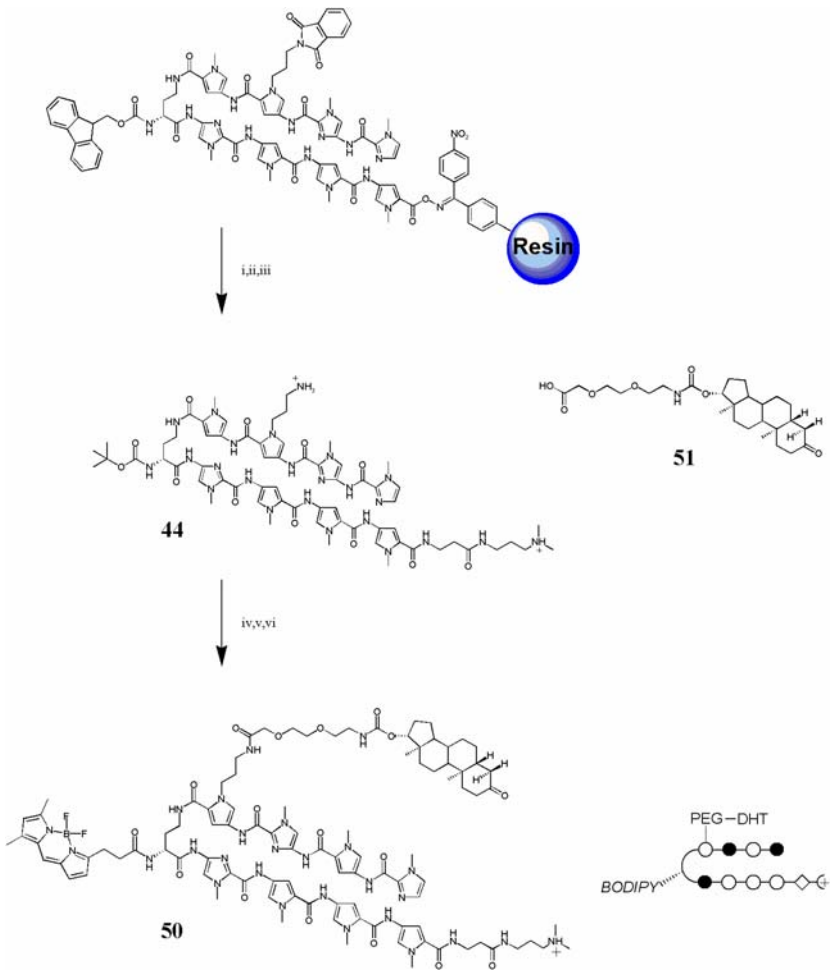


Figure 3.18 Synthesis of DHT conjugate. (i) 20% piperidine/DMF (v/v), 30 min at room temperature. (ii) Boc_2O, DIEA, DMF. (iii) 3-dimethylaminopropylamine, 6 hrs at 37ºC. (iv) **51** (for synthesis, see chapter 5, compound **13**), HOBt, DCC, DIEA, DMF. (v) 20% TFA in CH_2Cl_2 (v/v), 30 min at room temperature. (vi) Bodipy-FL, HOBt, DCC, DIEA, DMF.

Experimental Section

Synthesis of Polyamide Conjugates 1-18, 36-40

These conjugates were synthesized by P. S. Arora and J. M. Belitsky.[25,37]

Synthesis of 26[38]

Maleimide (1 g, 10 mmol), methyl chloroformate (0.8 mL, 10 mmol), and *N*-methylmorpholine (1.13 mL, 10 mmol) were dissolved in EtOAc (50 mL) at 0ºC and stirred for 1 hr. The resulting white precipitate was removed by filtration and the filtrate concentrated to obtain a pink solid. Boc- (*D*) Lys-OH (0.9 g, 3.5 mmol) was dissolved in sat. $NaHCO_3$ solution (35 mL) at 0ºC, to which was added the pink solid and the resulting suspension was stirred for 4 hrs, with slow warming to room temperature. The solution was extracted with $CHCl_3$ (2 x 150 mL), and the aqueous layer was acidified to pH~3 with concentrated H_2SO_4. The solution was washed with $CHCl_3$ (3 x 100 mL). The organic fractions were combined, dried with $MgSO_4$, and concentrated to yield a translucent yellow oil.

Synthesis of Peptide-Polyamide-Bodipy Conjugates 19-24

The synthesis of the 1-[3-(*tert*-butoxycarbonyl)amino]propyl-imidazole-2-carboxylic acid monomer utilized in the synthesis of **19-24** has been recorded.[29] Polyamide was synthesized on Kaiser oxime resin by standard solid phase methods as previously reported and liberated upon treatment with methylamine to provide **25**.[27,28] **26** (230 μmol) was dissolved in 300 μL DMF, to which was added DCC (230 μmol, 100 μL DMF) and HOBt (230 μmol, 100 μL DMF), and the resulting solution allowed to react at

37ºC for 30 min. This solution was added to **25** (70 μmol), dissolved in DMF (500 μL) and DIEA (40 μL, 230 μmol), and allowed to react at room temperature for 3 hrs, until complete as monitored by analytical HPLC. The crude product was dissolved into ~10 mL CH_2Cl_2, to which was added ~10 mL TFA and the resulting solution stirred at room temperature for 30 min. The solution was concentrated, purified by C_{18} reverse-phase preparatory HPLC, and lyophilized to yield a white powder. To 10 μmol of this polyamide dissolved in DMF (185 μL), was added first a solution of Bodipy-FL (20 μmol), DCC (20 μmol), and HOBt (20 μmol) dissolved in 300 μL DMF allowed to react at room temperature for 30 min, then DIEA (3.5 μL, 20 μmol), and the resulting solution was allowed to react at room temperature for 3 hrs. To 100 μL of this solution was added a solution of the appropriate peptide bearing an N-terminal cysteine residue (5 μmol) and TCEP (7.5 μmol) dissolved in 6M Gn·HCl (100 μL, pH = 7.3). The resulting solution was allowed to react at room temperature for 1 hr, and then purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **19-24** as orange powders. Characterization: **19** [(*D*) C-TAT*-PA-Bodipy], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2913.9, obsd 2914.4; **20** [(*L*) C-NLS-PA-Bodipy], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2594.2, obsd 2594.4; **21** [(*L*) C-TAT-PA-Bodipy], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 3270.0, obsd 3270.6; **22** [(*D*) C-R_7-PA-Bodipy], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2821.4, obsd 2821.7; **23** [(*D*) C-R_9-PA-Bodipy], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 3132.5, obsd 3132.9; **24** [(*D*) C-$(RAhx)_6$R-PA-Bodipy], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 3500.3, obsd 3500.4.

Synthesis of Polyamide-Peptide Conjugates 27-32

25 (35 μmol) was dissolved in 600 μL DMF and treated with DIEA (59 μL, 350 μmol) and a solution of 3-maleimidopropionic acid (0.028 g, 175 μmol), DCC (0.035 g, 165 μmol) and HOBt (0.024 g, 175 μmol) reacted for 30 min at room temperature in 200 μL DMF. The resulting solution was allowed to react at room temperature for 3 hrs then purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield a white powder. The maleimido-polyamide (2 μmol) was dissolved in 160 μL DMF, to which was added a solution of TCEP (4 μmol) and peptide (10 μmol) in 200 μL Gn·HCl (6 M, pH = 7.3). The resulting solution was allowed to react at room temperature for 1 hr, and then purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **27-32** as white powders. Characterization: **27** [(*L*) C-NLS-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2261.4, obsd 2261.1; **28** [(*L*) C-TAT-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2936.1, obsd 2935.5; **29** [(*D*) C-TAT*-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2582.1, obsd 2581.8; **30** [(*D*) C-R_7-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2489.3, obsd 2489.2; **31** [(*D*) C-R_9-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2801.4, obsd 2801.1; **32** [(*D*) C-$(RAhx)_6$R-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 3168.0, obsd 3166.8.

Synthesis of Polyamide-Disulfide-Peptide Conjugates 33-34

Polyamide was synthesized on Kaiser oxime resin employing the 1-[3-(fluorenylmethyloxycarbonyl)amino]propyl-4-(*tert*-butoxycarbonyl)amino-pyrrole-2-carboxylic acid monomer, whose synthesis has been detailed elsewhere.[29] The Fmoc protecting group was removed upon treatment of the resin with 20% (v/v)

piperidine/DMF, and replaced with the *t*-Boc protecting group by treatment of the resin with Boc_2O (0.5 g, 2.3 mmol) and DIEA (400 μL, 2.3 mmol) in 4 mL NMP. The polyamide was liberated from resin by treatment with 3,3'-diamino-*N*-methyldipropylamine for 10 hrs at room temperature and purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **35** as a white powder.

S-trityl-3-mercaptopropionic acid (0.026 g, 75 μmol), DCC (75 μmol), and HOBt (75 μmol) were dissolved in 225 μL DMF and allowed to activate for 3 hrs at room temperature. The resulting solution was centrifuged to remove precipitated DCU and added to a solution of **35** (11 μmol), dissolved in 300 μL DMF, and DIEA (15 μL, 90 μmol) was added. This was allowed to react for 18 hrs at room temperature. Dp (8 μL) was added (to quench any remaining *S*-trityl-3-mercaptopropionic acid –OBt ester), then the crude product was precipitated by addition of Et_2O and treated with 50% (v/v) TFA in CH_2Cl_2 for 1 hr at room temperature. The resulting solution was purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **36** as a white powder.

A solution of Bodipy-FL (17 μmol), DCC (17 μmol), and HOBt (17 μmol) was prepared in 200 μL DMF and allowed to activate for 2 hrs at room temperature. This was added to a solution of **36** (9 μmol) in DMF (300 μL) and DIEA (10 μL, 60 μmol) and allowed to react at room temperature for 1 hr. This solution was purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **37** and the thioester of **37** and Bodipy (1:1) as orange powders. The thioester was treated with 10 μL Dp to cleave the Bodipy thioester and purified to yield **37**.

Each peptide (2 μmol) was dissolved in 75 μL DMF. 2-Aldrithiol (0.5 μmol) was added and the solutions allowed to react for 1 min. A solution of **37** in 75 μL DMF was

then added to each peptide solution and allowed to react for 4 hrs. They were then purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **33** and **34** as orange powders. Characterization: **33** [(*D*) C-TAT*-SS-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2905.4, obsd 2907.8; **34** [(*D*) C-R_9-PA], MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 3124.7, obsd 3126.3.

Synthesis of Bodipy-Polyamide-Chlorambucil Conjugate (42)

Polyamide was synthesized on Boc-β-Ala-PAM resin employing the 1-[3-phthalimido)amino]propyl-4-(*tert*-butoxycarbonyl)amino-pyrrole-2-carboxylic acid monomer, whose synthesis has been reported elsewhere.[39] The Fmoc protecting group on the DABA turn was removed upon treatment of the resin with 20% (v/v) piperidine/DMF, and replaced with the *t*-Boc protecting group by treatment of the resin with Boc_2O (0.64 g, 3.0 mmol) and DIEA (200 μL, 1.7 mmol) in 3 mL DMF. The polyamide was liberated from resin by treatment with Dp for 18 hrs at 37ºC and purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **44** as a white powder.

A solution of Bodipy-FL (22 μmol), DCC (22 μmol), and HOBt (22 μmol) was prepared in 265 μL DMF and allowed to activate for 1 hr at room temperature. To this was added a solution of **44** (11 μmol) and DIEA (6 μL, 34 μmol), and the resulting solution was allowed to react at room temperature for 10 hrs. The crude product was precipitated by addition of Et_2O and treated with 20% (v/v) TFA in CH_2Cl_2 for 30 min at room temperature. The resulting solution was purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield **45** as an orange powder.

A solution of chlorambucil (0.009 g, 30 μmol), DCC (30 μmol), and HOBt (30 μmol) was prepared in 200 μL DMF and allowed to activate for 1 hr at room temperature. To the resulting solution was added a solution of **45** (3 μmol) in 350 μL DMF and DIEA (2 μL, 10 μmol), which was allowed to react for 3 hrs at room temperature, then purified by C_{18} reverse-phase preparatory HPLC and lyophilized to provide **42** as an orange powder. Characterization: ESI peaks of $[M+H]^+$ (monoisotopic mass ^{35}Cl) calcd 1840.9, obsd 1841.1; $[M+H]^+$ (isotopic mix of one ^{35}Cl and one ^{36}Cl) calcd 1841.9, obsd 1842.4; $[M+H]^+$ (monoisotopic mass ^{36}Cl) calcd 1842.9, obsd 1843.3.

Synthesis of Bodipy-Polyamide-Chlorambucil Conjugate (43)

Polyamide was synthesized on Kaiser oxime resin using azide pyrrole monomer **46** and liberated by cleavage with **47** (~1g, ~4 mmol) in 1 mL NMP. This solution was purified by C_{18} reverse-phase preparatory HPLC and lyophilized to provide polyamide as a white powder. The polyamide (8.8 μmol) was treated with an activated solution of chlorambucil (45 μmol), DCC (45 μmol), and HOBt (45μmol), as well as with DIEA (15 μL, 90 μmol). The resulting solution was allowed to react at room temperature for 3 hrs, then the crude product **48** was precipitated by addition of Et_2O.

Solutions of **48** (4.5 μmol) and triphenylphosphine (4.5 μmol) were prepared in anhydrous THF, then mixed and allowed to react for 3 hrs at room temperature. To this was then added aqueous NaOH (4.5 μmol) and the resulting solution heated at 70ºC for 2 hrs, until complete by analytical HPLC. The crude product **49** was precipitated by addition of Et_2O, then treated with an activated solution of Bodipy-FL (11 μmol), DCC (11 μmol), and HOBt (11 μmol), and DIEA (15 μL, 11 μmol). The crude product was

precipitated by addition of Et_2O, then resuspended in 100 μL anhydrous CH_2Cl_2, to which was added 10 μL $BF_3{\cdot}(Et_2O)_2$ and the solution allowed to react at room temperature for 30 min. The crude conjugate was precipitated upon addition of Et_2O and purified by C_{18} reverse-phase preparatory HPLC and lyophilized to provide **43** as an orange powder. Characterization: ESI peaks of $[M+H]^+$ (monoisotopic mass ^{35}Cl) calcd 1812.9, obsd 1812.5; $[M+H]^+$ (isotopic mix of one ^{35}Cl and one ^{36}Cl) calcd 1813.9, obsd 1813.5; $[M+H]^+$ (monoisotopic mass ^{36}Cl) calcd 1814.9, obsd 1814.6.

Synthesis of 1-[3-(azido)propyl]-4-(*tert*-butoxycarbonyl)amino-pyrrole 2-carboxylic acid (46)

Ethyl-4-nitropyrrole-2-carboxylate (2.5g, 13 mmol) was dissolved in acetone, to which was added K_2CO_3, and the resulting suspension stirred vigorously at room temperature for 2 hrs. To this was added 3-iodopropanol (2.6 mL, 26 mmol), then the suspension was stirred vigorously under reflux for 2 hrs, 30 min. The slurry was cooled to room temperature, filtered, and concentrated to yield a yellow oil. To this was added 100 mL H_2O and the pH was adjusted to ~3 with 10% (v/v) H_2SO_4. The mixture was extracted with EtOAc (3 x 100 mL), dried ($MgSO_4$), concentrated, and purified on silica with 2:1 hexanes:EtOAc to yield a clear oil (3.6 g, 56% yield).

The ethyl-1-(3-hydroxyl)propyl-4-nitropyrrole-2-carboxylate (1.6 g, 6.8 mmol) was dissolved in 20 mL EtOAc, to which was added 10% (wt/wt) Pd/C (0.3 g, 0.28 mmol). The suspension was added to a Parr bomb, pressurized to 100 psi H_2, and stirred at room temperature for 2h. The mixture was filtered through celite to remove the Pd/C and concentrated. The resulting yellow oil was dissolved in 20 mL DMF, to which was

added DIEA (2.44 mL, 14 mmol) and Boc_2O (1.53 g, 7 mmol), and stirred at room temperature for 2 hrs. The solution was concentrated and purified on 75g silica with 1:1 hexanes:EtOAc to yield a clear oil (1.77g, 84% yield).

This oil (1g, 3.2 mmol) was dissolved in 10 mL CH_2Cl_2, to which was added triethylamine (0.67 mL, 4.9 mmol), and the resulting solution was cooled to 0ºC. The solution was then treated with mesyl chloride (0.33 mL, 4.2 mmol) and DMAP (0.16 g, 1.3 mmol), stirred at 0ºC for 15 min, then allowed to warm to room temperature while stirring for 1 hr. 40 mL CH_2Cl_2 was added and the solution was washed with 1M $KHSO_4$, H_2O, sat. $NaHCO_3$, H_2O, and brine (50 mL each), dried ($MgSO_4$), and concentrated.

The resulting oil was dissolved in DMF (10 mL), to which was added NaN_3 (1g, 16 mmol) and the resulting solution stirred at 55ºC under Ar for 1 hr 30 min. The solution was partitioned between 50 mL brine and 50 mL Et_2O. The aqueous layer was washed with an additional 30 mL Et_2O. The organic portions were combined and washed with 10% (v/v) citric acid, brine, 1M $NaHCO_3$, and brine (2 x 20 mL each), dried, and concentrated to a yellow oil. The oil was dissolved in 10 mL EtOH, to which was added 5 mL KOH, and stirred at 40ºC for 6 hrs. The solution was cooled to room temperature and acidified with 10% (v/v) H_2SO_4 to pH~3. The mixture was extracted with CH_2Cl_2 (3 x 40 mL), washed with brine, dried ($MgSO_4$), and concentrated to yield **46** as a pale yellow solid. ^1H NMR (300 mHz, $CDCl_3$): δ 7.25 (s, 1 H), 6.66 (s, 1 H), 4.40 (t, 2 H), 3.53 (t, 2 H), 1.96 (m), 1.50 (s, 9 H).

Synthesis of 3-(*tert*-butoxycarbonyl), 3'-diamino-*N*-methyldipropylamine (47)

3,3'-diamino-*N*-methyldipropylamine (32.2 mL, 200 mmol) was cooled to 0ºC. A solution of Boc_2O (4.4g, 20 mmol) was prepared in CH_2Cl_2 and added to cooled amine dropwise while stirring for 30 min. The solution was allowed to cool to room temperature while stirring for 1 hr 30 min. The solution was then extracted with 200 mL ¼ sat. $NaHCO_3$ (a volume of sat. $NaHCO_3$ diluted to 25% concentration). The aqueous solution was back-extracted with CH_2Cl_2 (4 x 50 mL), and the organic portions combined and washed with 50 mL sat. $NaHCO_3$ and 50 mL brine. The solution was dried with $MgSO_4$ and concentrated to yield a pearly-white oil, which was azeotroped with toluene (2.5 g, 55% yield). ^{1}H NMR (300 mHz, $CDCl_3$): δ 3.19 (br. q, 2 H), 2.81 (t, 2 H), 2.38 (m, 4 H), 1.65 (m, 4 H), 1.43 (s, 9 H).

Synthesis of Bodipy-Polyamide-DHT Conjugate (50)

Polyamide **44** was prepared as above. DHT-linker molecule **51** was prepared as detailed in Chapter 5. A solution of **51** (17 μmol), DCC (17 μmol), and HOBt (17 μmol) was prepared and allowed to react at room temperature for 30 min. To this was added a solution of **44** (3 μmol) and DIEA (1.6 μL, 10 μmol), and the resulting solution was allowed to react at room temperature for 2hrs. Dp (4 μL) was added to quench excess activated acid and the solution was concentrated with speedvac for 1 hr. The concentrated oil was treated with 1 mL 80% (v/v) TFA in CH_2Cl_2 for 2 hrs at room temperature, then was azeotroped with benzene 3x. The resulting oil was purified by C_{18} reverse-phase preparatory HPLC and lyophilized to yield DHT-polyamide as a white powder.

A solution of Bodipy-FL (3.6 μmol), DCC (3.6 μmol), and HOBt (3.6 μmol) was prepared and allowed to react for 30 min. To it was added the DHT-polyamide (1.8 μmol) and DIEA (0.5 μL, 5 μmol), and the resulting solution was allowed to react at 37ºC for 16 hrs. The solution was purified by C_8 reverse-phase preparatory HPLC and lyophilized to yield **50** as an orange powder. Characterization: MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 2017.0, obsd 2017.8.

References

69) Dickenson, 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.

70) Coull, J.J.; He, G.C.; Melander, C.; Rucker, V.C.; Dervan, P.B. Margolis, D.M. *J. Virol*. **2002**, *76*, 12349.

71) Janssen, S.; Durussel, T.; Laemmli, U.K *Mol. Cell* **2000**, *6*, 999.

72) Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U.K. *Mol. Cell* **2000**, *6*, 1013.

73) Chiang, S.Y.; Bürli, R.W.; Benz, C.C.; Gawron, L.; Scott, G.K.; Dervan P.B.; Beerman, T.A. *J. Biol. Chem*. **2000** *275*, 24246.

74) Dervan, P.B.; Gottesfeld, J.M.; Beerman, T.A. Unpublished results.

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

76) Fischer, P.M.; Krausz, E.; Lane, D.P. *Bioconjugate Chem*. **2001**, *12*, 825.

77) Wunderbaldinger, P.; Josephson, L.; Weissleder, R. *Bioconjugate Chem*. **2002**, *13*, 264.

78) Lee, H.J.; Pardridge, W.M. *Bioconjugate Chem*. **2001**, *12*, 995.

79) Polyakov, V.; Sharma, V.; Dahlheimer, J.L.; Pica, C.M.; Luker, G.D.; Piwnica-Worms, D. *Bioconjugate Chem*. **2000**, *11*, 762.

80) Schwarze, S.R.; Ho, A.; Vocero-Akbani, A.; Dowdy, S.F. *Science* **1999**, *285*, 1569.

81) Mann, D.A.; Frankel, A.D. *EMBO J*. **1991**, *10*, 1733.

82) Benimetskaya, L.; Guzzo-Pernell, N.; Lui, S.-T.; Lai, J.C.H.; Miller, P.; Stein, C.A. *Bioconjugate Chem*. **2002**, *13*, 177.

83) Cutrona, G.; Carpaneto, E.M.; Ulivi, M.; Roncella, S.; Landt, O.; Ferrani, M.; Boffa, L.C. *Nature Biotech*. **2000**, *18*, 300.

84) Brandem, L.J.; Mohamed, A.J.; Smith, C.I.V. *Nature Biotech*. **1999**, *17*, 784.

85) Zanta, M.A.; Belguise-Valladier, P.; Behr, J.-P. *Proc. Natl. Acad. Sci. U.S.A*. **1999**, *96*, 91.

86) Chook, Y.M.; Blobel, G. *Curr. Opin. Struct. Biol*. **2001**, *11*, 703.

87) Conti, E.; Uy, M.; Leighton, L.; Blobel, G.; Kuryan, J. *Cell* **1998**, *94*, 193.

88) Schwarze, S.R.; Dowdy, S.F. *Trends in Pharmacological Sciences* **2000**, *21*, 45.

89) Efthymiadis, A.; Briggs, L.; Jans, D. *J. Biol. Chem*. **1998**, *273*, 1623.

90) Wender, P.A.; Mitchell, D.J.; Pattabiraman, K.; Pelkey, E.T.; Steinman, L.; Rothbard, J.B. *Proc. Natl. Acad. Sci. U.S.A*. **2000**, *97*, 13003.

91) Ho, A.; Schwarze, S.R.; Mermelstein, S.J.; Waksman, G.; Dowdy, S.F. *Cancer Res*. **2001**, *61*, 474.

92) Trauger, J.W.; Dervan, P.B. *Methods Enzymol*. **2001**, *340*, 450.

93) Arora, P.S. Postdoctoral Report **2002**.

94) Rothbard, J.B.; Kreider, E.; VanDeusen, C.L.; Wright, L.; Wylie, B.L; Wender, P.A. *J. Med. Chem*. **2002**, *45*, 3612.

95) Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1996**, *118*, 6141.

96) Belitsky, J.M.; Nguyen, D.H.; Wurtz, N.R.; Dervan, P.B. *Bioorg. Med. Chem*. **2002**, *10*, 2767.

97) Wurtz, N.R. California Institute of Technology: Pasadena, CA, 2002.

98) Boffa, L.C.; Scarfi, S.; Mariani, M.R.; Damonte, G.; Allfrey, V.G.; Benatti, U.; Morris, P.L. *Cancer Res*. **2000**, *60*, 2258.

99) Zhang, X.; Simmons, C.G.; Corey, D.R. *Bioorg. Med. Chem. Lett*. **2001**, *11*, 1269.

100) Leamon, C.P.; Low, P.S. *J. Biol. Chem*. **1992**, *267*, 24966.

101) Leamon, C.P.; Low, P.S. *Proc. Natl. Acad. Sci. U.S.A*. **1991**, *88*, 5572.

102) Luthman, H.; Magnusson, G. *Nucleic Acids Res*. **1983**, *11*, 1295.

103) Larsen, A.; Escargueil, A.; Skladanowski, A. *Pharmacol. Ther*. **2000**, *85*, 217.

104) Hurwitz, S.; Terashima, M.; Mizunuma, N.; Slapak, C. *Blood* **1997**, *89*, 3745.

105) Belitsky, J.M. California Institute of Technology: Pasadena, CA, 2002.

106) Keller, O.; Rudinger, J. *Helv. Chim. Acta* **1975**, *58*, 531.

107) Foister, S. California Institute of Technology: Pasadena, CA, 2003.

Chapter 4

Nuclear Uptake of Polyamide-Fluorophore Conjugates in Mammalian Cell Lines

Chapter 4A: *The text of this chapter was taken in part from a manuscript coauthored with Benjamin S. Edelson, Nicholas G. Nickols, and Professor Peter B. Dervan (Caltech).*

(Best, T.P.; Edelson, B.S.; Nickols, N.G.; Dervan, P.B. "Nuclear Localization of Pyrrole-Imidazole Polyamide-Fluorescein Conjugates in Cell Culture" *Proc. Natl. Acad. Sci. U.S.A*. **2003**, *100*, 12063.)

Chapter 4B: *The text of this chapter was taken in part from a manuscript coauthored with Benjamin S. Edelson, Bogdan Olenyuk, Nicholas G. Nickols, Raymond M. Doss, Shane Foister, Alexander Heckel, and Professor Peter B. Dervan (Caltech).*

(Edelson, B.S.; Best, T.P.; Olenyuk, B.; Nickols, N.G.; Doss, R.M.; Foister, S.; Heckel, A.; Dervan, P.B. "Influence of Structural Variation on Nuclear Localization of DNA-Binding Polyamide-Fluorophore Conjugates" *Nuc. Acids Res.* **2004**, *32*, 2802.)

Chapter 4A

Polyamide-FITC Conjugates for Cellular Uptake Studies

Abstract

A series of hairpin pyrrole-imidazole polyamide-fluorescein conjugates were synthesized and assayed for cellular localization. Thirteen cell lines, representing eleven human cancers, one human transformed kidney cell line, and one murine leukemia cell line, were treated with 5 μM polyamide-fluorescein conjugates for 10-14 hrs, then imaged by confocal laser scanning microscopy. A conjugate containing a β-alanine residue at the C-terminus of the polyamide moiety showed no nuclear localization, while an analogous compound lacking the β-alanine residue was strongly localized in the nuclei of all cell lines tested. The localization profiles of several other conjugates suggest that pyrrole-imidazole sequence and content, dye choice and position, linker composition, and molecular weight are determinants of nuclear localization. The attachment of fluorescein to the C-terminus of a hairpin polyamide results in an approximate 10-fold reduction in DNA-binding affinity, with no loss of binding specificity with reference to mismatch binding sites.

Introduction

Small molecules that preferentially bind to predetermined DNA sequences inside living cells would be useful tools in molecular biology, and perhaps human medicine. The effectiveness of these small molecules requires not only that they bind to chromosomal DNA site-specifically, but also that they be permeable to the outer membrane and gain access to the nucleus of living cells. Polyamides containing the aromatic amino acids *N*-methylpyrrole (Py), *N*-methylimidazole (Im) and *N*-methyl-3-hydroxypyrrole (Hp) bind DNA with affinities and specificities comparable to naturally occurring DNA-binding proteins.[1,2] A set of pairing rules describes the interactions between pairs of these heterocyclic rings and Watson-Crick base pairs within the minor groove: Im/Py is specific for G•C, Hp/Py is specific for T•A and Py/Py binds both A•T and T•A. Exploitation of polyamides to target sequences of biological interest has yielded results in a number of cell-free systems.[3-8]

Extension of these *in vitro* biological results to cellular systems has proven to be cell-type dependent. Polyamides exhibited biological effects in primary human lymphocytes,[9] human cultured cell lines,[10] and when fed to *Drosophila* embryos.[11,12] However, attempts to inhibit the transcription of endogenous genes in cell lines other than insect or T-lymphocytes have met with little success. For example, polyamides that down-regulate transcription of the HER2/*neu* gene in cell-free experiments display no activity in HER2-overexpressing SK-BR-3 cells.[13]

To determine if these results were due to poor cellular uptake or nuclear localization, a series of polyamides incorporating the fluorophore Bodipy FL was synthesized. Intracellular distribution of these molecules in several cell lines was then

determined by confocal laser scanning microscopy.[14] Cells which demonstrated robust responses to polyamides, such as T-lymphocyte derivatives, showed staining throughout the cells, including the nucleus. In other cell lines studied, however, treatment with polyamide-Bodipy conjugates produced a punctate staining pattern in the cytoplasm, with no observable signal in the nuclei. Bashkin and coworkers recently reported that an eight-ring polyamide-Bodipy conjugate colocalized with LysoTracker Red DND 99 (a lysosome and *trans*-Golgi stain, **23**) in several cell lines, indicating that the punctate staining pattern was presumably due to trapping of the polyamide in acidic vesicles.[15] In contrast, an eight-ring polyamide-fluorescein conjugate, **1**, was shown to accumulate in the nuclei of HCT-116 human colon cancer cells.

We have found that a similar eight-ring polyamide-fluorescein conjugate, **2**, with a single Py to Im change, is excluded from the nuclei of thirteen different mammalian cell lines, whereas removal of the β-Ala residue in the linker, affording **3**, enables nuclear localization in all of these cell lines, with no obvious toxicity. This raises the issue: what are the molecular determinants—fluorophore, position of attachment, linker composition, polyamide sequence and size—of uptake and nuclear localization in cultured cells? Understanding nuclear accessibility in a wide variety of living cells is a minimum first step toward chemical regulation of gene expression with this class of molecules. A further issue will be whether polyamides modified for optimal cellular and nuclear uptake retain favorable DNA-binding affinity and sequence specificity. We have synthesized twenty-two polyamide-fluorophore conjugates with incremental changes in structure and examined their intracellular distribution in thirteen cell lines.

Results

Structures for all of the compounds synthesized are listed in Figure 4.1. Sample images for compounds **2** and **3** in two cell lines are shown in Figure 4.2a. To show that compound **3** localizes to the nuclei of live cells, samples were treated with **3**, the nuclear stain Hoechst 33342, and the dead-cell stain Sytox Orange (Figure 4.2b). The uptake characteristics of compounds **1-22** were examined in thirteen cell lines by confocal microscopy, and each sample was rated qualitatively for the extent of nuclear localization (Figure 4.3).

The molecules showing the highest degree of nuclear uptake in most cell lines stained nuclei very brightly with reference to the background fluorescence caused by fluorescent agent in the medium. Agents **1**, **3**, **5**, **6**, **11-14**, and **22** showed such high levels of uptake in many cell lines. Many compounds exhibited a very scattered uptake profile in the series of cells studied, and compounds **2**, **4**, **15**, and **21**, showed no significant nuclear staining in any cell line. Of the entries in Figure 4.3 indicating poor nuclear uptake properties, most reflect lysosomal staining, as indicated by costaining with LysoTracker Red DND-99, as well as fluoresence in the intercellular medium. Only compound **2** was found exclusively in the medium, showing no uptake in cells.

The DNA-binding properties of **3** were assessed by DNase I footprinting titrations on the plasmid pDEH9, which bears the 5'-TGGTCA-3' match site and discrete single base-pair mismatch sites. Conjugate **3** bound to the match site with a K_a of 1.6 (± 0.3) x 10^9 M^{-1}, and showed specificity over mismatch sites by >100-fold (Figure 4.4).

To investigate the energy dependence of the cellular uptake mechanism of compound **3**, HeLa cells growing under normal conditions were incubated for 30 min in

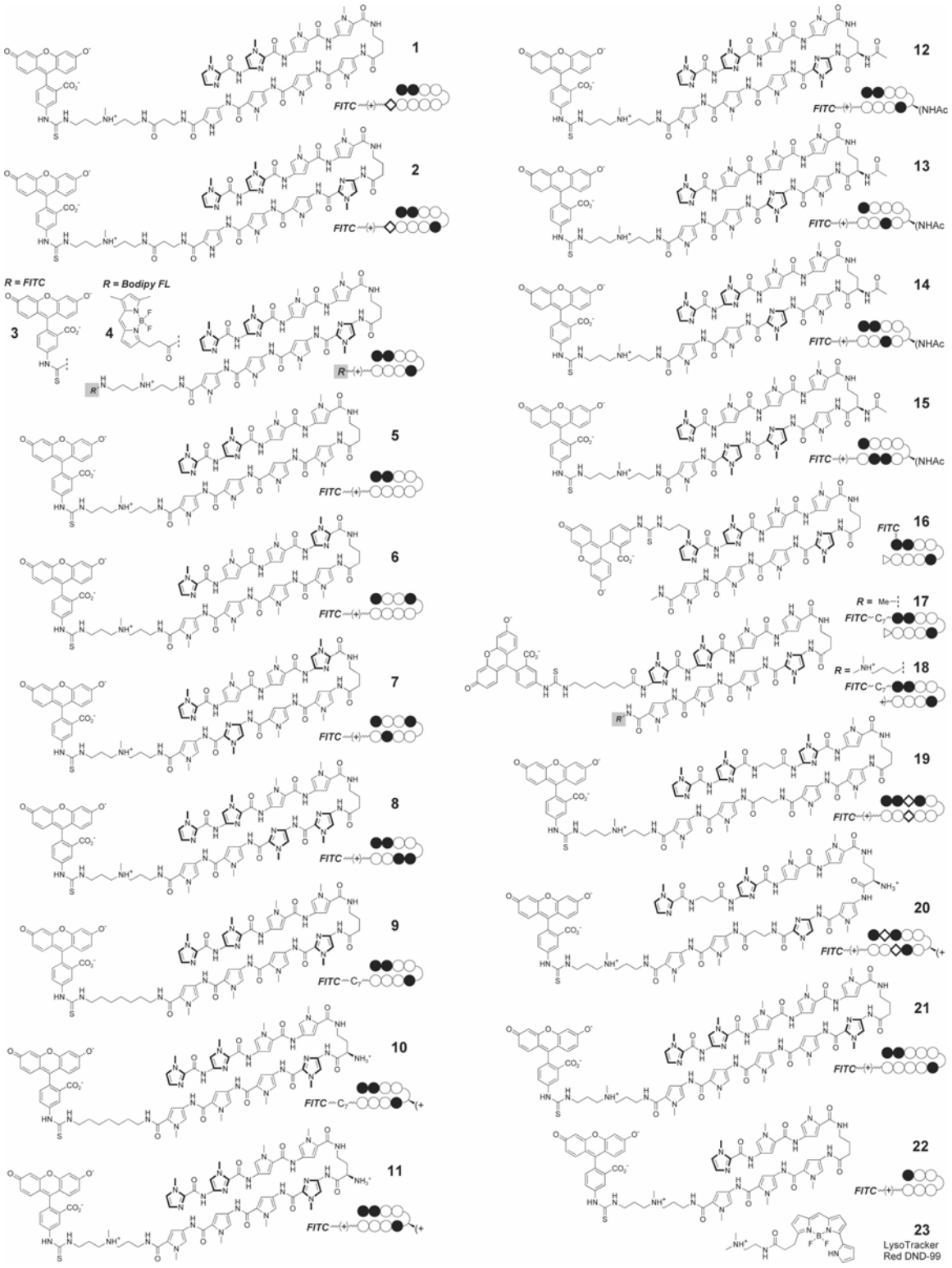


Figure 4.1 Structures of compounds used in uptake experiments.

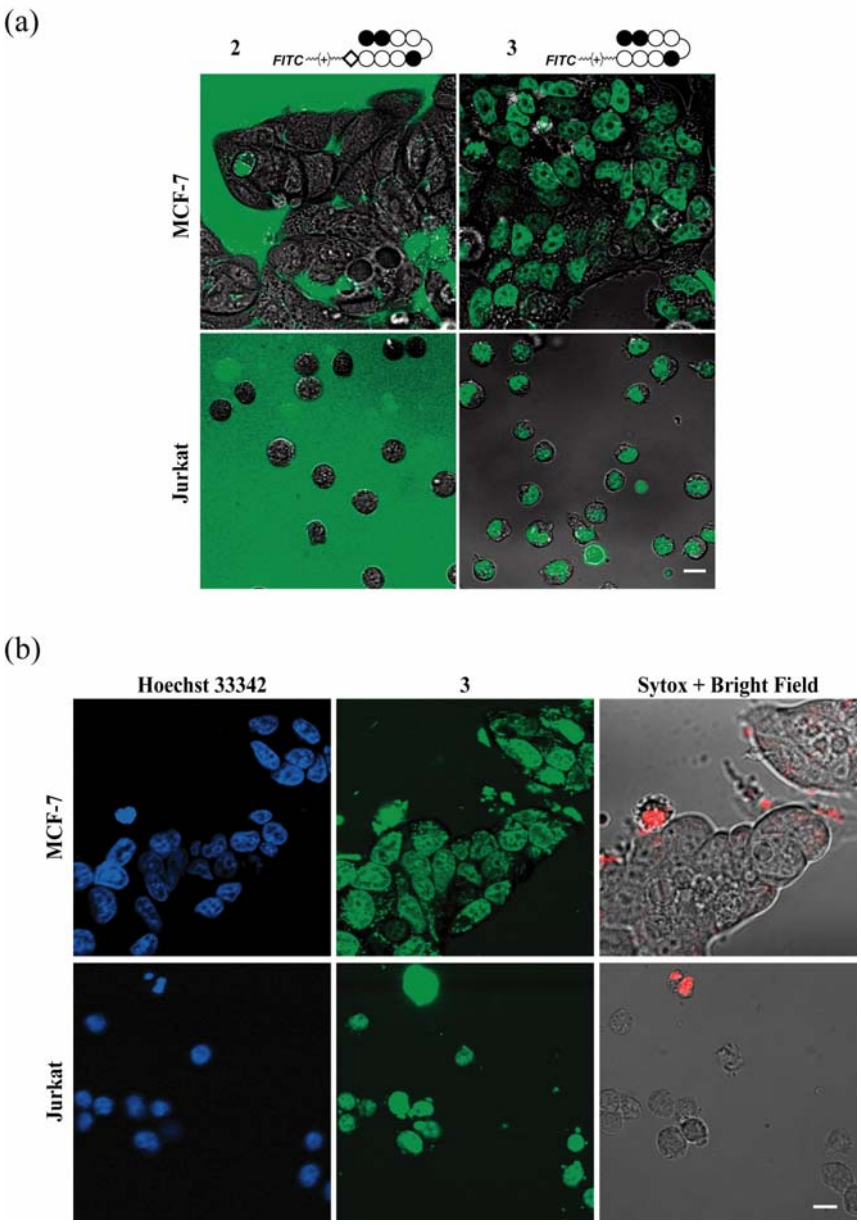


Figure 4.2 Cellular localization of polyamide-fluorescein conjugates. (a) (*Top*) Adherent MCF-7 cells were treated with compound **2** (*Upper left*) or **3** (*Upper right*) for 10-14 h at 5 μM. Compound **2** was excluded from the cells entirely, while compound **3** localized to the nucleus. (*Bottom*) Suspended Jurkat cells were similarly treated, and show similar results. (b) Colocalization of polyamide **3** and Hoechst in live cells, imaged using sequential single- and two-photon excitation. (*Top*) MCF-7 cells were treated with the nuclear stain Hoechst 33342 (15 μM), compound **3** (5 μM), and the dead cell stain Sytox Orange (0.5 μM). Fluorescence signals from Hoechst (*Upper left*, blue) and compound **3** (*Upper center*, green) colocalize in cell nuclei. (*Upper right*) Overlay of the visible light image (grayscale) and the Sytox Orange fluorescence image (red), indicating that the majority of cells are alive. (*Bottom*) Jurkat cells were treated similarly, and show similar results. (Scale bar = 10 μm.)

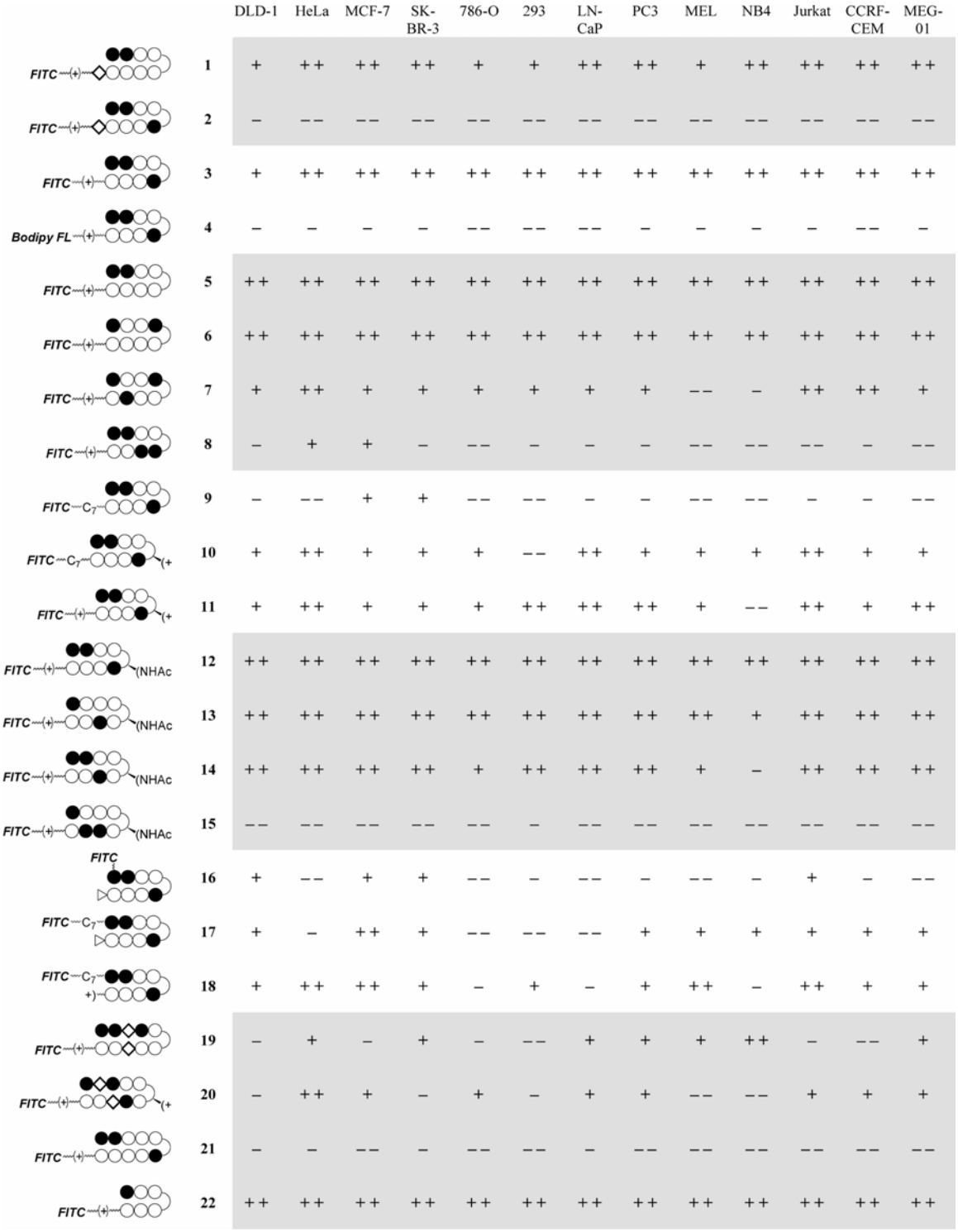


Figure 4.3 Uptake profile of compounds **1**-**22** in 13 cell lines: "+ +" Indicates nuclear staining exceeds that of the medium; "+" indicates nuclear staining ≤ that of the medium, but still prominent; "-" indicates very little nuclear staining, with the most fluorescence seen in the cytoplasm and/or medium; "--" indicates no nuclear staining.

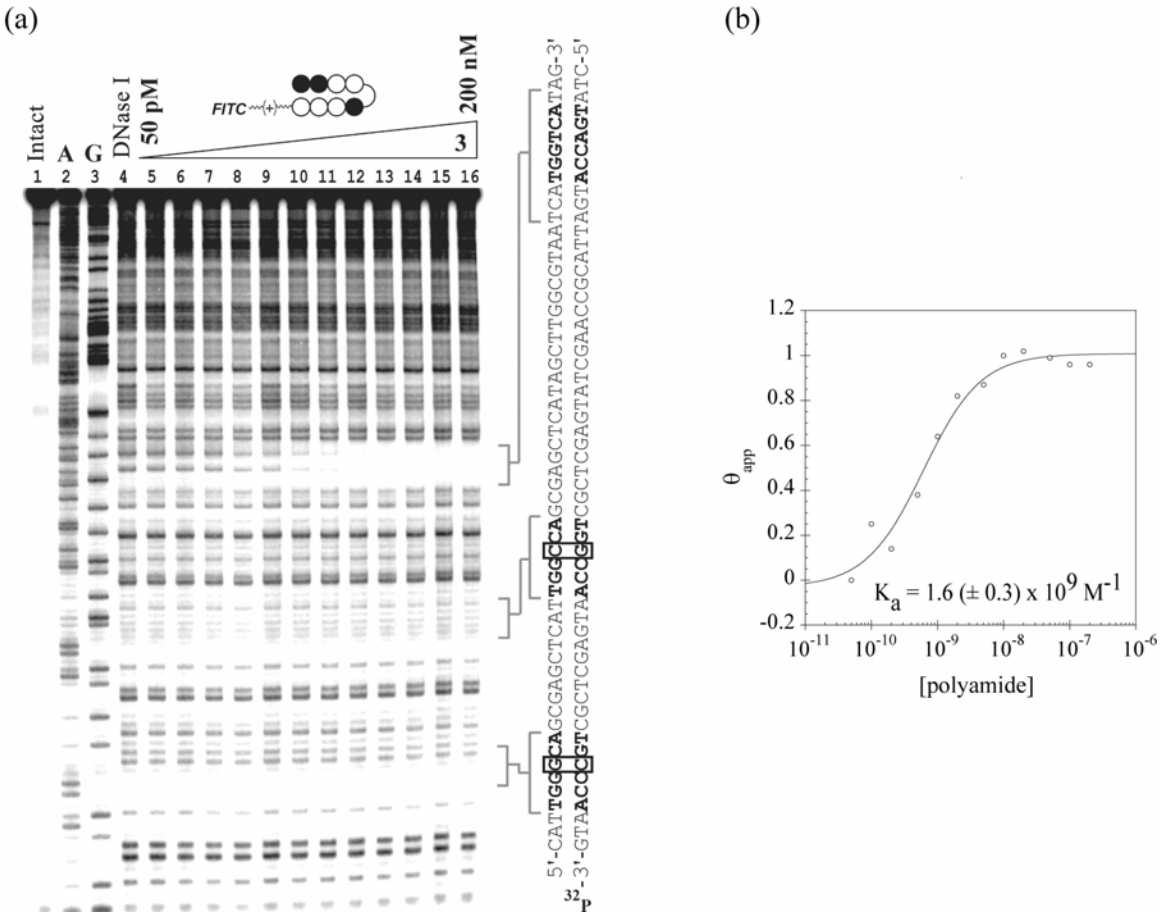


Figure 4.4 Quantitative DNase I footprinting titration. (a) Compound **3** binds the 3'-TGGTCA-5' site with an affinity K_a = 1.6 (± 0.3) x 10^9 M^{-1}. Compound **3** does not bind the single base-pair mismatch sites shown at concentrations ≤ 200 nM. (b) DNA binding isotherm for compound **3**.

either normal DMEM medium or inhibitory DMEM medium, then treated with **3** for 1 hr prior to confocal imaging (Figure 4.5). The cells growing in normal medium showed clear nuclear staining, whereas the cells growing in inhibitory medium displayed very little to no discernable staining. Subsequent washing and replacement of the inhibitory medium with normal medium (supplemented with 5 μM **3**), resulted in nuclear staining after 1 h, comparable to that seen in the sample grown continuously in the normal medium.

It had been shown previously that polyamide-Bodipy conjugates stain the nuclei of T-lymphocytes, but no other cell type tested, and most commonly produced a punctate cytoplasmic staining pattern.[14] Our studies indicated that a polyamide-fluorescein conjugate, **2**, uniformly proved refractory to nuclear uptake in several human cancer cell lines. Elimination of the β-alanine residue at the carboxy-terminal end of the polyamide afforded compound **3**, which, surprisingly, showed excellent nuclear staining properties in the cell lines examined. Attachment of the Bodipy-FL fluorophore to the polyamide precursor of compound **3** provided compound **4**. This molecule showed no nuclear staining, sequestering itself in cytoplasmic vesicles, indicating that some characteristic of polyamide-Bodipy conjugates differing from that of polyamide-FITC conjugates, and not the C-terminal β-alanine, prevents their trafficking into the nucleus. It is interesting to note that the structure of LysoTracker Red DND-99 (**23**) includes both a tertiary alkyl amine, similar to that often used in hairpin polyamide tails, and a Bodipy moiety.

It became our intent to explore the structure-space of polyamide-fluorophore conjugates to overcome cellular exclusion and lysosomal trapping, allowing the polyamides to travel to the nucleus. To explore the criteria that permit uptake of **3**, and

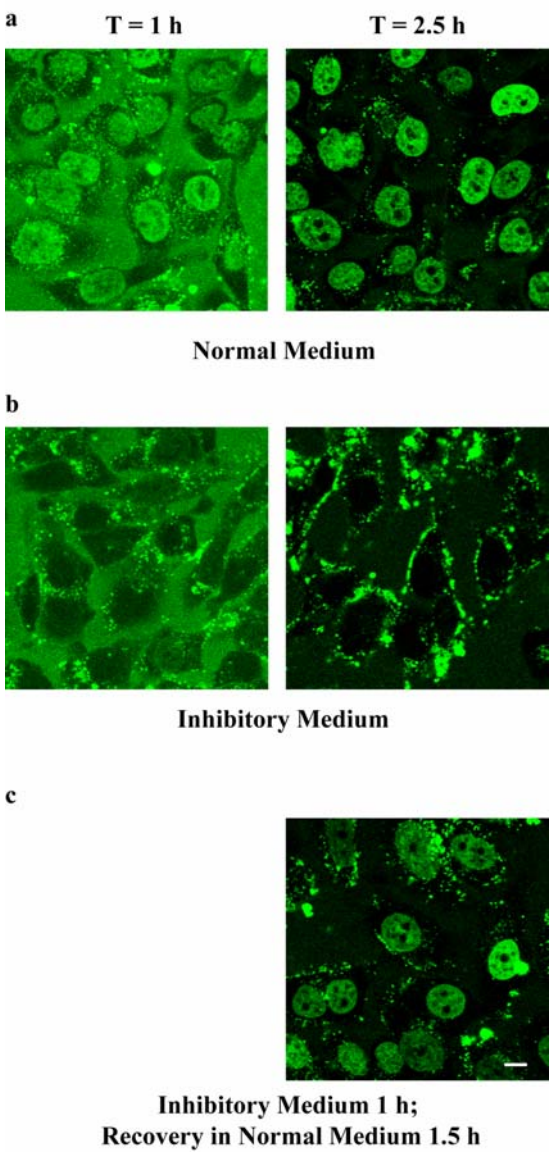


Figure 4.5 Energy dependence of nuclear localization of compound **3**. (a) HeLa cells were washed, incubated in normal DMEM medium for 30 min, supplemented with 5 mM **3**, incubated for 1 h, and imaged, showing roughly equivalent amounts of compound localized both in the nucleus and in the medium (*left*). The cells were incubated for a further 1 h 30 min, and then imaged once more (*right*), showing exclusively nuclear localization. (b) HeLa cells were washed, incubated in inhibitory DMEM medium for 30 min, supplemented with 5 mM **3**, incubated for 1 h, and imaged, showing localization at the cellular membranes, as well as in the medium (*left*). The cells were incubated for a further 1 h 30 min, and then imaged once more (*right*), showing exclusively membranous localization. (c) Recovery experiment: HeLa cells treated as in b were washed 2x with normal DMEM medium, incubated for 1h 30 min in normal DMEM medium supplemented with 5 mM **3**, then imaged, showing recovery of nuclear localization upon replacement of inhibitory medium with normal medium. (Scale bar = 10 μm.)

that prevent that of **4**, several variations on the structure of the compound were made and their effects determined by confocal microscopy.

Wishing to explore the effect of Py/Im sequence and content of hairpin polyamides on cellular trafficking, we synthesized compounds **5**-**8**. Compounds **5** and **6**, both containing two imidazole residues, showed a high degree of nuclear staining in all cell lines studied. Three-imidazole compound **7** showed intermediate levels of nuclear staining in the cell lines studied. The nuclear localization of four-imidazole compound **8** was quite poor in all cell lines tested. The difference in nuclear staining levels exhibited by compounds **3** and **7** shows that Py/Im sequence alone is an important determinant of nuclear uptake, though overall content (in terms of the number of Py and Im residues) may also be a factor.

To explore the effect of the positively-charged linker on nuclear uptake, compound **9** was synthesized. This agent showed nearly global abrogation of nuclear uptake efficiency versus the analog containing a tertiary amine. Further, substitution of the γ-aminobutyric acid turn (γ-turn) of **9** with the [(*R*)-α-amino]-γ-diaminobutyric acid turn (^{H_2N}γ-turn) provided **10**, which restored most of the nuclear uptake properties of **3**. This suggests that the overall charge of the molecule, and perhaps the placement of that charge, are important variables in nuclear uptake of polyamide-dye conjugates.

The ^{H_2N}γ-turn is a structural element commonly included in polyamide design. Replacement of the γ-turn of **3** with the ^{H_2N}γ-turn provided **11**. This molecule showed nuclear staining in all cell lines tested, save NB4, though often to a lesser degree than **3**. It is unclear whether this reduction in uptake efficiency is due to an increase in the overall

positive charge of the molecule, a more branched structure than the linear γ-linked hairpin, or the positioning of a positive charge medial in the molecule.

Selective acetylation of the polyamide precursor to **11**, followed by FITC conjugation, provided **12**. This molecule exhibited excellent nuclear uptake, as good or better than both **11** and **3**. The excellent uptake of **12** argues against branching as a negative determinant of nuclear staining. This result also prompted us to synthesize **13**-**15** to probe the generality and flexibility of turn acetylation across several polyamide sequences. The uptake of **13** and **14** was mostly nuclear. Compound **15**, on the other hand, was a very poor nuclear stain. This result reaffirms the importance of polyamide sequence on nuclear uptake.

We next synthesized several conjugates to explore the uptake effects of different points of attachment of the fluorophore-linker moiety to the polyamide. An *N*-propylamine linkage from the terminal imidazole and a methylamide tail were incorporated into **16**, which showed poor nuclear uptake in nearly all cell lines tested. We thought it possible that the poor uptake properties of conjugate **16** was due to some non-linearity introduced into the overall structure of the molecule by the linkage at the 1-position of the N-terminal ring of the polyamide. Consequently, conjugates **17** and **18** were synthesized, appending the linker-fluorophore moiety to the N-terminal amine. Although compound **17** has an overall structure and charge distribution very similar to that of **9**, it possesses better nuclear uptake properties, particularly in suspended cell lines. Subsequent addition of a positive charge at the C-terminal end of the polyamide through employment of an *N*,*N*-dimethylaminopropylamine tail (compound **18**) boosts uptake in all cell lines studied, save NB4. Once again, this suggests that the addition or deletion of

a single charge may (but not necessarily will) have a strong effect on nuclear uptake. In fact, a comparison of compounds **3**, **10**, and **18** (differing in the placement of a positive charge) with **9** and **17** (which lack a positive charge) seems to indicate that the positioning of this charge is less influential than its presence.

Another common structural feature utilized in DNA-binding polyamides is the addition of one or more β-alanine residues. This amino acid is often used to relax overcurvature of a polyamide backbone with reference to the minor groove of DNA and permits the targeting of longer sequences than that of eight-ring hairpins. Compound **19**, which includes this structural element, and **20**, which includes both β-alanines and the $^{H_2N}\gamma$-turn, showed fair to poor nuclear uptake in the cell line series. The reduced uptake levels could be due to the increased molecular flexibility imparted by the β-alanine residues, increased molecular weight, decreased recognition by a cellular import protein, or some combination of these criteria.

In order to determine whether or not the added size imparted to **19** and **20** by the β-alanine residues was likely a major contributor to their poor uptake, the ten-ring compound **21** was synthesized. Its ubiquitously poor nuclear uptake suggests that molecular weight may, indeed, play a prominent role in uptake character. The excellent uptake character of the six-ring compound **22** furthers this hypothesis, though Py/Im sequence in the 6-ring and 10-ring contexts is also likely to be important.

This study demonstrates the cellular localization profile of a host of polyamide-dye conjugates in a wide variety of cell lines. In general, compounds exhibiting good nuclear uptake properties have several common elements: an eight-ring polyamide DNA-binding domain, one or more positive charges incorporated within either the linker or the

turn residue, and a conjugated fluorescein fluorophore. This study also demonstrates that each cell line possesses a unique uptake profile for the panel of compounds presented to it. These profiles will be important in choosing a cell line and compound architecture appropriate to a given experiment.

Conjugation of the fluorophore to the polyamide DNA-recognition domain results in an approximately 10-fold reduction in DNA-binding affinity as compared with the parent polyamide, with retention of binding specificity over mismatch sites. This quality, along with the nuclear uptake results, suggests that fluorophore-conjugated polyamides may be employed directly in experiments designed to take place in living mammalian cells.

Clearly, there are many criteria at work, each one having its share of influence upon nuclear uptake of polyamide-dye conjugates. The extension of the structure-space of compounds known to both bind chromosomal DNA specifically and with high affinity, as well as to traffic to the nucleus of living cells is the object of current investigations. The illumination of these possibilities will permit the further study of transcriptional regulation in living systems.

Experimental

Chemicals

Polyamides **1** and **2** were prepared by solid-phase methods on Boc-β-alanine-PAM-resin (Peptides International; Louisville, KY).[16] All other polyamides were synthesized by solid phase methods on the Kaiser oxime resin (Nova Biochem; Laufelfingen, Switzerland).[17] After cleavage with the appropriate amine and reverse-

phase HPLC purification, polyamides were allowed to react at room temperature for ~3 h at ~0.01 M in *N*,*N*-dimethylformamide with fluorescein isothiocyanate (FITC; compounds **1**-**3** and **5**-**22**) or the *N*-hydroxysuccinimidyl ester of BODIPY-FL (**4**), as well as 20 eq of Hünig's base, to yield polyamide-dye conjugates. The purity and identity of the dye conjugates were verified by analytical HPLC, UV-vis spectroscopy and MALDI-TOF mass spectrometry. All fluorescent dye reagents were from Molecular Probes. Chemicals not otherwise specified were from Aldrich.

Mass Spectra

MALDI mass spectra were obtained on a Voyager De PRO time-of-flight mass spectrometer (Applied BioSystems) operated at an accelerating voltage of +20 kV. Samples were applied to the target in an α-cyanohydroxycinnamic acid matrix. The mass spectrometer was calibrated with a calibration mixture provided by the instrument manufacturer. The mass spectra data are summarized in Table 4.1.

Preparation of Polyamide Solutions

To make 100 μM solutions of each compound, dry HPLC-purified aliquots of each conjugate were dissolved to 1-10 mM in DMSO, then diluted with PBS buffer (pH 7.4) or 18 MΩ cm water to ~200 μM. A small portion of this solution was diluted to ~10 μM and the concentration determined by UV quantitation at 310 nm (compounds **1**-**20**: ε = 69500; compound **21**: ε = 86875; compound **22**: ε = 52125) the concentrated solution was then diluted to 100 μM. The final concentration of DMSO in any sample was less than 1.5%, usually less than 0.1%.

Cell Cultures

The human cancer cell lines DLD-1, MCF-7, 786-O, LNCaP, PC3, MEG-01, NB4, Jurkat, and CCRF-CEM were cultured in a 5% CO_2 atmosphere at 37°C in supplemented RPMI-1640 medium. The human cancer cell line HeLa, the murine leukemia cell line MEL, and the transformed human kidney cell line 293 were grown as above in supplemented Dulbecco's Modified Eagle medium (DMEM). The human cancer cell line SK-BR-3 was cultured as above in supplemented McCoy's medium. All media were supplemented with 10% fetal bovine serum (Irvine Scientific) and 1% penicillin/streptomycin solution (Sigma).

Confocal Microscopy

Adherent cell lines were trypsinized for 5-10 min at 37°C, centrifuged for 5 min at 5°C at 2000 rpm, and resuspended in fresh medium to a concentration of 1.25 x 10^6 cells/mL. Suspended cell lines were centrifuged and resuspended in fresh medium to the same concentration. Incubations were performed by adding 150 μL cells into culture dishes equipped with glass bottoms for direct imaging (MatTek Corporation). Adherent cells were grown in the glass-bottom culture dishes for 24 h. The medium was then removed and replaced with 142.5 μL of fresh medium. Then 7.5 μL of the 100 μM polyamide solution was added and the cells were incubated in a 5% CO_2 atmosphere at 37°C for 10-14 h. Suspended cell line samples were prepared in a similar fashion, omitting trypsinization. These samples were then incubated as above for 10-14 h. Imaging was performed with a Zeiss LSM 5 Pascal inverted laser scanning microscope.

Images were line-averaged 4, 8, or 16 times and were obtained at a 0.8 μs/pixel scanning rate. Polyamide-fluorescein conjugate fluorescence and visible light images were obtained using 488 nm laser excitation with a standard fluorescein filterset and a pinhole of 181 μm. For two-color experiments with LysoTracker Red DND-99, cells were treated with 0.75 μL of a 100 μM DMSO stock solution, affording a final dye concentration of 500 nM LysoTracker. After ~10 min incubation with LysoTracker, images were obtained by simultaneous 488 nm and 543 nm laser excitation and a filterset appropriate for simultaneous visualization of fluorescein and rhodamine.

Hoechst colocalization experiments were performed on a Zeiss LSM 510 META NLO upright laser scanning microscope with a Coherent Chameleon 2-photon laser. For the MCF-7 cell line, the cells were grown in glass-bottom culture dishes in a volume of 150 μL and treated with polyamide **3** as usual. Approximately 30 min prior to imaging, 0.75 μL of Sytox Orange solution (100 μM in DMSO) and 3 μL of Hoechst 33342 solution (750 μM in sterile water) were added to the cell medium, affording final dye concentrations of 500 nM Sytox and 15 μM Hoechst. Immediately prior to imaging, a coverslip was placed over the cells, and the dish inverted. Three sequential images were obtained in succession, at a 0.8 μs/pixel scanning rate. The images were line-averaged 8 times. The Sytox and visible light overlay was collected using 543 nm laser excitation with a standard rhodamine filterset and a pinhole of 499 μm. The polyamide-fluorescein conjugate was imaged using 488 nm laser excitation with a standard fluorescein filterset and a pinhole of 181 μm. Hoechst was imaged using 800 nm laser two-photon excitation with an HFT KP 680 dichroic and a 390-465 nm bandpass filter with a fully open pinhole.

For the Jurkat cell line, cells were treated as above, but immediately prior to imaging, 30 μL of the cell solution was placed between a clean glass slide and a coverslip separated by a coverslip spacer. Images were collected in multitrack mode, at a 0.64 μs/pixel scanning rate. The images were line-averaged 8 times. Sytox and Hoechst were imaged using 800 nm laser two-photon excitation with an HFT KP 680 main dichroic, an NFT 545 secondary dichroic, a 565-615 nm bandpass filter for the Sytox signal (fully open pinhole), and a 390-465 bandpass filter for the Hoechst signal (363 μm pinhole). The polyamide-fluorescein conjugate and visible light images were obtained using 488 nm laser excitation with a standard fluorescein filterset and a pinhole of 249 μm.

Energy Dependence Experiments

Inhibitory medium was prepared by supplementing glucose- and sodium pyruvate-free DMEM (Gibco #1196025) with 2-deoxyglucose (6 mM) and sodium azide (10 mM).[18] Cells were grown, trypsinized, resuspended, plated and incubated as above. After 24 h of growth at 37°C in 5% CO_2, the medium was removed and replaced with either 142.5 μL fresh normal DMEM medium or 142.5 μL inhibitory DMEM medium. The cells were incubated for 30 min, then treated with 7.5 μL of 100 μM compound **3**. The cells were then incubated for 1 h, followed by confocal imaging as above. Samples in inhibitory medium were then treated by removal of the medium, washing and removal of 200 μL of normal medium, replacement with 142.5 μL of normal medium and addition of 7.5 μL of 100 μM compound **3**. These samples were incubated for 1 h and then imaged once more.

DNase I Footprinting Titration Experiments

A 3'-[^{32}P]-labeled restriction fragment from the plasmid pDEH9 was generated in accordance with standard protocols and isolated by nondenaturing gel electrophoresis.[19,20]

1	H-FITC	1265.6	1265.8
2	H-FITC	1266.6	1267.3
3	H	1584.6	1584.5
4	Na	1490.7	1490.7
5	H	1583.6	1583.6
6	H-FITC	1194.6	1194.6
7	H	1584.6	1585.1
8	H	1585.6	1585.2
9	H	1569.6	1569.6
10	H	1584.6	1584.7
11	H-FITC	1210.6	1210.6
12	H	1641.6	1641.5
13	H	1640.6	1641.4
14	H	1641.6	1642.4
15	H	1641.6	1641.5
16	H	1513.5	1513.3
17	H	1634.6	1634.4
18	H-FITC	1294.6	1294.8
19	H	1797.7	1797.9
20	H	1812.7	1812.8
21	H –FITC	1461.7	1461.8
22	H	1339.5	1339.6

Table 4.1 Mass spectra for compounds **1**-**22**.

Acknowledgements

We are grateful to the National Institutes of Health for support for predoctoral support for T.P.B. under T32-GM08501 and to the Howard Hughes Medical Institute for a fellowship to B.S.E. Mass spectral analyses were performed in the Mass Spectrometry Laboratory of the Division of Chemistry and Chemical Engineering of Caltech, supported in part by NSF MRSEC program. We thank M. Waring for helpful discussions.

References

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

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

110) Dickinson, L.A.; Trauger, J.W.; Baird, E.E.; Ghazal, P.; Dervan, P.B.; Gottesfeld, J.M. *Biochemistry* **1999**, *38*, 10801.

111) Gottesfeld, J.M.; Turner, J.M.; Dervan, P.B. *Gene Expression* **2000**, *9*, 77.

112) Wang, C.C.C.; Dervan, P.B. *J. Am. Chem. Soc*. **2001**, *123*, 8657.

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

114) Gottesfeld, J.M.; Melander, C.; Suto, R.K.; Raviol, H.; Luger, K.; Dervan, P.B. *J. Mol. Biol*. **2001**, *309*, 615.

115) Ehley, J.A.; Melander, C.; Herman, D.; Baird, E.E.; Ferguson, H.A.; Goodrich, J.A.; Dervan, P.B.; Gottesfeld, J.M. *Mol. Cell. Biol*. **2002**, *22*, 1723.

116) 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.

117) Coull, J.J.; He, G.C.; Melander, C.; Rucker, V.C.; Dervan, P.B.; Margolis, D.M. *J. Virol*. **2002**, *76*, 12349.

118) Janssen, S.; Durussel, T.; Laemmli, U.K. *Mol. Cell* **2000**, *6*, 999.

119) Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U.K. *Mol. Cell* **2000**, *6*, 1013.

120) Chiang, S.Y.; Bürli, R.W.; Benz, C.C.; Gawron, L.; Scott, G.K.; Dervan, P.B.; Beerman, T.A. *J. Biol. Chem*. **2000**, *275*, 24246.

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

122) Crowley, K.S.; Phillion, D.P.; Woodard, S.S.; Schweitzer, B.A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J.K. *Bioorg. Med. Chem. Lett*. **2003**, *13*, 1565.

123) Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1996**, *118*, 6141.

124) Belitsky, J.M.; Nguyen, D.H.; Wurtz, N.R.; Dervan, P.B. *Bioorg. Med. Chem*. **2002**, *10*, 2767.

125) Richardson, W.D.; Mills, A.D.; Dilworth, S.M.; Laskey, R.A.; Dingwall, C. *Cell* **1998**, *52*, 655.

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

127) Trauger, J.W.; Dervan, P.B. *Methods Enzymol*. **2001**, *340*, 450.

Chapter 4B

Influence of Structural Variation on Nuclear Localization of DNA-Binding Polyamide-Fluorophore Conjugates

Abstract

A pivotal step forward in chemical approaches to controlling gene expression is the development of sequence-specific DNA-binding molecules that can enter live cells and traffic to nuclei unaided. DNA-binding polyamides are a class of programmable, sequence-specific small molecules that have been shown to influence a wide variety of protein-DNA interactions. We have synthesized over 100 polyamide-fluorophore conjugates and assayed their nuclear uptake profiles in thirteen mammalian cell lines. The compiled dataset, comprising 1300 entries, establishes a benchmark for the nuclear localization of polyamide-dye conjugates.

Compounds in this series were chosen to provide systematic variation in several structural variables, including dye composition and placement, molecular weight, charge, ordering of the aromatic and aliphatic amino-acid building blocks, and overall shape. Nuclear uptake does not appear to be correlated with polyamide molecular weight or with the number of imidazole residues, although the positions of imidazole residues affect nuclear access properties significantly. Generally negative determinants for nuclear access include the presence of a β-Ala-tail residue and the lack of a cationic alkyl amine moiety, whereas the presence of an acetylated 2,4-diaminobutyric acid-turn is a positive factor for nuclear localization. We discuss implications of this data on the design of polyamide-dye conjugates for use in biological systems.

Introduction

Cell-permeable small molecules that preferentially bind to predetermined DNA sequences inside living cells would be useful tools in molecular biology, and perhaps human medicine. Minor groove-binding polyamides containing the aromatic amino acids *N*-methylpyrrole (Py), *N*-methylimidazole (Im), *N*-methyl-3-hydroxypyrrole (Hp), and other related aromatic heterocycles bind DNA with affinities and specificities comparable to naturally occurring DNA-binding proteins.[1-4] DNA recognition by polyamides is described by a code of side-by-side amino acid pairings that are oriented N→C with respect to the 5'→3'-direction of the DNA helix in the minor groove: Im/Py is specific for G•C, Hp/Py is specific for T•A, and Py/Py binds both A•T and T•A.[1]

Polyamides have been shown to influence a wide variety of protein-DNA interactions in solution,[5-15] yet similar experiments in cell culture have proven to be dependent on cell-type.[16-18] Studies with fluorescent Bodipy-labeled polyamides indicate that these conjugates are excluded from the nuclei of most cells, with the notable exceptions of lymphoid and myeloid cell types.[19,20] One report suggested that the fluorophore itself may play a role in cellular uptake: an eight-ring polyamide-fluorescein conjugate was shown to accumulate in the nuclei of HCT-116 colon cancer cells.[21] Expanding upon this lead, we recently described a set of polyamide-fluorescein conjugates that localized to the nuclei of thirteen live mammalian cell lines.[22] Relatively small structural alterations, such as differences in Py/Im sequence, caused dramatic changes in nuclear localization. The current study seeks to elucidate more fully the structural requirements for nuclear localization of polyamide-fluorophore conjugates. This is a minimum first step in exploring the use of these DNA-binding ligands for gene

modulation in cell culture experiments.

Results

We have synthesized 100 polyamide-fluorophore conjugates, whose chemical structures are shown along with schematic ball-and-stick representations (e.g., Figure 4.6). We examined the nuclear localization properties of these compounds in thirteen mammalian cell lines, representing eleven human cancers, one human transformed kidney cell line, and one murine leukemia cell line. Cells were allowed to incubate with polyamide-fluorophore conjugate in the surrounding medium (2 or 5 μM, as indicated, 10-14 h) and the extent of nuclear uptake was analyzed by confocal microscopy. Polyamide-dye conjugates display a range of uptake efficiencies, varying from strong nuclear concentration, producing brightly stained nuclei with little or no signal in the medium or cytoplasm, to an absence of nuclear accumulation, in which case the molecules may be trapped in vesicles or excluded from cells entirely. Within a single sample, staining is usually remarkably homogenous, displaying minimal cell-to-cell variation.

The extent of nuclear localization for each sample was assigned one of four qualitative ratings, and these data are organized into tables showing the ball-and-stick structure of each compound and the level of uptake in each cell line (e.g. Figure 4.7). A thin vertical line separates data collected in adherent cell lines (*left*, MCF-7 through 293) from data collected in suspended cell lines (*right*, Jurkat through NB4). Within these two sets, more permissive cell lines – those that generally display stronger nuclear staining with polyamide conjugates – are towards the left. Thus, MCF-7 and Jurkat are the most

permissive cell types of the adherent and suspended cells, respectively. Shaded groupings of compounds indicate structurally related conjugates, as discussed in the text.

Dye composition

We synthesized conjugates **1**-**20** to investigate the effect of the fluorophore moiety on nuclear localization (Figure 4.6). Figure 4.7 shows the uptake profiles of these compounds. Polyamides **1**-**6** are conjugated to fluorescein derivatives, and all show strong nuclear staining in the cell lines tested. There appears to be flexibility in the specific type of chemical linkage, as thiourea (**1**, resulting from FITC conjugation) and amide (**2**-**6**) linkers are effective, including the extended thioether amide-linkage of Oregon Green® (OG) 514 in compound **5**. Fluorination of the aromatic rings is well tolerated (**3**-**5**), and replacing the carboxylic acid with a sulfonic acid (as in OG 500 conjugate **4**) appears to be somewhat beneficial in several cells lines (compare compounds **3** and **4**). The greater photostability of the fluorinated OG derivatives, as compared to fluorescein, may be advantageous for fluorescence imaging applications.[28,29] The JOE fluorophore, a fluorescein derivative with chloro and methoxy substituents, has red-shifted absorption and emission spectra compared to fluorescein, such that their spectra can be distinguished. JOE conjugate **6** shows moderate to strong uptake in all cell lines, providing a second color for multicolor fluorescence applications.

Compounds **7**-**11**, conjugated to rhodamine derivatives, display poorer nuclear localization profiles than analogous fluorescein conjugates. The sulfonated rhodamine-dye Alexa 488 is a highly photostable alternative to fluorescein; however, conjugate **7** was completely excluded from the nuclei of all cells tested. Rhodamine green derivative

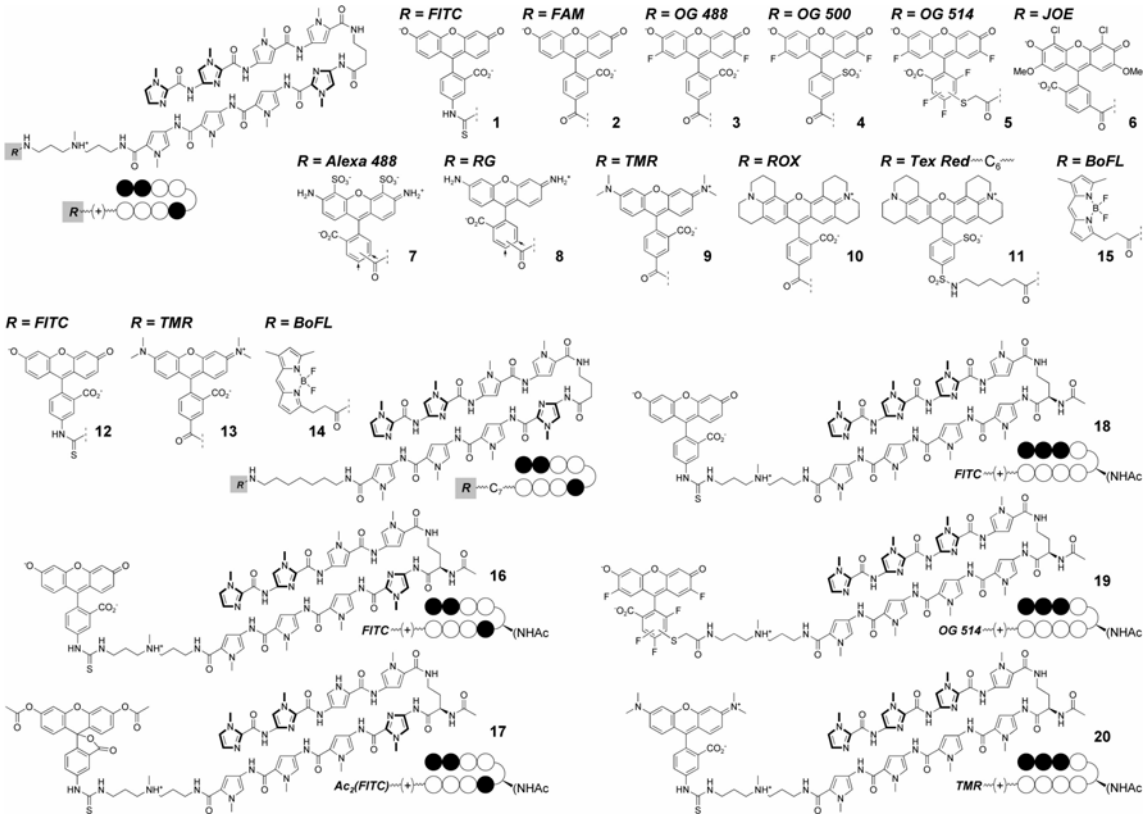


Figure 4.6 Dye composition. Chemical and ball-and-stick structures of compounds **1**-**20**, testing the effect of differing dye composition on cellular uptake.

8, lacking sulfonates, showed slightly stronger nuclear localization, particularly in adherent cell lines. The additional alkyl groups in tetramethyl rhodamine (TMR) conjugate **9** impart enhanced uptake properties in most cell lines, compared to **8**. Depending upon the cell type (as well as the ring sequence and composition of the polyamide, *vide infra*), polyamide-TMR conjugates may be useful because of the photostable, red/orange fluorescence of the TMR dye. Compound **10** is conjugated to the related ROX fluorophore, and conjugate **11** incorporates a Texas Red fluorophore with a hexanoic acid spacer, a dye/linker combination that has been used to label polyamides for telomere staining experiments.[30] Both polyamides, **10** and **11**, are consistently excluded

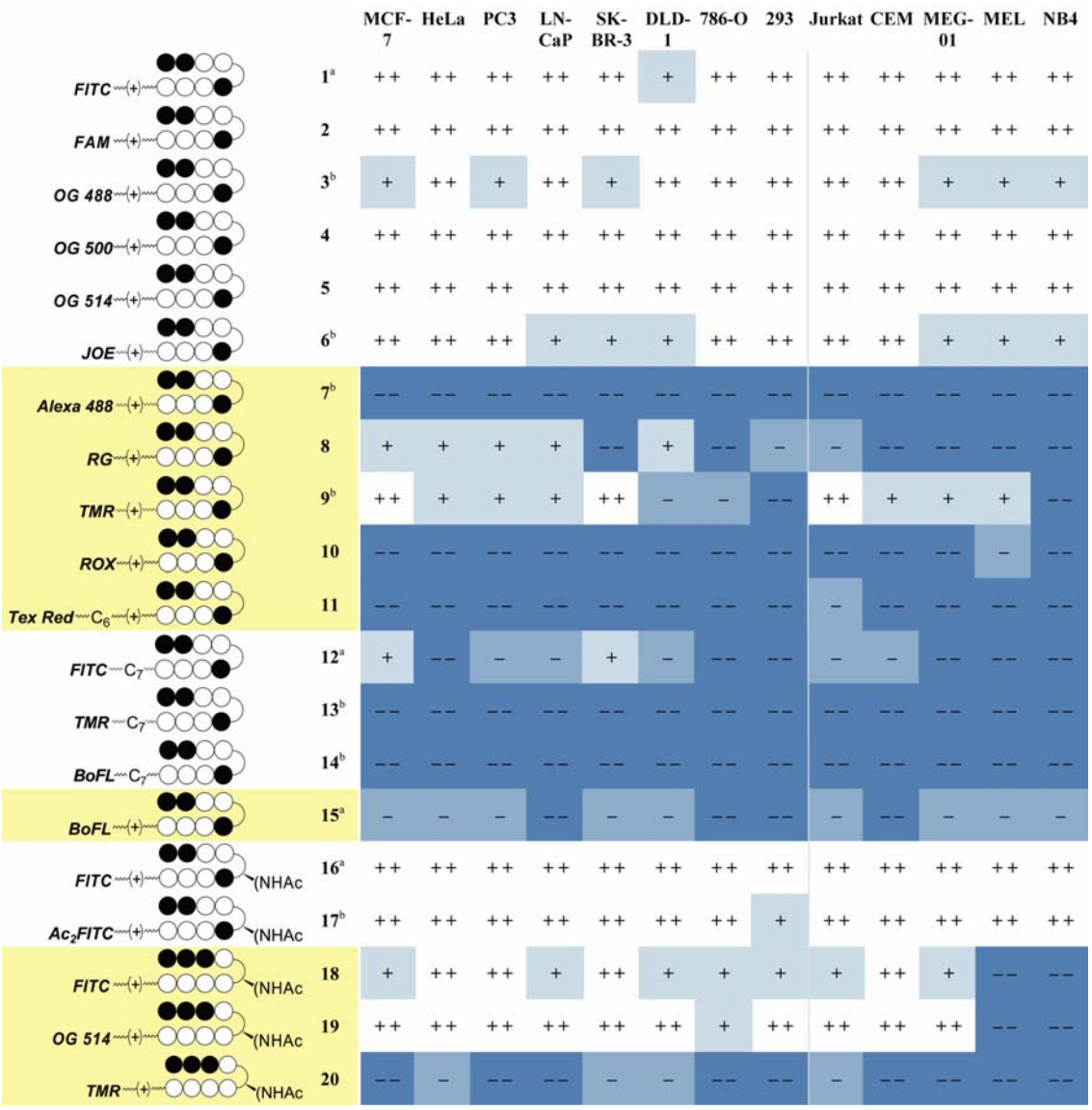


Figure 4.7 Dye composition: uptake profile of polyamides **1**-**20** in thirteen cell lines. Alternating highlighting indicates groups of chemically similar polyamides, as described in the text. Data for adherent and suspended cells are towards the left and right, respectively, separated by a light gray vertical line. Lighter shades of blue represent stronger nuclear localization. + +, 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. Except where noted, polyamide concentration was 2 μM. [a]Polyamide described previously,[22] assayed at 5 μM. [b]Assayed at 5 μM.

from live cells.

In compounds **12-14**, a seven-carbon alkyl chain links the dye—FITC, TMR, and Bodipy FL (BoFL), respectively—to the polyamide. For comparison, BoFL derivative **15**, incorporating the standard cationic triamine linker, is also included. As described previously, the BoFL conjugate in this motif displays only minimal evidence of nuclear access in live cells.[22] For each alkyl-linked conjugate, the loss of the tertiary amine in the linker worsened the uptake profile (compare **1** and **12**, **9** and **13**, and **15** and **14**). It was shown previously that alkyl-linked conjugate **12** performed much more poorly than triamine-linked conjugate **1**, and that addition of an amino group at the turn residue, using the chiral $^{H_2N}\gamma$-turn, restored most of the uptake properties (see compound **78**, below).[22] In contrast, adding amino groups to the dye, as in alkyl-linked TMR conjugate **13**, abolishes nuclear access.

Acetylating fluorescein produces a non-fluorescent, uncharged moiety; cleavage of the acetates by esterases subsequently unmasks the fluorophore. Polyamide **16**, as described previously,[22] demonstrates that the $^{AcHN}\gamma$-turn does not impede uptake, and the analogous diacetyl-FITC conjugate **17** stains nuclei with similar effectiveness. It should be noted that no special effort was taken to remove esterases from the growth medium, such that some amount of **16** is likely to be in solution when cells are treated with **17**. Polyamides **18-20** have the same number of Im and Py residues as compounds **1-17** but in a different sequence. With the exception of the less permissive suspended cells MEL and NB4, FITC conjugate **18** (an isomer of **16**) and OG 514 conjugate **19** display modest to strong nuclear staining. In contrast, TMR conjugate **20** is a very poor nuclear stain. Comparing analogous compounds, conjugates **18** and **19** display only slightly poorer

uptake profiles than those of conjugates **1** and **5**, respectively, whereas TMR conjugate **20** is a much worse nuclear stain than polyamide **9**. Such comparisons indicate that modifications such as ring sequence and fluorophore structure interact to affect nuclear uptake in ways that are not yet predictable.

Second-Generation Rings

We have recently described several second-generation aromatic ring systems that improve the DNA-binding properties of polyamides.[3,4] When paired against Py, a chlorothiophene (Ct) residue at the N-terminal cap position targets a T•A base pair.[4] The hydroxybenzimidazole (Hz) moiety, incorporated as part of a dimeric subunit, appears to be a chemically stable replacement for the Hp residue,[3] and a CtHz-cap dimer has recently been found to target the sequence 5'-TT-3' (R.M. Doss, M.A. Marques, S. Foister, and P.B. Dervan, unpublished results). Figure 4.8 shows the chemical structures of polyamide-fluorophore conjugates incorporating second-generation rings, and Figure 4.9 displays their uptake properties.

Ct-cap compounds **21**-**24** have the same ring sequence but different dyes. Polyamides **21**-**23**, conjugated to fluorescein derivatives, display strong nuclear localization, whereas TMR conjugate **24** displays a moderate uptake profile that is generally similar to **9** and better than **20**. Conjugates **25**-**28** demonstrate the effects of different ring sequences on nuclear uptake of thiophene-cap conjugates. Whereas **25** displays good to excellent nuclear staining, its isomer **26** is excluded from all cells tested. Other conjugates with three contiguous Im residues – methylthiophene-cap polyamide **27**, and ten-ring hairpin **28** – also display very poor uptake properties. The only

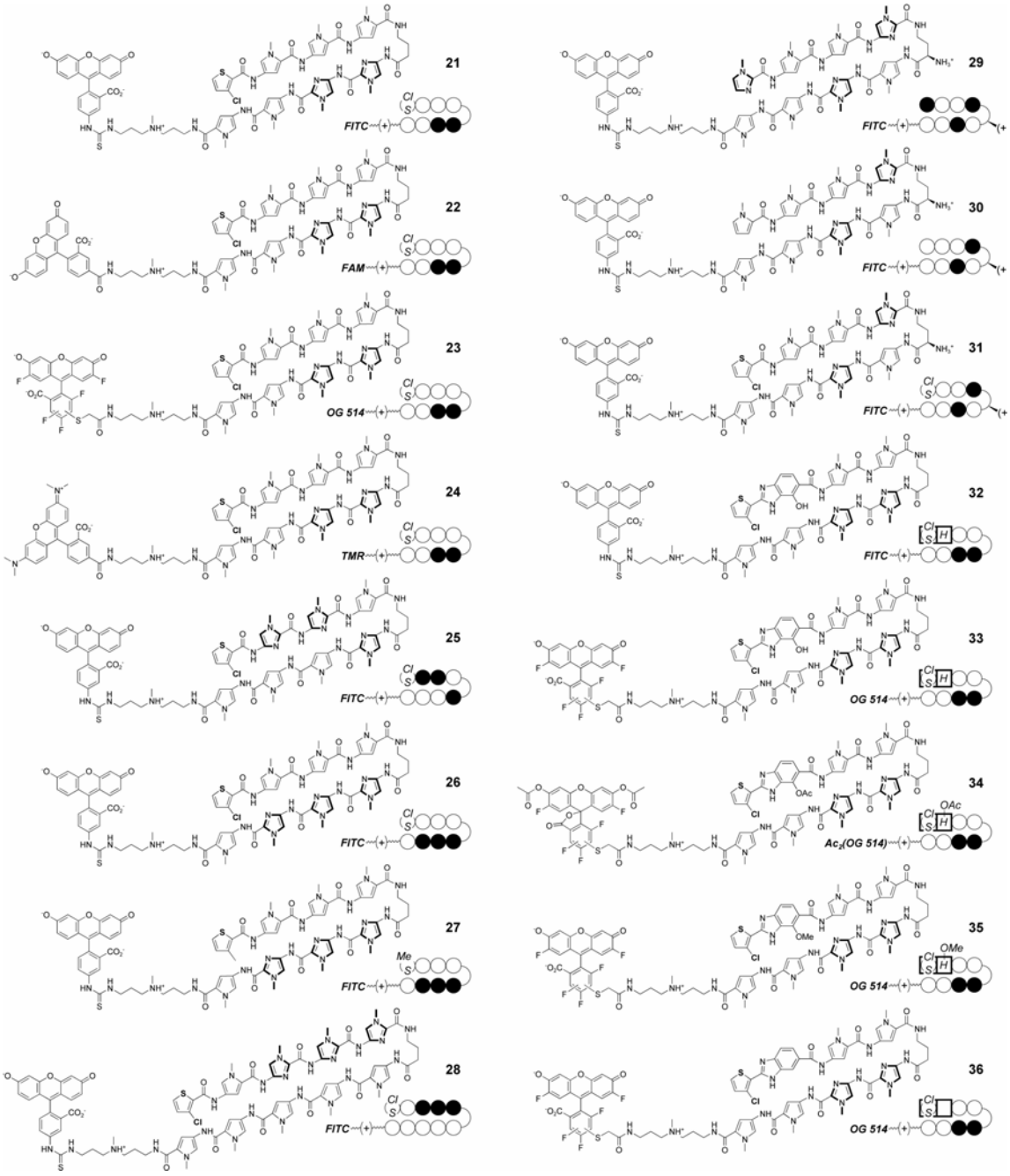


Figure 4.8 Second-generation rings. Chemical and ball-and-stick structures of compounds **21**-**36**, testing the effect of new aromatic rings on cellular uptake.

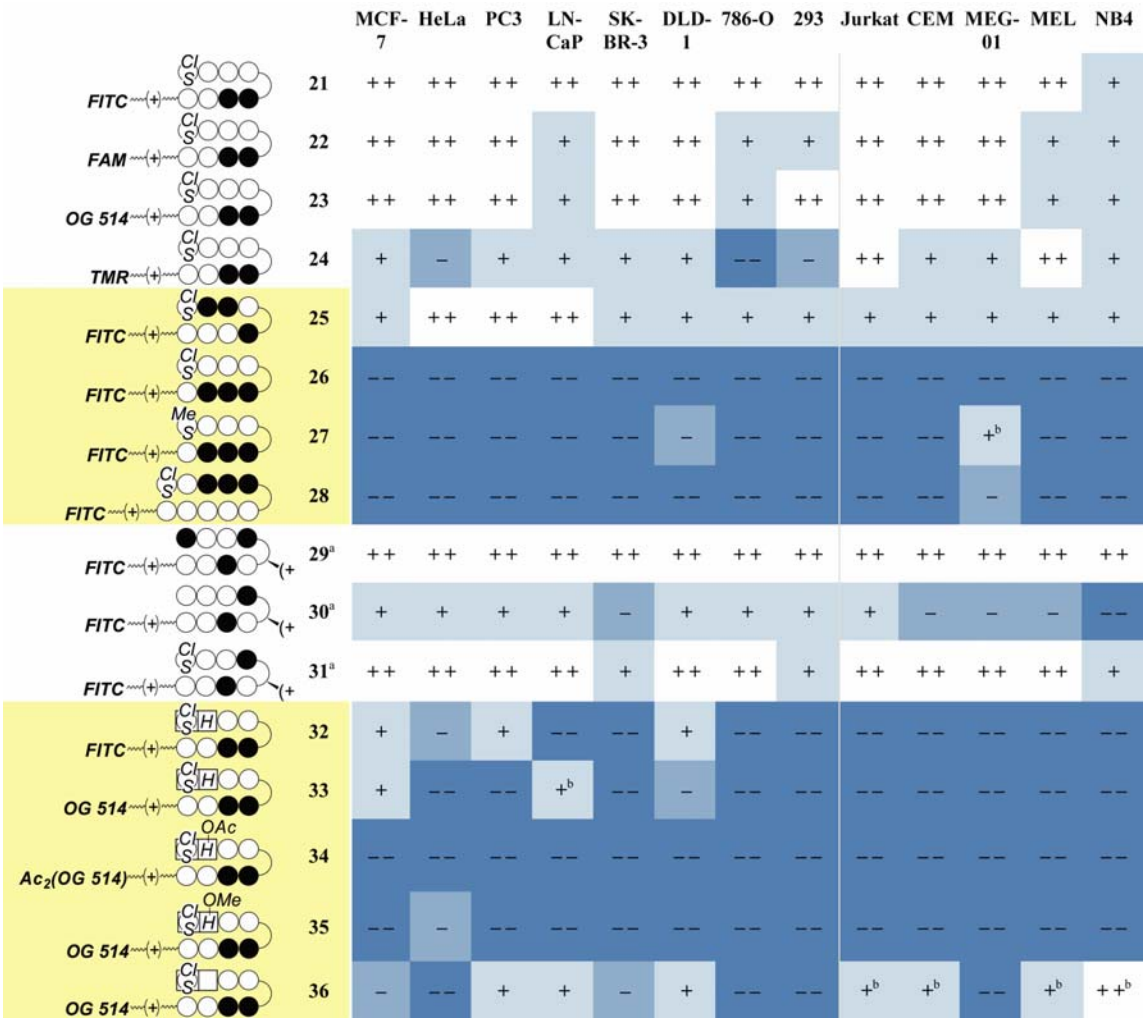


Figure 4.9 Second-generation rings: uptake profile of polyamides **21-36**. Symbols are defined in Figure 4.7. Except where noted, polyamide concentration was 2 μM. [a]Assayed at 5 μM. [b]Highly heterogeneous uptake profile. Some cells display clear nuclear localization, while others show none. The value shown is an approximate average.

exception is relatively modest and heterogeneous uptake of **27** in MEG-01 cells. Contrasting these data with the generally good to excellent nuclear staining displayed by **18** and **19**, which also contain three contiguous imidazole residues, reemphasizes the complexity of the interactions between polyamide-fluorophore conjugates and cells.

Polyamides **29-31**, each with a chiral $^{H_2N}\gamma$-turn, differ only in cap-residue

structure. Compounds **29** and **31**, with Im and Ct caps, respectively, are excellent nuclear stains, whereas Py-cap polyamide **30** is significantly less effective. It is remarkable that an Im→Py conversion (**29** to **30**), formally an N→C-H shift, reduces nuclear uptake, but the more dramatic Im→Ct conversion affects uptake very little.

Based on the ring sequence of **21**-**24**, compounds **32-36** incorporate variants of the benzimidazole ring system. Although polyamides incorporating the Hz unit display favorable DNA-binding properties, FITC conjugate **32** stains the nuclei of only a few cell lines, and OG 514 conjugate **33** stains only MCF-7 and LNCaP cells. Modifications to the hydroxyl group, including acetylation (**34**) and conversion to a methoxy group (**35**), did not improve the uptake profile. Complete removal of the hydroxyl group (**36**) allows nuclear localization in many cell lines, though the uptake profile does not correlate with the general permissiveness of the cell types. Furthermore, removing the hydroxyl eliminates the ability to distinguish between A•T and T•A base pairs at that position.[3]

Extended Hairpin Motif

Figure 4.10 shows the chemical structures of polyamide-dye conjugates **37-46** with C-terminal unpaired rings, and Figure 4.11 shows their uptake profiles. The molecules are grouped into pairs that differ with respect to the conjugated dye. Although the DNA-binding properties of the extended hairpin motif have not been studied as extensively as those of fully ring-paired hairpin polyamides,[31,32] extended hairpin polyamides have been shown to interfere with NF-κB—DNA interactions[13] and to label specific heterochromatic regions of human chromosomes.[29] Furthermore, tandem polyamides with extended hairpin subunits have been used to stain telomeres[30] and to

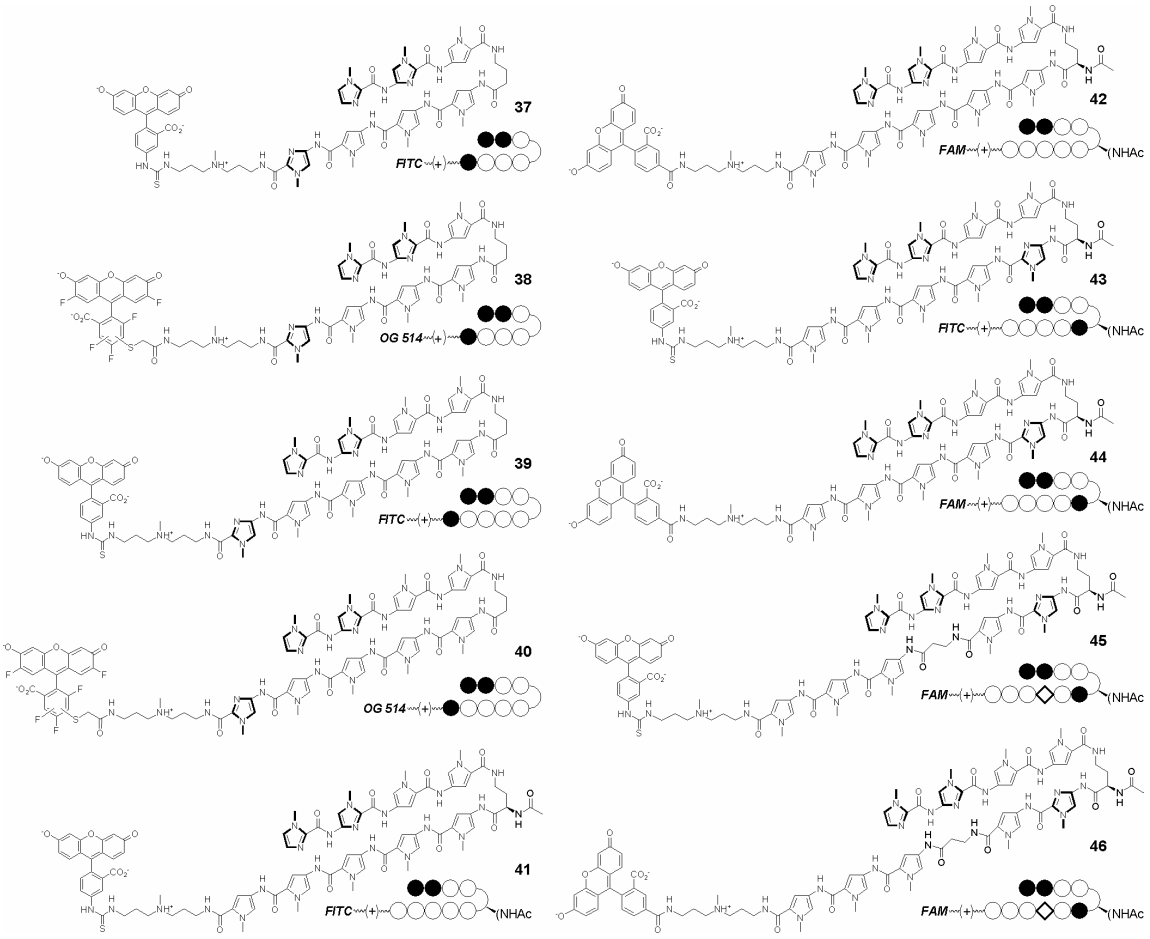


Figure 4.10 Extended hairpins. Chemical and ball-and-stick structures of compounds **37**-**46**, testing the effect of extensions to the hairpin motif on cellular uptake.

displace a viral transcription factor from DNA.[33] Overall, conjugates **37**-**46** display good to excellent nuclear localization properties.

Seven-ring polyamides **37** and **38**, conjugated to FITC and OG 514, respectively, are moderately effective nuclear stains. The larger nine-ring FITC conjugate **39** displays a significantly improved uptake profile, whereas OG 514 conjugate **40** is similar to the seven-ring hairpins. Polyamides **41** and **42** have a similar nine-ring core, but with a single unpaired Py residue and a chiral $^{AcHN}\gamma$-turn. Both the FITC and FAM conjugate (**41** and **42**, respectively) are excellent nuclear stains. From these compounds, one

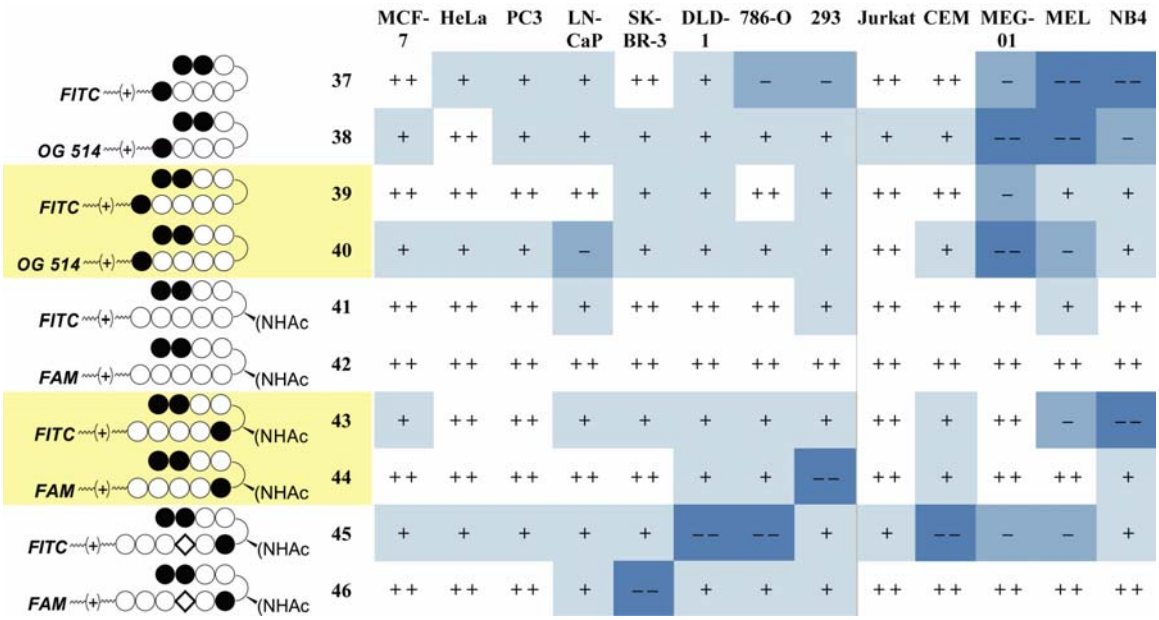


Figure 4.11 Extended hairpins: uptake profile of polyamides **37**-**46**. Symbols are defined in Figure 4.7.

Py→Im substitution gives **43** and **44**, respectively. In this case, FAM conjugate **44** performs better than FITC conjugate **43**. Based on a motif that was used in studies with NF-κB,[13] compounds **45** and **46** have two unpaired Py residues. Again, the FAM conjugate displays the better uptake profile, strongly staining the nuclei of eight of the thirteen cell lines.

Larger Polyamides

Somewhat larger than the extended hairpin polyamides **37**-**46** are fully ring-paired hairpins **47**-**54** and **56**-**62**. Figure 4.12 shows the structures of dye conjugates **47**–**62**, and Figure 4.13 displays their uptake profiles. Many of these compounds incorporate the flexible β-Ala residue, which is often substituted for Py to allow polyamides to adapt to sequence-dependent DNA microstructure and flexibility.[34,35] In general, conjugates of

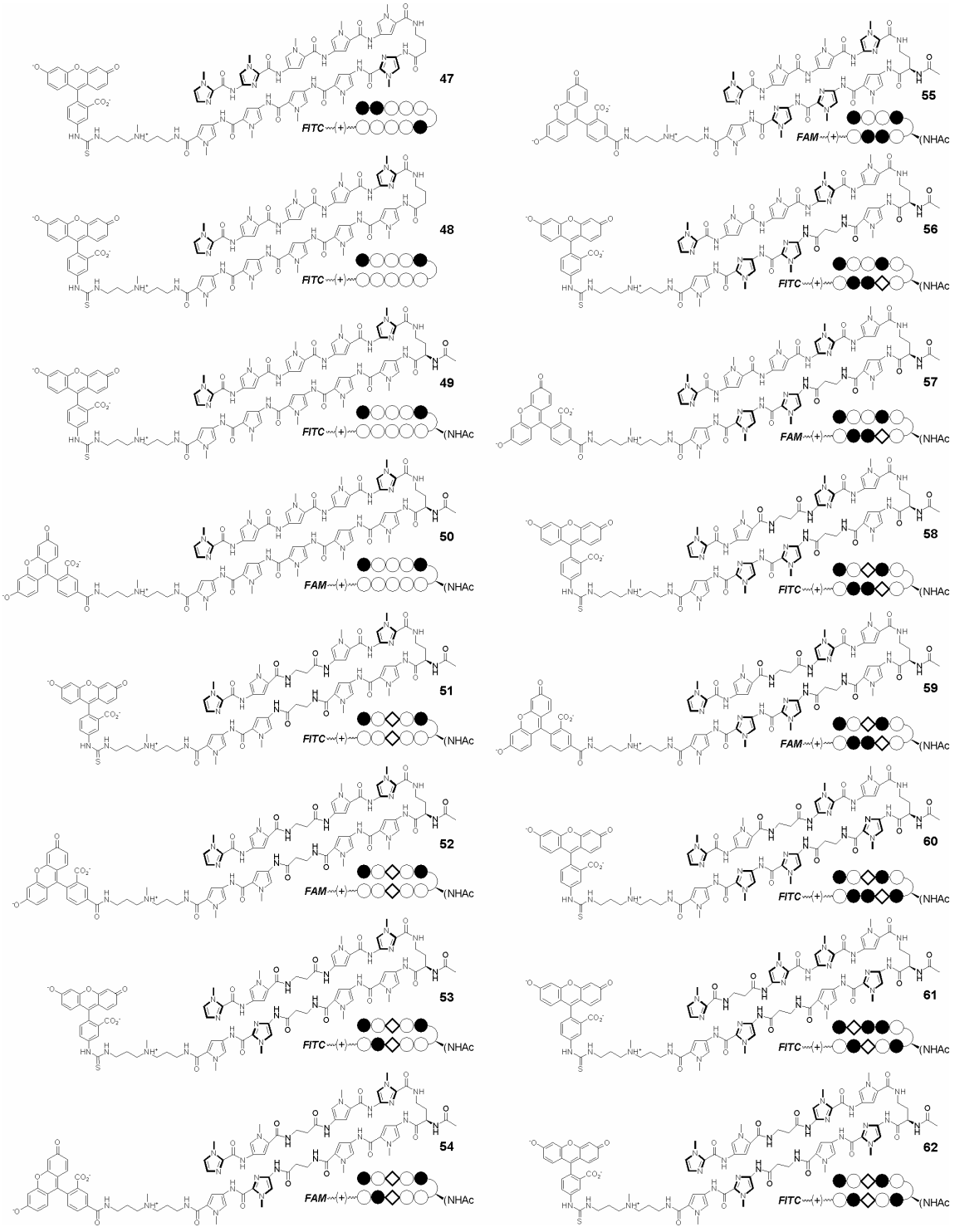


Figure 4.12 Larger polyamides. Chemical and ball-and-stick structures of compounds **47**-**62**, testing the effect of molecular size on cellular uptake.

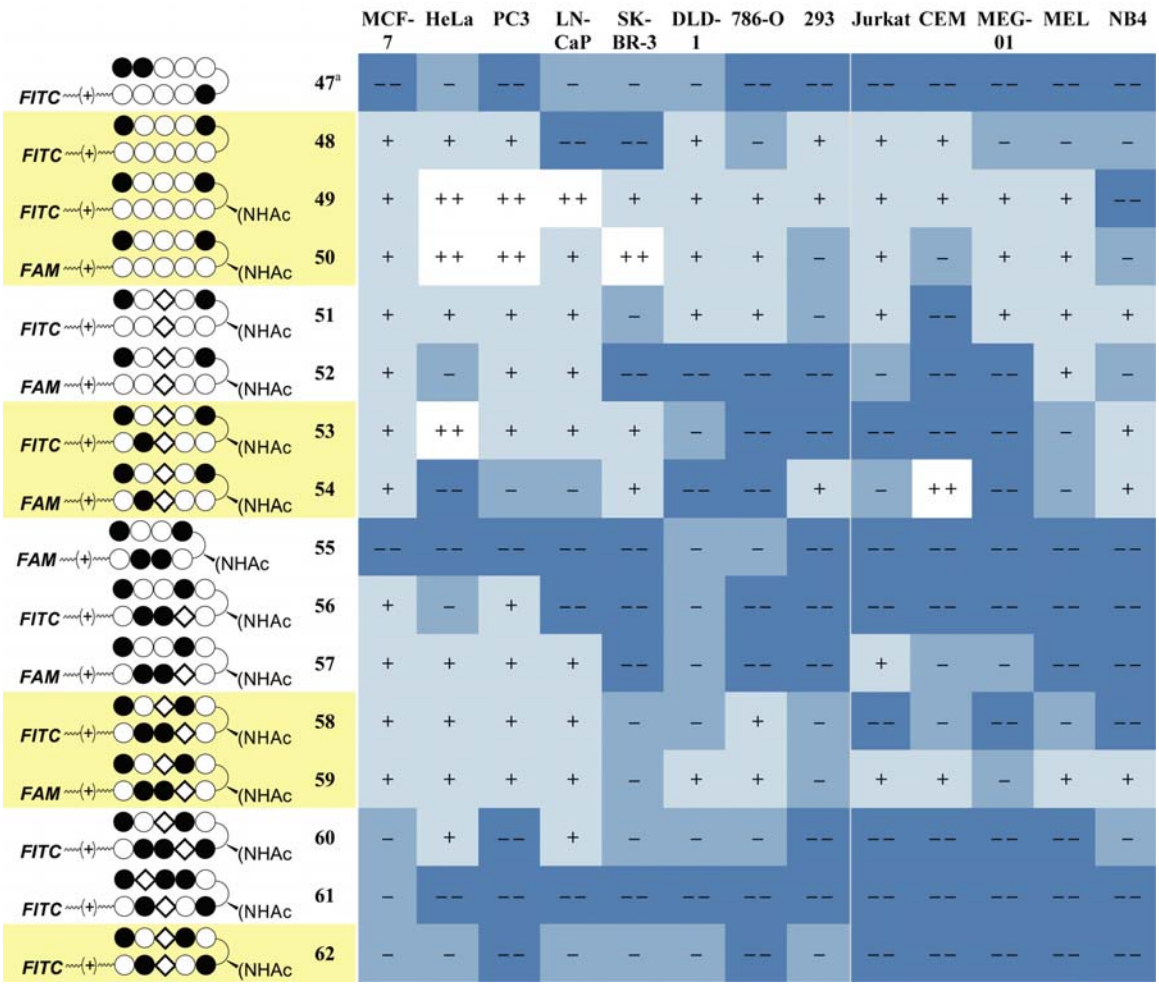


Figure 4.13 Larger polyamides: uptake profile of polyamides **47**-**62**. Symbols are defined in Figure 4.7. Except where noted, polyamide concentration was 2 μM. [a]Polyamide described previously,[22] assayed at 5 μM.

this motif display poor to moderate nuclear staining, although there are scattered exceptional cases.

Conjugate **47**, described previously, displayed a particularly poor uptake profile. From **47**, an Im→Py substitution produces **48**, which displays some nuclear staining in eight cell lines. Chiral $^{AcHN}\gamma$-turn analogs of **48**, conjugates **49** and **50**, display improved uptake properties, accessing the nuclei of most of the cell types, and showing excellent staining in some of the more permissive adherent cells. Polyamides **51** and **52** (related to

49 and **50**, respectively, by Py→β-Ala substitutions of the central residues), are less efficient nuclear stains, with FAM conjugate **52** performing significantly more poorly. Compounds **53** and **54** result from an additional Py→Im substitution. FITC conjugate **53** accesses the nuclei of the more permissive adherent cells, whereas FAM conjugate **54** displays a scattered uptake profile.

Polyamides **55**-**59** are designed to target essentially the same 6-bp DNA sequence. Eight-ring hairpin **55**, included as a reference, is excluded from all cell nuclei, whereas the larger polyamides access nuclei somewhat more successfully. In this set, internal β-Ala substitution and FAM-conjugation are both positive determinants for nuclear uptake; accordingly, conjugate **59** is the most effective nuclear stain. Polyamide **60** is related to **58** by a Py→Im substitution, and **60** displays a much poorer uptake profile. Inverting the strands of **60** produces **61** – the result of a formal 180° rotation of the ring core – which is excluded from the nuclei of all cells tested. Conjugate **62** is based on a motif that has been employed in a variety of studies;[18,20,36] however this FITC conjugate is excluded from cell nuclei.

C-terminal β-Ala Residues

In a previous study, we noted that adding β-Ala residue to the C-terminal ("tail") position of a polyamide-dye conjugate could have dramatic effects on nuclear localization.[22] We therefore synthesized the molecules shown in Figure 4.14 to investigate the effects of β-Ala-tail residues. To allow for the possibility that the added β-Ala would enhance uptake, precursor ring systems were chosen to have a range of uptake efficiencies (Figure 4.15). The polyamides are grouped by their core ring

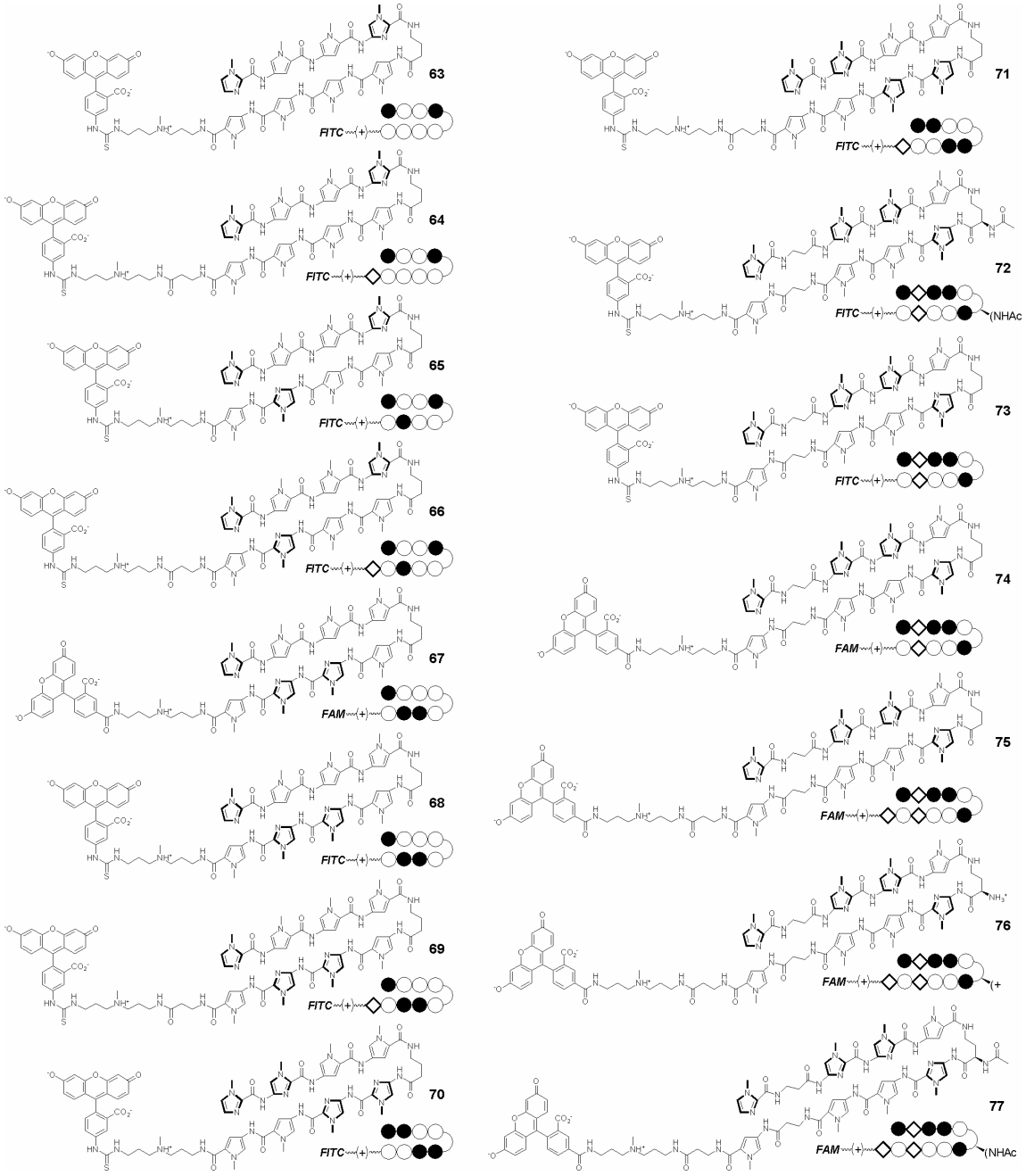


Figure 4.14 β-Alanine tails. Chemical and ball-and-stick structures of compounds **63**-**77**, testing the effect of molecular size on cellular uptake.

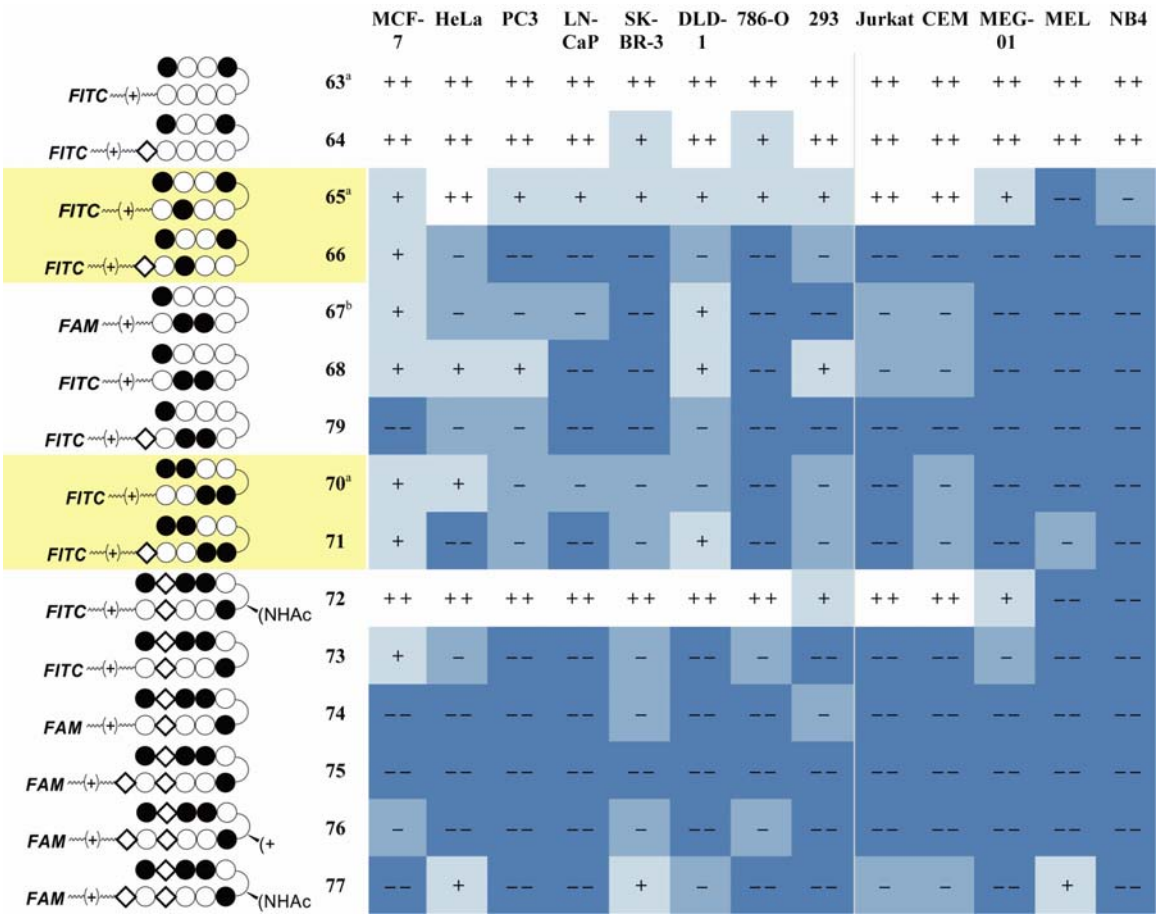


Figure 4.15 β-Alanine tail residues: uptake profile of polyamides **63**-**77**. Symbols are defined in Figure 4.7. Except where noted, polyamide concentration was 2 μM. [a]Polyamide described previously,[22] assayed at 5 μM. [b]Assayed at 5 μM.

sequence. In almost every case, adding a β-Ala-tail residue reduces nuclear uptake efficiency. Indeed, even in cases in which both samples have the same qualitative rating, the conjugate with a β-Ala-tail shows slightly less intense nuclear staining.

Conjugate **63** is an excellent nuclear stain, and addition of a β-Ala-tail residue (compound **64**) affects uptake only minimally. In contrast, polyamide **65** displays a moderate to good uptake profile, and its β-Ala-tail analog **66** shows very poor nuclear staining. Although FITC conjugate **68** (an isomer of polyamide **1**) accesses nuclei slightly more effectively than FAM conjugate **67**, it is nonetheless a rather weak stain,

and addition of a β-Ala-tail residue (conjugate **69**) completely inhibits nuclear access. Conjugate **70**, described previously,[22] accesses the nuclei of only two cell lines, and addition of a β-Ala-tail residue (compound **71**) does not change the uptake profile significantly. Polyamides **72**-**77** have the same ring sequence, but vary with respect to dye structure, turn substitution, and the presence of a β-Ala-tail residue. Of this series, only FITC conjugate **72** (with an $^{AcHN}\gamma$-turn and lacking a β-Ala-tail) displays an excellent uptake profile. Polyamide **72** is isomeric with compounds **58** and **62**, and of these three conjugates, **72** displays by far the most efficient nuclear staining. Interestingly, although the overall uptake profile of FAM conjugate **77** is relatively poor, this β-Ala-tail polyamide is one of three compounds in Figure 4.15 that can access the nuclei of MEL cells.

Turn-linked conjugates

To investigate other dye-attachment points, conjugates were synthesized linked to the chiral turn element (**80**-**88**, Figure 4.16). For comparison, this set includes tail-conjugated polyamides **78**, **79**, and **89**. Figure 4.17 presents the uptake profiles of **78**-**89**. Polyamides **78** and **79** are conjugated to FITC through seven- and six-methylene spacers, respectively. Neither compound stains nuclei as effectively as the analogous conjugate **1**, although **79** is quite close. Compared to **78**, polyamide **80** exchanges the positions of the cationic amine group and the dye linker. These turn-linked FITC conjugates are excellent nuclear stains, and the presence of an additional amine (in **81**) or a β-Ala-tail residue (in **82**) does not impede nuclear uptake. The sole exception is compound **81** in NB4, a cell line which was previously observed to exclude polyamide-dye conjugates with added

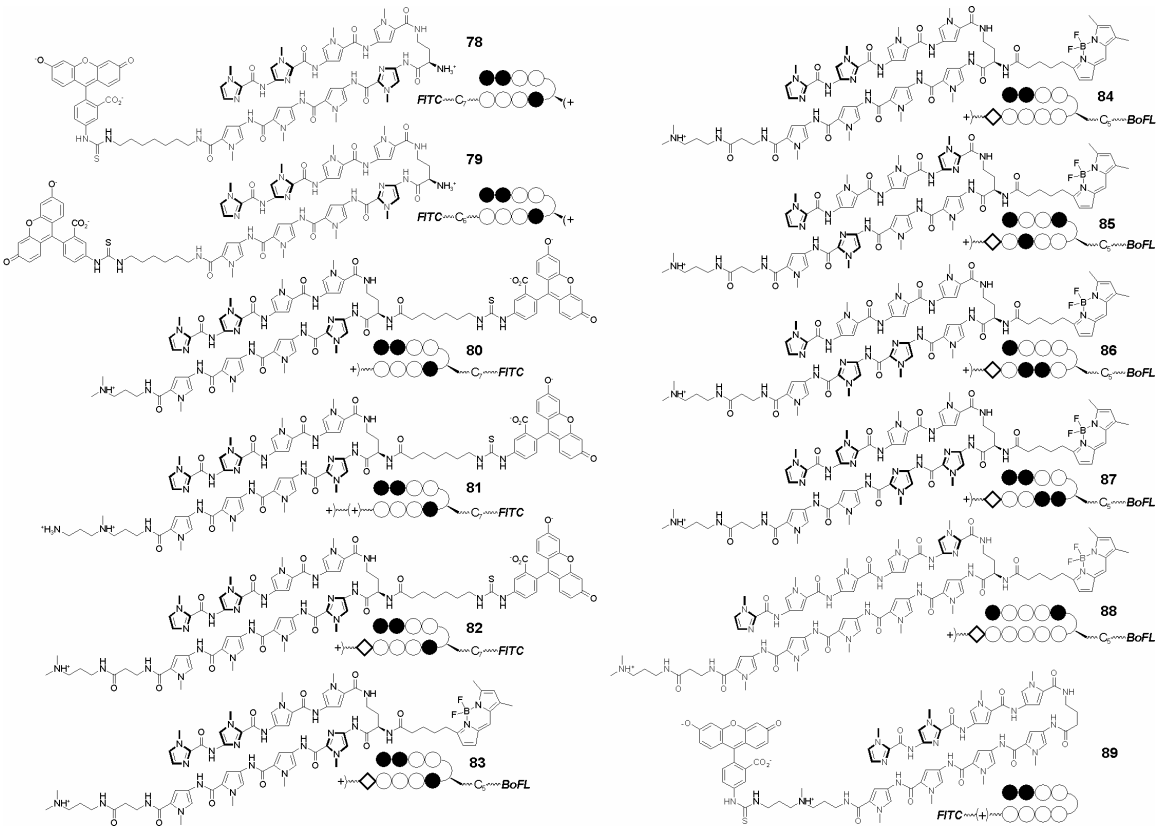


Figure 4.16 $^{H_2N}\gamma$-turn-linked conjugates. Chemical and ball-and-stick structures for polyamides **78-89**, testing the effect of removing the dye moiety to the $^{H_2N}\gamma$-turn.

positive charges.[22]

Hairpin polyamides **83-88**, linked to BoFL through the turn, differ with respect to Py/Im composition, and are analogous to the tail-linked FITC conjugates **1**, **89**, **65**, **68**, **70**, and **48**, respectively. Although each eight-ring BoFL turn-conjugate is a less effective stain than the corresponding FITC tail-conjugate, relative uptake efficiency within each motif appears to be influenced primarily by ring composition. Indeed, ranked by average effectiveness, **83** ≈ **84** > **85** > **86** ≈ **87**, and correspondingly, **1** ≈ **89** > **65** > **68** ≈ **70**. Ten-ring BoFL-turn conjugate performs similarly to its FITC-tail analog **48**, although the $^{AcHN}\gamma$-turn derivative **49** is a significantly better nuclear stain than either

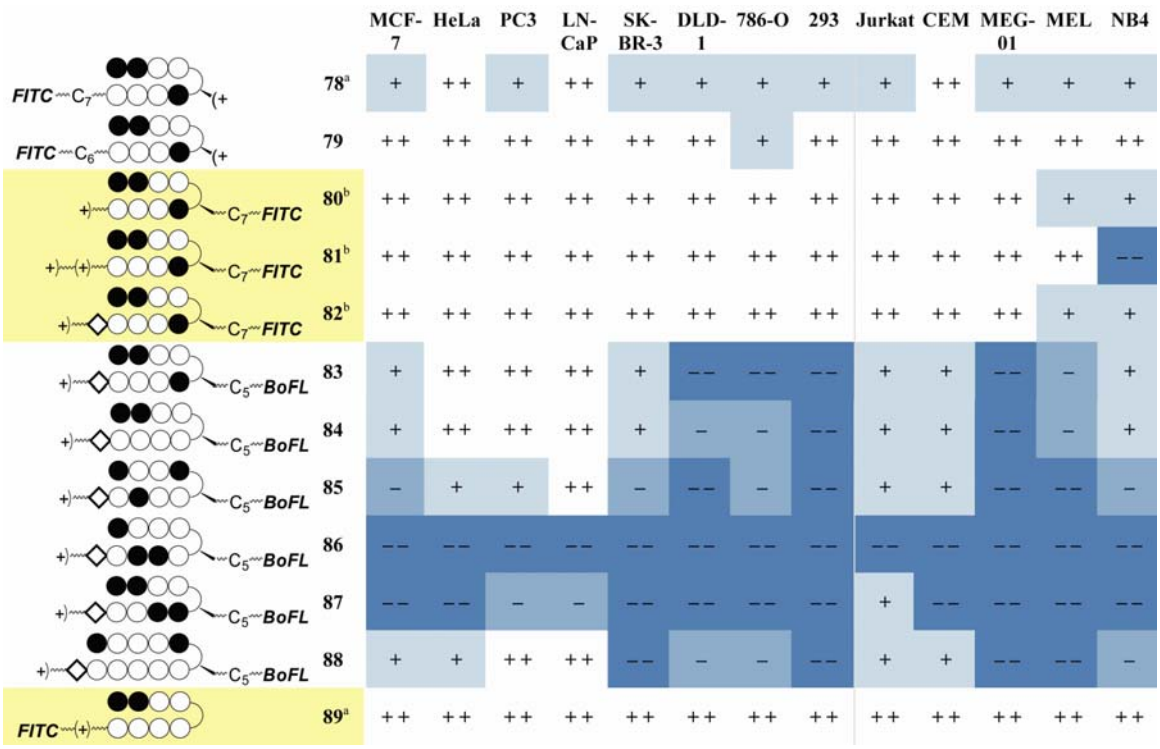


Figure 4.17 $^{H_2N}\gamma$-Turn-linked conjugates: uptake profile of polyamides **78**-**89**. Symbols are defined in Figure 4.7. Except where noted, polyamide concentration was 2 μM. [a]Polyamide described previously,[22] assayed at 5 μM. [b]Assayed at 5 μM.

conjugate. The favorable uptake properties of some of these BoFL-turn conjugates–particularly in more permissive cell lines–is somewhat surprising in comparison to the poor nuclear localization profile of tail-linked BoFL conjugate **15** (Figure 4.7).

Shapes

Beyond hairpin polyamides derivatized at the turn, a variety of unique polyamide shapes have been described, including *N*-methyl substituted hairpins,[13-15,37,38] cycles,[39,40] U-pins,[26] H-pins,[27,41] and tandem hairpins.[30,33,42] Figure 4.18 shows the chemical structures of dye conjugates of various shapes, and Figure 4.19 presents their uptake profiles.

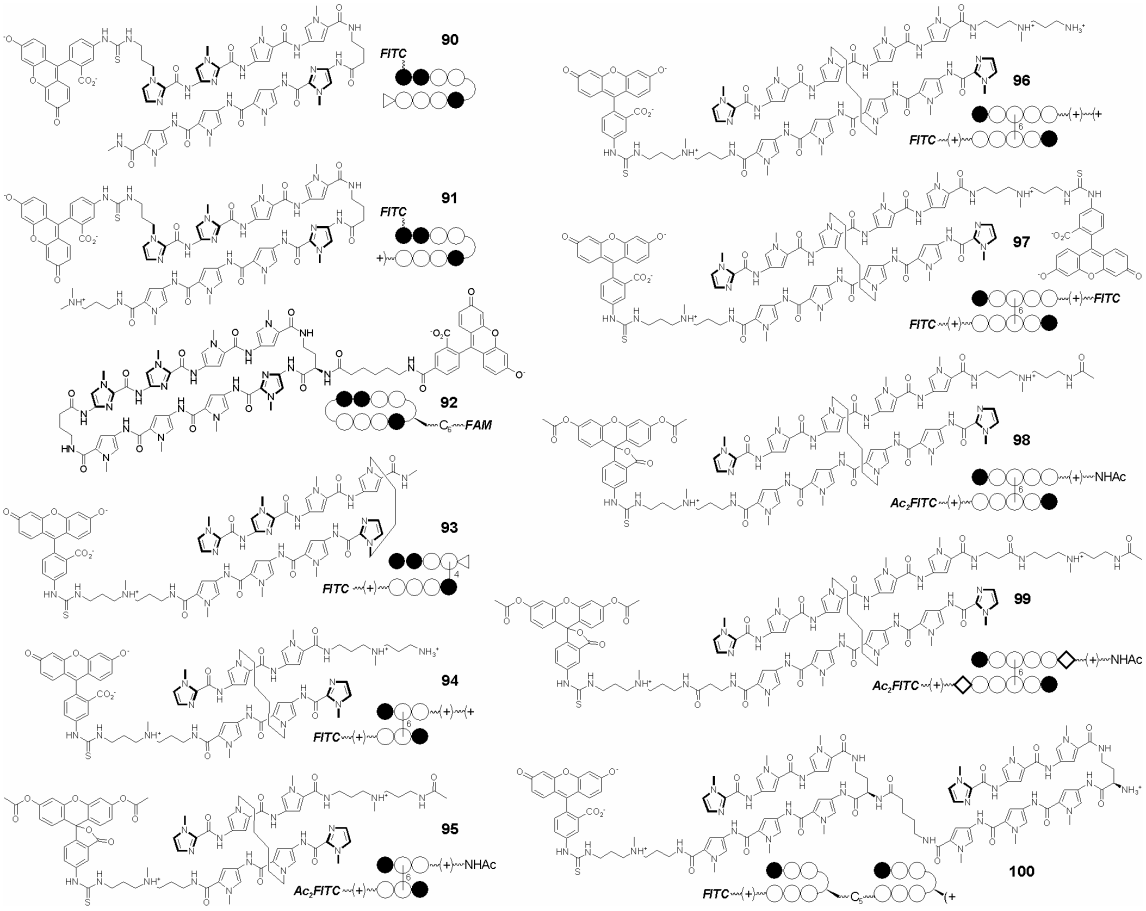


Figure 4.18 Shape variation. Chemical and ball-and-stick structures of compounds **90**-**100**, testing the effect of molecular shape on cellular uptake.

FITC conjugate **90**, linked through an *N*-propylamine Im-cap residue, is a poor nuclear stain, as described previously.[22] After evaluating other compounds, and noting that effective compounds incorporated a cationic amine group, we synthesized compound **91** (an isomer of conjugate **1**), which demonstrated good to excellent uptake properties. Correspondingly, cyclic polyamide **92** displayed a very poor uptake profile. In contrast, U-pin polyamide **93** is a moderately good nuclear stain, especially in adherent cells.

Uptake data for H-pin polyamides indicates that a motif cannot be evaluated by a handful of molecules alone. Of the six-ring H-pins, **95** (the peracetylated derivative of

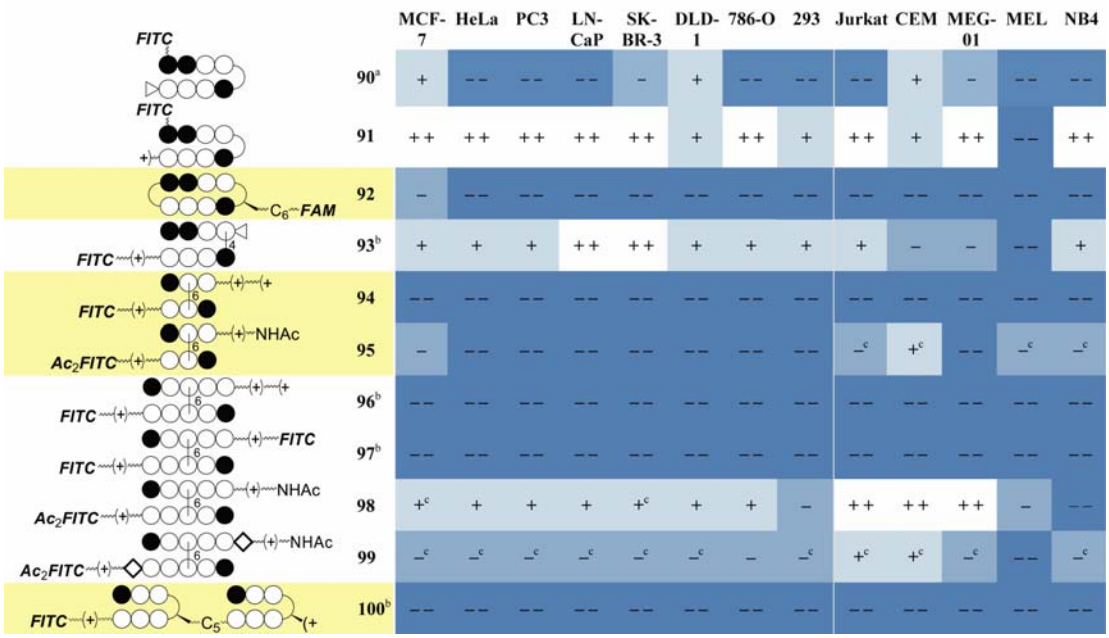


Figure 4.19 Shape variations: uptake profile of polyamides **90-100**. Symbols are defined in Figure 4.7. Except where noted, polyamide concentration was 2 μM. [a]Polyamide described previously,[22] assayed at 5 μM. [b]Assayed at 5 μM. [c]Highly heterogeneous uptake profile. Some cells display clear nuclear localization, while others show none. The value shown is an approximate average.

94) showed highly heterogeneous staining in a small number of cell lines. Within a single microscope frame, some cells were brightly stained, while others were completely dark – where indicated, the data in Figure 4.17 represent a rough average across the cells. Ten-ring H-pins **96** and **97** were excluded from all cells tested, whereas **98** (the ten-ring analog of **95**) showed good to excellent staining in all but the least permissive cell lines. In line with trends observed earlier, adding C-terminal β-Ala residues (compound **99**) diminishes the uptake efficiency.

Tandem hairpins can target longer sequences of DNA than standard hairpin polyamides, so the complete exclusion of FITC conjugate **100** from cells is somewhat disappointing. However, it is possible that another tandem, containing a different Py/Im sequence, dye, or other variation might access the nuclei of some cell types.

Short Peptides

The successful staining of the nuclei of mammalian cells with polyamides labeled with fluorescein and other similar dyes permits many transcription inhibition experiments to be attempted with confidence that a negative result is not due to exclusion of active compound from genomic DNA. However, there are many applications requiring polyamides to be conjugated to peptide and small molecule moieties. The uptake characteristics of these compounds must be good in order to proceed with *in vivo* experiments. As a first-pass attempt, a small set of polyamide-peptide-fluorophore conjugates have been prepared (Figure 4.20) and their uptake characteristics measured. The synthetic scheme for **101**-**105** is shown in Figure 4.21. In general, the uptake of these conjugates is poor, with the compounds being excluded from or sequestered in the cytopoasm of all cell lines studied (Figure 4.22). Further work will be necessary to determine whether this is a general characteristic of polyamide-peptide conjugates, or if some privileged structures might permit good nuclear uptake.

Discussion

From this sea of data, certain trends can be discerned. There does not appear to be any clear correlation between molecular weight and nuclear uptake. Indeed, additional residues can improve uptake (compare, e.g., **37** and **39**, or **55** and **59**), or impede it (compare **1** and **47**, or **48** and **63**). Similarly, a Py→β-Ala exchange, which reduces molecular weight, may intensify or diminish nuclear staining (Figure 4.15). The acetylated ^{AcHN}γ-turn, which adds molecular weight compared to an unsubstituted γ-turn, is one of the few consistently positive factors for nuclear localization (for example,

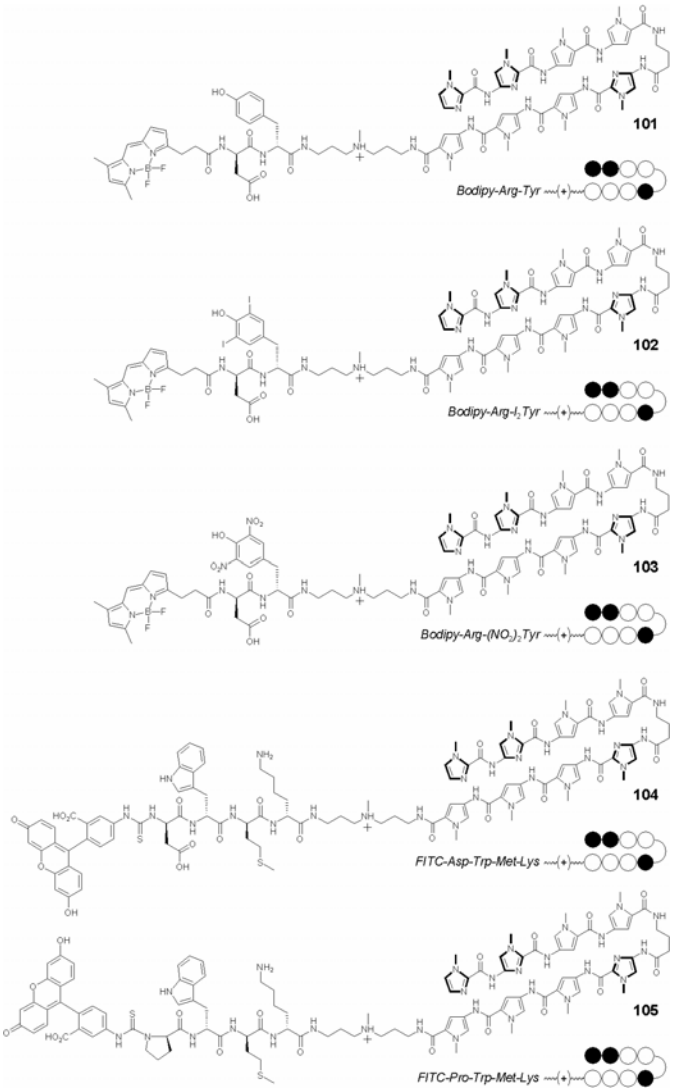


Figure 4.20 Small peptides. Chemical and ball-and-stick structures of compounds **101**-**105**, testing the effect on cellular uptake of the addition of small peptide moieties to polyamide-fluorophore conjugates.

compare **1** and **16**, **48** and **49**, or **72** and **73**). Generally negative determinants for nuclear access include the presence of a β-Ala-tail residue and the lack of a cationic alkyl amine moiety.

Although there is no general correlation between the number of Im residues and uptake efficiency, the positions of Im residues clearly affect nuclear access. For example,

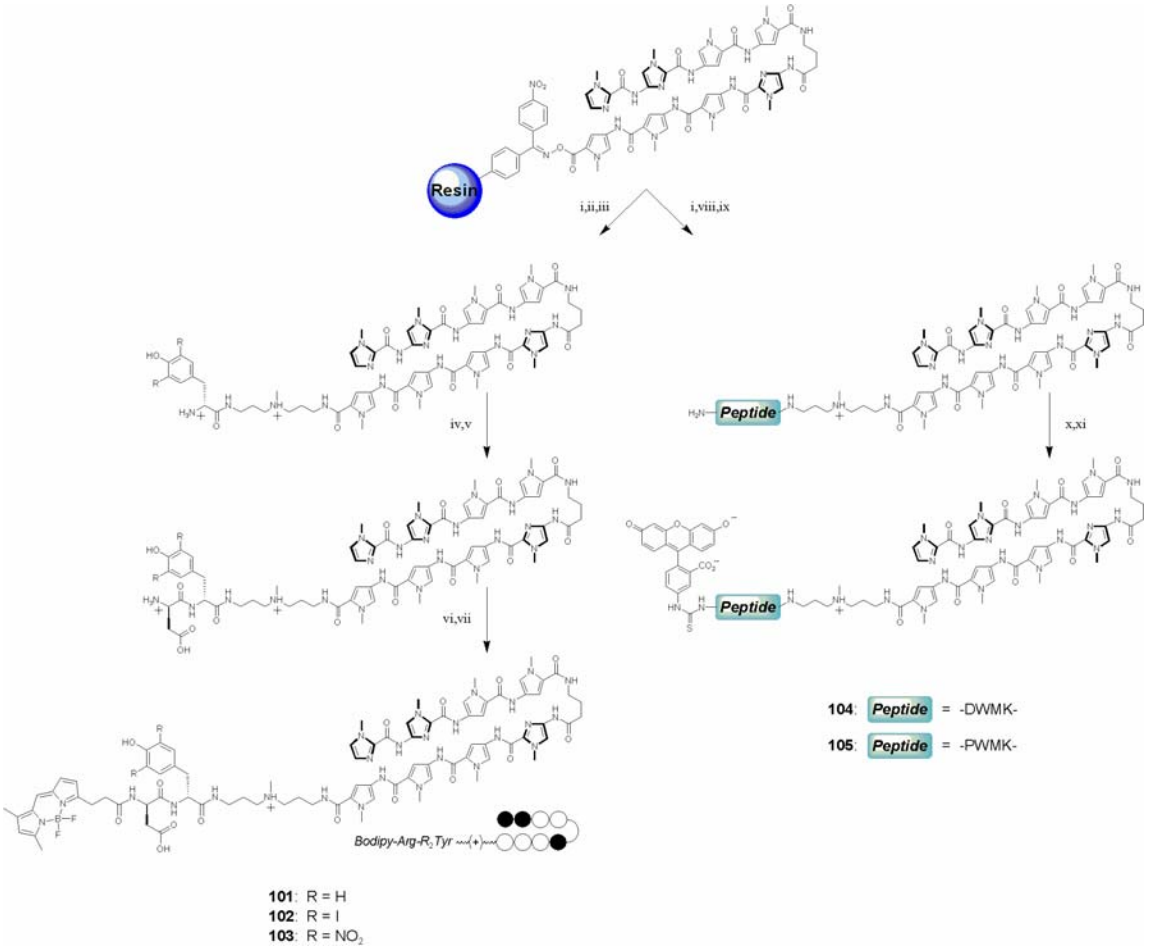


Figure 4.21 Synthesis of small peptide conjugates. (i) 3,3'-diamino-*N*-methyl-dipropylamine, 6 hrs at 37ºC. (ii) R_2Tyr, DCC, HOBt, DIEA, DMF. (iii) 20% (v/v) piperidine/DMF. (iv) (*D*) Asp-Ph_2*i*pr, DCC, HOBt, DIEA, DMF. (v) 20% (v/v) piperidine/DMF. (vi) Bodipy-FL, DCC, HOBt, DIEA, DMF. (vii) 20% (v/v) TFA/CH_2Cl_2. (viii) Peptide, DCC, HOBt, DIEA, DMF. (ix) 20% (v/v) piperidine/DMF. (x) FITC, DIEA, DMF. (xi) 20% (v/v) TFA/CH_2Cl_2.

structural isomers often display very different uptake profiles (compare eight-ring polyamides **1**, **65**, and **68**, or larger conjugates **58**, **62**, and **72**). Furthermore, a Py→Im exchange is nearly always a negative determinant for nuclear access (except at the cap position, compare **29** and **30**), and the impact is highly dependent on the position of the residue and the overall polyamide motif. For example, in eight-ring polyamide-dye conjugates the effect of a single Py→Im exchange may be very small (**89** versus **1**),

modest (**63** versus **65**), or severe (**1** versus **70**).

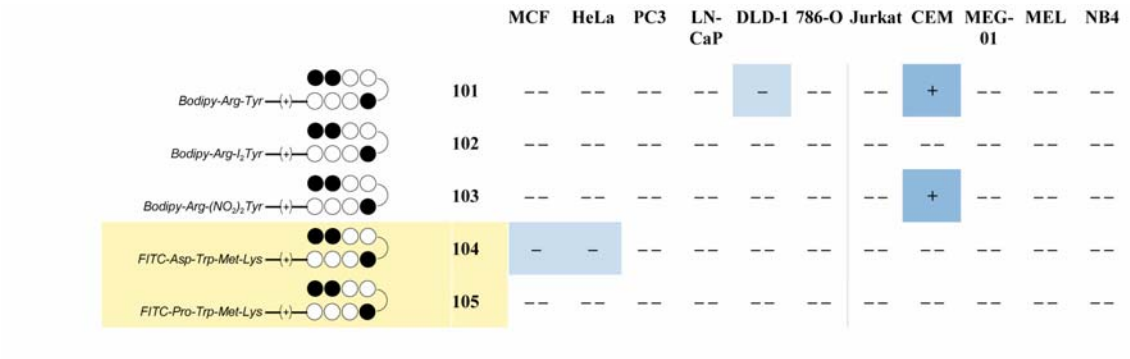


Figure 4.22 Small peptides: uptake profile of polyamides **101**-**105**. Symbols are defined in Figure 4.7. Polyamide concentration was 2 μM.

The role of charge in nuclear uptake is particularly intriguing, though uptake does not appear to correlate with net charge. At neutral pH, fluorescein conjugates that efficiently access cell nuclei have a net charge of -1 (the majority of conjugates, exemplified by **1**) or a net charge of 0 (such as **29**, **31**, and **81**). However, polyamide-fluorescein conjugates have been observed to accumulate in acidic vesicles (reference 21 and data not shown), in which the pH is between ~4.5 and 6.[43] Because the pK_a's of the phenolic protons of fluorescein and OG derivatives are ~6.4 and ~4.7, respectively,[44] the phenol moiety will be fully or partially protonated in acidic vesicles. Thus, in such subcellular structures, polyamide-fluorescein conjugates such as **1** would have a net charge of 0, whereas conjugates such as **29** would have a net charge of +1.

It may be that the ability to change protonation state enhances the nuclear localization properties of fluorescein conjugates relative to rhodamine and bodipy derivatives, which are pH insensitive. TMR and BoFL conjugates having a net charge of 0 (compounds **13** and **14**) or +1 (compounds **8**-**11**, **15** and **83**-**88**) all perform significantly more poorly than analogous fluorescein conjugates. It is remarkable that moving the

positive charge from the linker to the dye is not tolerated whatsoever (compare triamine-linked FAM conjugate **2** and alkyl-linked TMR conjugate **13**, Figure 4.7), whereas relatively well-tolerated alterations include moving the positive charge to the turn (compare FITC conjugates **1** and **78**) and moving the dye to the cap residue, effectively increasing the distance between the anionic and cationic moieties (compare FITC conjugates **1** and **91**).

Conjugation to fluorescein appears to facilitate uptake for this class of molecules, though it is not essential: TMR conjugate **9** is an excellent nuclear stain in several cell lines, and turn-linked BoFL conjugates **83**, **84**, and **88** display improved uptake profiles compared to the tail-linked BoFL derivative **15**. Indeed, polyamides without any attached fluorophore have been shown to induce biological effects in living cells, albeit in a limited number of cell types. Whereas **62** was excluded from cells, an analog of **62** without a dye-label altered gene expression in live lymphoid cells, and a non-dye-labeled chorambucil derivative of **62** was shown to alkylate genomic DNA in live cells.[20] This example highlights the distinct properties of dye-conjugated versus unlabeled polyamides. One cannot be considered as a "proxy" for the other, regarding cellular localization characteristics or other properties. Fluorescent polyamide conjugates are unique because their ability to access nuclei can be assessed directly, and it is our intent to use such labeled polyamides in biological studies involving live cell systems.

Although the uptake data presented here do not allow for prediction of nuclear uptake properties *a priori*, rough design guidelines are apparent. Synthesizing and analyzing a small, focused library of polyamide-fluorophore conjugates appears to be the optimal approach. Key points of variation are ring sequence (if possible on a given

target), dye composition (FITC, FAM, OG 514, etc.) and position of conjugation, turn substitution (γ-turn or ^{AcHN}γ-turn), and β-Ala incorporation. Examples of such focused libraries are compounds **55**-**59**, **72**-**77**, and **96**-**99**.

For use in live-cell studies, polyamides must now be optimized along three axes: DNA-binding affinity and specificity, *in vitro* biochemical activity, and nuclear localization. For the first and second axes, assays such as DNase footprinting, gel shift, and *in vitro* transcription are well established, and studies based on these techniques continue to expand the scope and utility of DNA-binding polyamides.[3,12,15,45] The current study employs an assay based on confocal microscopy to establish a benchmark dataset, comprising 1300 entries, for nuclear localization of DNA-binding polyamide-dye conjugates. Live-cell studies with these compounds are currently in progress, as are efforts to elucidate the energy-dependent mechanism of uptake.[22] Understanding nuclear accessibility in a wide variety of living cells is a minimum first step toward chemical regulation of gene expression with this class of molecules.

Experimental

Chemicals

Polyamides were synthesized by solid phase methods on Boc-β-ala-PAM resin (Peptides International, Louisville, KY)[23] or on Kaiser oxime resin (Nova Biochem, Laufelfingen, Switzerland).[24] Synthetic protocols for second-generation building blocks (incorporated into polyamides shown in Figure 4.8 and polyamides **90** and **91**, Figure 4.17) are essentially as described.[2-4,25] The polyamide precursor to compound **92** was synthesized as follows:

Cbz- γ -ImImPyPy- (*R*) $^{FmocHN}\gamma$ -ImPyPyPy-Oxime was prepared using the Kaiser oxime resin as previously described.[24] The Fmoc-protecting group was cleaved by 20% (v/v) piperidine/DMF cleavage (4 x 3 min) and replaced by a *t*-Boc group upon treatment with Boc_2O, DIEA, and DMF. The polyamide was cleaved from resin by first swelling dried resin in dioxane (500 mg resin, ~1 mL dioxane) for 1 hr, and then adding a solution of 3:1 dioxane:NaOH (1M aqueous) and allowing the resin to cleave at 37ºC for 36 hrs. Filtration of the resin and a second treatment with fresh cleavage solution yields further crude product, all of which was purified by C_{18} reverse-phase HPLC and lyophilized to provide polyamide as a white powder (~10% yield).

Cbz- γ -ImImPyPy- (*R*) $^{BocHN}\gamma$ -ImPyPyPy-COOH (8 μmol) was dissolved in 1:1 absolute EtOH:CH_2Cl_2 (5 mL), to which was added first $Pd(OAc)_2$ (0.2 g, 950 μmol), then ammonium formate (0.3 g, 4.7 mmol). The solution was stirred at room temperature under Ar for 20 min. The solution was centrifuged to remove precipitated PdO, purified by C_{18} reverse-phase HPLC, and lyophilized to provide polyamide as a white powder (~25% yield).

H_2N- γ -ImImPyPy- (*R*) $^{BocHN}\gamma$ -ImPyPyPy-COOH (2.2 μmol) was dissolved in DMF (4 mL). K_2CO_3 (0.11 g, 0.8 mmol) that had been dried at 130ºC was added and the mixture was stirred for 1 hr. DPPA (14.5 μL, 67 μmol) was added and the mixture was stirred for 4 hrs. The solution was concentrated to a brown film, which was treated with 80% (v/v) TFA/CH_2Cl_2 for 45 min, concentrated, purified by C_{18} reverse-phase HPLC, and lyophilized to provide polyamide as a white powder (~25% yield). Polyamide cyclo-(γ-ImImPyPy- (*R*)$^{H_2N}\gamma$-ImPyPyPy-) was treated with fluorescein amide-C_6 linker

succinimidyl ester as with other dye additions, purified, and lyophilized to provide **92** as an orange powder (~50% yield).

U-pin **93** and H-pins **94**-**99** were synthesized according to methods described in references (26) and (27), respectively. After cleavage with the appropriate amine and reverse-phase HPLC purification, polyamides were dissolved in DMF and treated with diisopropylethylamine (DIEA) (20 eq) followed by the fluorophore in the form of an isothiocyanate, an *N*-hydroxysuccinimidyl ester, or a free acid activated *in situ* with HBTU. Fluorophores were delivered as solutions in DMF or DMSO. After reacting at rt for ~3 h, the resulting dye conjugates were purified by HPLC.

Peracetylated polyamides **17**, **34**, **95**, **98**, and **99** were obtained by treating the precursor polyamides with acetic anhydride and DIEA in DMF for ~30 min. For compounds with a free (*R*)-2,4-diaminobutyric acid ($^{H_2N}\gamma$-turn) moiety (**29**-**31**, **76**, **78**, **79**, and **100**), the protected $^{FmocHN}\gamma$-turn amine was deprotected (20% piperidine-DMF) and reprotected as the Boc derivative (Boc_2O, DIEA, DMF) immediately prior to cleavage of the polyamide from resin. After conjugation of the dye moiety, these compounds were treated with TFA and purified by HPLC.

Compounds **101**-**105** were prepared as shown in Figure 4.21. The peptides for **104** and **105** were prepared on SASRIN resin (Bachem), cleaved with 1% (v/v) TFA in CH_2Cl_2, and purified on silica, retaining the side-chain protection on the aspartic acid and tryptophan residues. The identity and purity of each compound was verified by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF). Table 4.2 lists the masses calculated and found for all conjugates. All fluorescent dye reagents were from Molecular Probes. Chemicals

not otherwise specified were from Sigma-Aldrich.

Cell Cultures

The human cancer cell lines MCF-7, PC3, LNCaP, DLD-1, 786-O, Jurkat, CCRF-CEM (CEM), MEG-01, and NB4 were cultured in a 5% CO_2 atmosphere at 37°C in supplemented RPMI medium 1640. The human cancer cell line HeLa, the murine leukemia cell line MEL, and the transformed human kidney cell line 293 were grown as above in supplemented DMEM. The human cancer cell line SK-BR-3 was cultured as above in supplemented McCoy's medium. All media were supplemented with 10% fetal bovine serum (Irvine Scientific) and 1% penicillin/streptomycin solution (Sigma).

Confocal Microscopy

Adherent cell lines (MCF-7, HeLa, PC3, LNCaP, SK-BR-3, DLD-1, 786-O, and 293) were trypsinized for 5–10 min at 37°C, centrifuged for 5 min at 5°C at 900 *g*, and resuspended in fresh medium to a concentration of 1.25 x 10^6 cells per ml. Suspended cell lines (Jurkat, CEM, MEG-01, MEL, and NB4) were diluted in fresh medium to the same concentration. Incubations were performed by adding 150 µl of cells into culture dishes equipped with glass bottoms for direct imaging (MatTek, Ashland, MA). Adherent cells were grown in the glass-bottom culture dishes for 24 h. The medium was then removed and replaced with 147 µl (or 142.5 µl) of fresh medium. Then 3 µl (or 7.5 µl) of the 100 µM polyamide solution was added for final polyamide concentration of 2 µM (or 5 µM), as noted in the data tables. Cells were incubated in a 5% CO_2 atmosphere at 37°C for 10–14 h. Suspended cell line samples were prepared in a similar fashion, omitting

trypsinization. These samples were then incubated as above for 10–14 h. Imaging was performed with a x40 oil-immersion objective lens on a Zeiss LSM 5 Pascal inverted laser scanning microscope. The optical slice was set to 2.2 μm. Images were line-averaged 2, 4, 8, or 16 times and were obtained at a 0.8 μs/pixel scanning rate. Polyamide-dye conjugate fluorescence and visible light images were obtained using standard filter sets appropriate for fluorescein, rhodamine, or Texas Red.

Acknowledgements

We are grateful to the National Institutes of Health for support (Grant GM57148) and for predoctoral support to T.P.B. and R.M.D. (Grant T32-GM08501), to the Howard Hughes Medical Institute for a fellowship to B.S.E., and to the National Science Foundation for a predoctoral fellowship to S.F. Mass spectral analyses were performed in the Mass Spectrometry Laboratory of the Division of Chemistry and Chemical Engineering of Caltech, supported in part by National Science Foundation Materials Research Science and Engineering program.

Table 4.2 Mass spectrometry data for compounds **1-105**.

Compound	M+(ion)	calc'd	found	Compound	M+(ion)	calc'd	found
1	H	1584.6	1584.5	**51**	H –FITC	1393.7	1393.9
2	H	1553.6	1553.4	**52**	H	1751.7	1751.9
3	H	1589.6	1589.5	**53**	H –FITC	1394.7	1395.0
4	H	1625.6	1625.6	**54**	H	1752.7	1752.9
5	H	1689.6	1689.5	**55**	H	1611.6	1611.7
6	Na	1703.6	1703.6	**56**	H –FITC	1446.7	1447.0
7	H	1711.6	1711.7	**57**	H	1804.7	1804.8
8	H	1551.7	1551.8	**58**	H –FITC	1395.7	1396.0
9	H	1607.7	1607.7	**59**	H	1753.7	1753.5
10	H	1711.8	1711.8	**60**	H –FITC	1396.7	1396.7
11	H	1896.8	1896.8	**61**	H –FITC	1396.7	1396.6
12	H	1569.6	1569.6	**62**	H –FITC	1395.7	1395.6
13	H	1592.7	1592.5	**63**	H –FITC	1194.6	1194.6
14	-F	1434.7	1434.1	**64**	H	1654.7	1654.8
15	Na	1490.7	1490.7	**65**	H	1584.6	1585.1
16	H	1641.6	1641.5	**66**	H	1655.7	1655.9
17	H –Ac_2FITC	1274.6	1274.6	**67**	H	1553.1	1553.1
18	H	1641.6	1640.7	**68**	H –FITC	1195.6	1195.7
19	H	1746.6	1746.6	**69**	H –FITC	1266.6	1266.7
20	H	1664.7	1664.7	**70**	H	1585.6	1585.2
21	H –FITC	1231.5	1231.6	**71**	H	1656.7	1656.8
22	H	1589.5	1589.6	**72**	Na –FITC	1417.7	1417.7
23	H	1725.5	1725.6	**73**	H	1727.7	1727.7
24	H	1643.6	1643.8	**74**	H	1696.7	1696.7
25	H –FITC	1232.5	1232.6	**75**	H	1767.7	1767.7
26	H –FITC	1232.5	1232.3	**76**	H	1782.7	1782.6
27	H –FITC	1212.5	1212.5	**77**	H	1824.8	1824.5
28	H –FITC	1476.6	1476.0	**78**	H	1584.6	1584.7
29	H	1599.6	1599.6	**79**	Na	1592.6	1592.4
30	H	1598.7	1598.6	**80**	H	1683.7	1683.7
31	H	1635.5	1636.1	**81**	H	1726.7	1726.7
32	H –FITC	1241.5	1241.6	**82**	H	1754.7	1754.6
33	H	1735.5	1735.4	**83**	Na	1562.7	1562.5
34	H –OG514	1283.5	1283.2	**84**	Na	1561.7	1561.6
35	H	1749.5	1749.6	**85**	Na	1562.7	1562.6
36	H	1719.5	1719.6	**86**	Na	1562.7	1562.6
37	H –FITC	1073.5	1073.5	**87**	Na	1563.7	1563.5
38	H	1567.5	1567.7	**88**	Na	1805.8	1805.8
39	H –FITC	1317.6	1317.3	**89**	H	1583.6	1583.6
40	H	1811.6	1812.0	**90**	H	1513.5	1513.3
41	H –FITC	1373.7	1373.7	**91**	H	1584.6	1584.6
42	H	1731.7	1731.7	**92**	H	1636.6	1636.4
43	H –FITC	1374.6	1374.8	**93**	H	1541.6	1541.6
44	H	1732.7	1732.6	**94**	H	1438.7	1438.4
45	H –FITC	1445.7	1445.6	**95**	H	1564.7	1564.3
46	H	1803.7	1804.0	**96**	H	1926.9	1926.3
47	H –FITC	1461.7	1461.8	**97**	H	2315.9	2319.0
48	H –FITC	1460.7	1460.7	**98**	H	2052.9	2052.2
49	H –FITC	1495.7	1495.9	**99**	H	2195.0	2194.4
50	H	1853.8	1853.9	**100**	H	2270.9	2271.3

101	**H**	**1747.8**	**1749.6**
102	**H**	**1999.6**	**1998.7**
103	**H**	**1837.8**	**1838.6**
104	**H**	**2144.9**	**2149.3**
105	**H**	**2126.9**	**2127.6**

References

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

129) Briehn,C.A.; Weyermann,P.; Dervan,P.B. *Chem.-Eur. J.* **2003**, *9*, 2110.

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

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

132) Dickinson, L.A.; Trauger, J.W.; Baird, E.E.; Ghazal, P.; Dervan, P.B.; Gottesfeld, J.M. *Biochemistry* **1999**, *38*, 10801.

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

134) McBryant,S.J.; Baird,E.E.; Trauger,J.W.; Dervan,P.B.; Gottesfeld,J.M. *J. Mol. Biol*. **1999**, *286*, 973.

135) Chiang, S.Y.; Bürli, R.W.; Benz, C.C.; Gawron, L.; Scott, G.K.; Dervan, P.B.; Beerman, T.A. *J. Biol. Chem*. **2000**, *275*, 24246.

136) Wang, C.C.C.; Dervan, P.B. *J. Am. Chem. Soc*. **2001**, *123*, 8657.

137) Ehley, J.A.; Melander, C.; Herman, D.; Baird, E.E.; Ferguson, H.A.; Goodrich, J.A.; Dervan, P.B.; Gottesfeld, J.M. *Mol. Cell. Biol*. **2002**, *22*, 1723.

138) Yang, F.; Belitsky, J.M.; Villanueva, R.A.; Dervan, P.B.; Roth, M.J. *Biochemistry*, **2003**, *42*, 6249.

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

140) Wurtz, N.R.; Pomerantz, J.L.; Baltimore, D.; Dervan, P.B. *Biochemistry*, **2002**, *41*, 7604.

141) Arora, P.S.; Ansari, A.Z.; Best, T.P.; Ptashne, M.; Dervan, P.B. *J. Am. Chem. Soc.* **2002**, *124*, 13067.

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

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

144) Janssen,S.; Cuvier,O.; Muller,M.; Laemmli,U.K. *Mol. Cell.* **2000**, *6*, 1013.

145) Wang, Y.D.; Dziegielewski, J.; Wurtz, N.R.; Dziegielewska, B.; Dervan, P.B.; Beerman, T.A. *Nucleic Acids Res*. **2003**, *31*, 1208.

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

147) Dudouet, B.; Burnett, R.; Dickinson, L.A.; Wood, M.R.; Melander, C.; Belitsky, J.M.; Edelson, B.S.; Wurtz, N.R.; Briehn, C.; Dervan, P.B.; Gottesfeld, J.M. *Chem. Biol.* **2003**, *10*, 859.

148) Crowley, K.S.; Phillion, D.P.; Woodard, S.S.; Schweitzer, B.A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J.K. *Bioorg. Med. Chem. Lett*. **2003**, *13*, 1565.

149) Best, T.P.; Edelson, B.S.; Nickols, N.G.; Dervan, P.B. *Proc. Natl. Acad. Sci. U.S.A*. **2003**, *100*, 12063.

150) Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1996**, *118*, 6141.

151) Belitsky, J.M.; Nguyen, D.H.; Wurtz, N.R.; Dervan, P.B. *Bioorg. Med. Chem*. **2002**, *10*, 2767.

152) Wurtz, N.R. California Institute of Technology: Pasadena, CA, 2002.

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

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

155) Delmotte, C.; Delmas, A. *Bioorg. Med. Chem. Lett*. **1999**, *9*, 2989.

156) Gygi, M.P.; Ferguson, M.D.; Mefford, H.C.; Lund, K.P.; O'Day, C.; Zhou, P.; Friedman, C.; van den Engh, G.; Stolowitz, M.L.; Trask, B.J. N*ucleic Acids Res*. **2002**, *30*, 2790.

157) Maeshima, K.; Jansses, S.; Laemmli, U.K. *EMBO J*. **2001**, *20*, 3218.

158) Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Chem. Biol*. **1996**, *3*, 369.

159) Trauger, J.W.; Baird, E.E.; Dervan, P.B. *Angew. Chem.-Int. Ed*. **1998**, *37*, 1421.

160) Schaal, T.D.; Mallet, W.G.; McMinn, D.L.; Nguyen, N.V.; Sopko, M.M.; John, S.; Parekh, B.S. Nucleic Acids Res. 2003, 31, 1282.

161) Swalley, S.E.; Baird, E.E.; Dervan, P.B. *Chem.-Eur. J*. **1997**, *3*, 1600.

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

163) Gottesfeld, J.M.; Belitsky, J.M.; Melander, C.; Dervan, P.B.; Luger, K. *J. Mol. Biol*. **2002**, *321*, 249.

164) Rucker, V.C.; Foister, S.; Melander, C.; Dervan, P.B. *J. Am. Chem. Soc*. **2003**, *125*, 1195.

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

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

167) Baliga, R.; Baird, E.E.; Herman, D.M.; Melander, C.; Dervan, P.B.; Crothers, D.M. *Biochemistry* **2001**, *40*, 3.

168) Greenberg, W.A.; Baird, E.E.; Dervan, P.B. *Chem.-Eur. J.* **1998**, *4*, 796.

169) Kers, I.; Dervan, P.B. *Bioorg. Med. Chem*. **2002**, *10*, 3339.

170) Lodish, H.; Berk, A.; Zipursky, L.S.; Matsudaira, P.; Baltimore, D.; Darnell, J. (2000) *Molecular Cell Biology*. 4th ed. W.H. Freeman and Co., New York, NY.

171) Lin, H.J.; Szmacinski, H.; Lakowicz, J.R. *Anal. Biochem*. **1999**, *269*, 162.

172) Marques, M.A.; Doss, R.M.; Urbach, A.R.; Dervan, P.B. *Helv. Chim. Acta* **2002**, *85*, 4485.

Chapter 4C

DNA-Binding Characteristics of Polyamide-Fluorophore Conjugates

Abstract

The successful staining of the nuclei of mammalian cells with polyamides labeled with fluorescein and other similar dyes permits many transcription inhibition experiments to be attempted with confidence that a negative result is not due to exclusion of active compound from genomic DNA. However, accompanying the need for good cellular uptake properties is the need for polyamides or polyamide conjugates to possess good DNA-binding properties. The effect of fluorophore conjugation on the affinity and specificity of polyamides for DNA is an important question pertaining to the use of these compounds as transcriptional regulators.

Several polyamide-fluorophore conjugates have been synthesized and their DNA-binding properties tested by DNase I footprinting titration experiments. In general, fluorescein conjugation is most tolerated at the *C*-terminal tail, resulting in a 5- to 25-fold decrease in affinity while retaining specificity. The effects of attachment off of the chiral turn residue and the N-terminal imidazole 4-amino group were also probed. Fluorescein conjugation to either of these positions resulted in a ~100-fold decrease in binding affinity with reference to the tail conjugate. Surprisingly, attachment of a Bodipy fluorophore from the turn residue results in a somewhat lesser 10-fold decrease in affinity, as compared to the tail-linked fluorescein conjugate.

Results

Compounds **1-6** were synthesized, linking the fluorescein moiety through the polyamide *C*-terminal tail (Figure 4.23). These polyamides target several different DNA binding sites, and were all shown to stain the nuclei of several different cell lines. Polyamides **1-3** were footprinted on a plasmid derived from the γ-globin promoter (Figure 4.24).

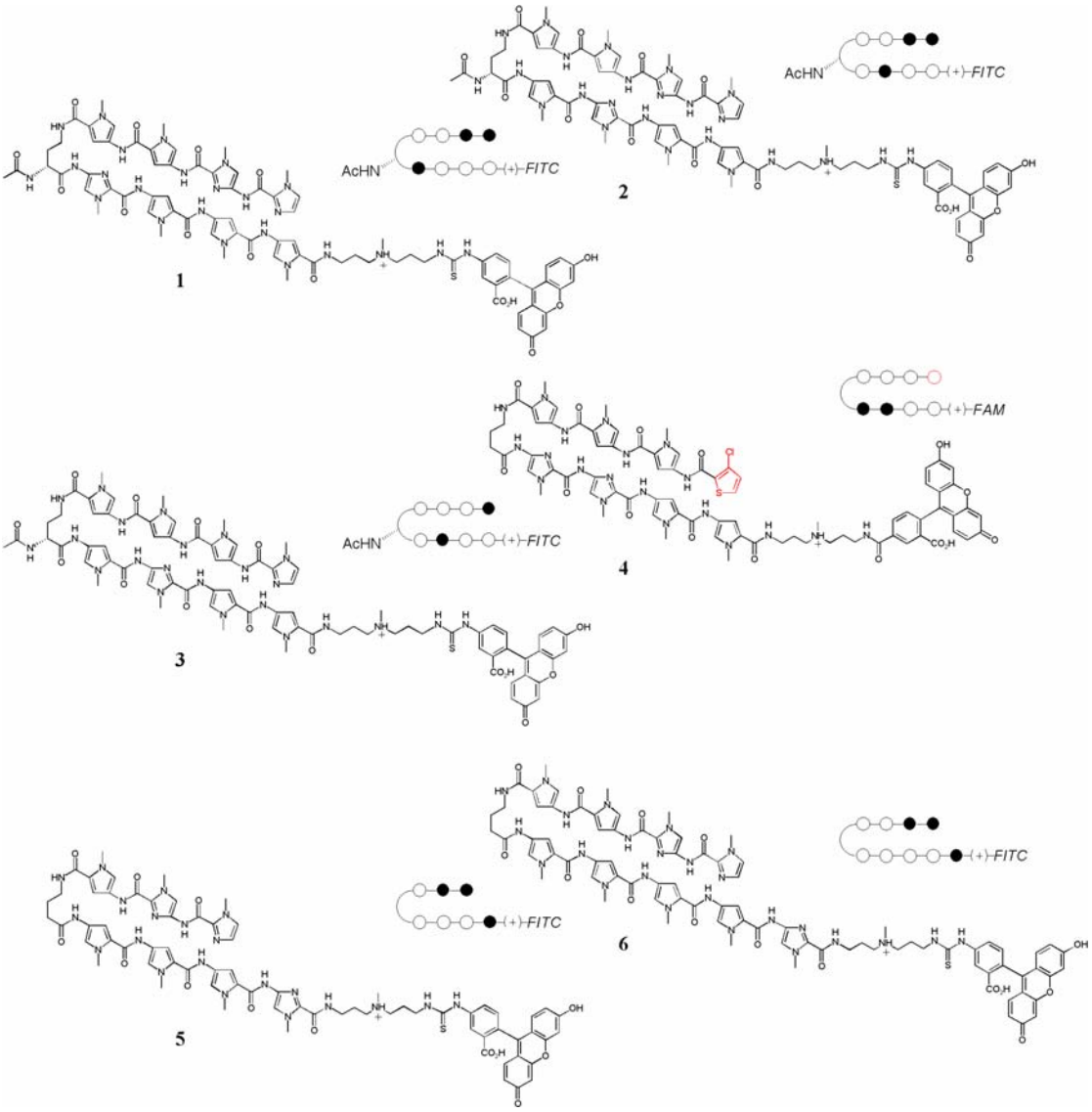


Figure 4.23 Chemical structures and ball-and-stick models of tail-linked polyamide-FITC conjugates **1-6**.

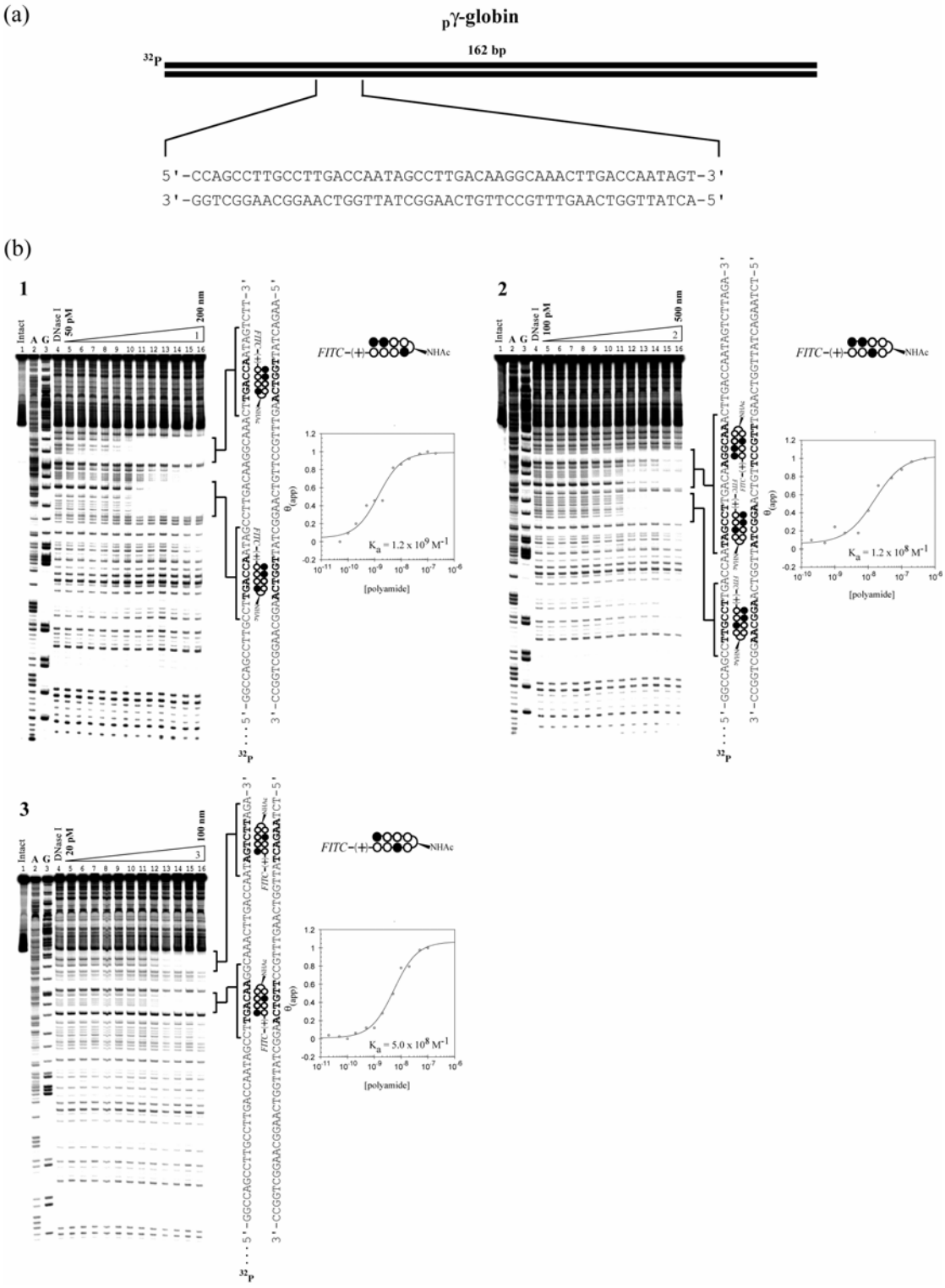
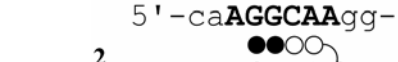
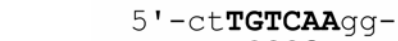
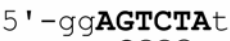


Figure 4.24 Footprinting of **1-3**. (a) γ-globin promoter PCR fragment used for footprinting reactions. (b) Footprinting gels and isotherms for compounds **1-3**.

The affinities of compounds **1-3** are reduced with reference to polyamides lacking the fluorescein moiety. The precise degree of this reduction is complicated by two factors. First, it is unclear which are the best choices of parent polyamides against which to compare binding characteristics. Second, for truly accurate comparisons to be made, both the polyamide and its conjugate must be footprinted in exactly the same sequence context, preferably on the same plasmid. Though imperfect, it is possible to make useful comparisons between the footprinting data reported here and past data.

Table 4.3 shows comparative data between compounds **1-3** and related polyamides **7-9**. There is a 4- to 25-fold decrease in binding affinity upon addition of the fluorescein moiety. The large differences in affinity decrease between compounds is most likely due to the imperfect comparisons that may be drawn between **1-3** and **7-9**, as well as the differing DNA contexts. However, the general effect of FITC addition seems to be an approximately 10-fold reduction in affinity, with good specificity.

Table 4.3 Comparison between the DNA-binding affinities of polyamides and polyamide-FITC conjugates.

	polyamide-FITC conjugate	K_a (M^{-1})		*polyamide*	K_a (M^{-1})
1	5'-at**TGGTCA**ag-3'  5'-ta**ACCAGT**tc-3'	1.2×10^9	**7**	5'-ta**TGGTCA**tg-3'  5'-at**ACCAGT**ac-3'	4.8×10^9
2	5'-ca**AGGCAA**gg-3'  5'-gt**TCCGTT**cc-3'	1.2×10^8	**8**	5'-gc**AGGCAA**cc-3'  5'-cg**TCCGTT**gg-3'	3.0×10^9
3	5'-ct**TGTCAA**gg-3'  5'-ga**ACAGTT**cc-3'	5.0×10^8	**9**	5'-gg**AGTCTA**ta-3'  5'-cc**TCAGAT**at-3'	2.0×10^9

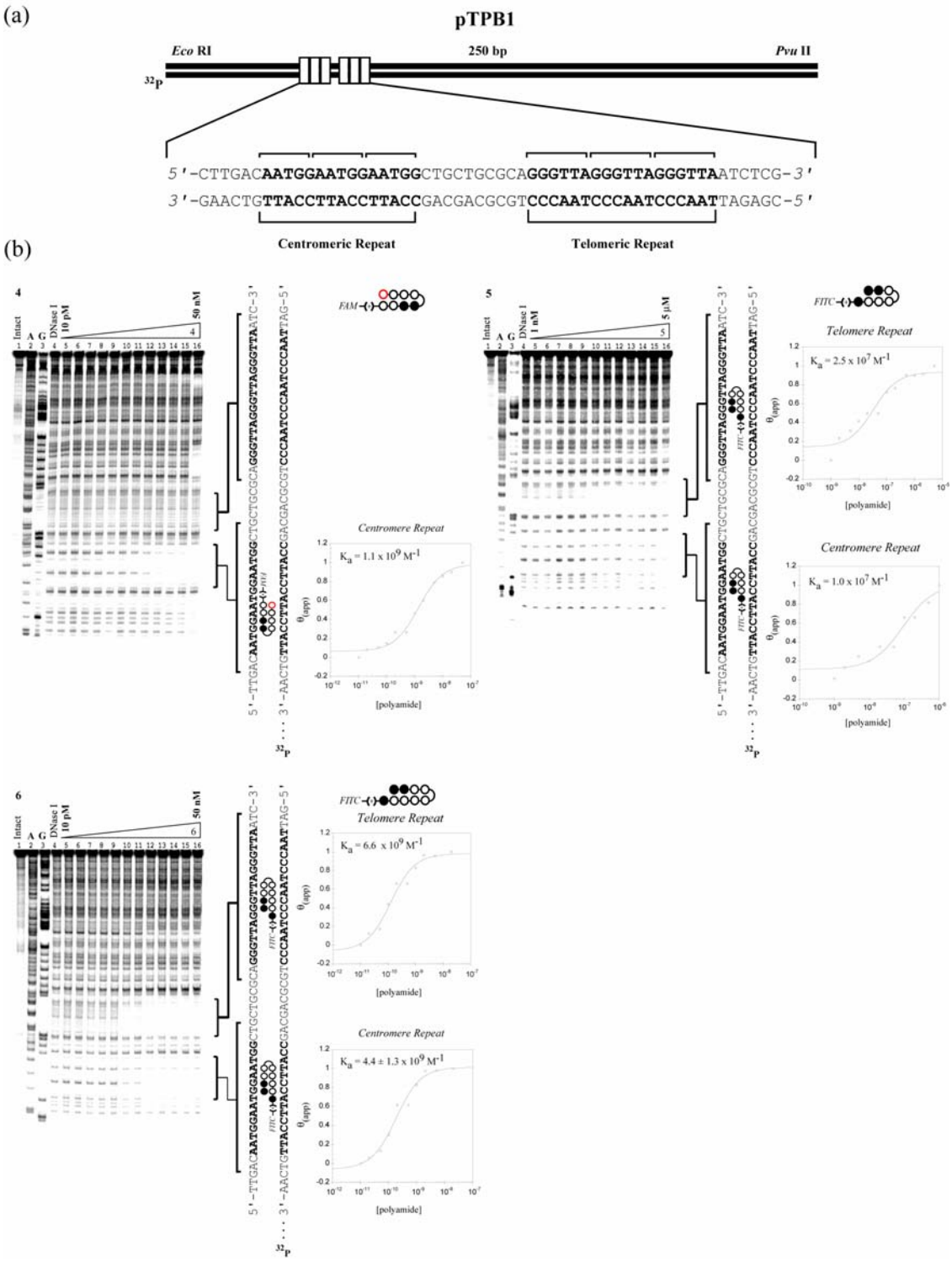


Figure 4.25 Footprinting of **4-6**. (a) pTPB1 restriction fragment used for footprinting reactions. (b) Footprinting gels and isotherms for compounds **4-6**.

The affinities of compounds **4** and **6** are similar in magnitude to those of compounds **1**-**3**, while the affinity of 7-ring compound **5** is substantially lower, probably reflecting a parent polyamide with poor affinity (Figure 4.25). Though their affinities are almost certainly reduced with reference to polyamides lacking the fluorescein moiety, there is no footprinting data available for likely parent polyamides. As for compounds **10**-**13**, they all are related to compound **7**, to varying degrees, and even more closely to compound **1**, differing mainly in the placement of the fluorophore moiety (Figure 4.26). Only compound **10**, which is structurally very similar to **1**, possesses a low nanomolar binding affinity. None of the other compounds have affinities as high as **1**, though surprisingly the Bodipy compound **13** has a 10-fold higher affinity than either **11** or **12**, which are both 100-fold lower than **1** (Figure 4.27). It should be noted that while compounds **10** and **13** were footprinted on the plasmid pDEH9, **11** and **12** were footprinted on the γ-globin plasmid (Figure 4.24a).

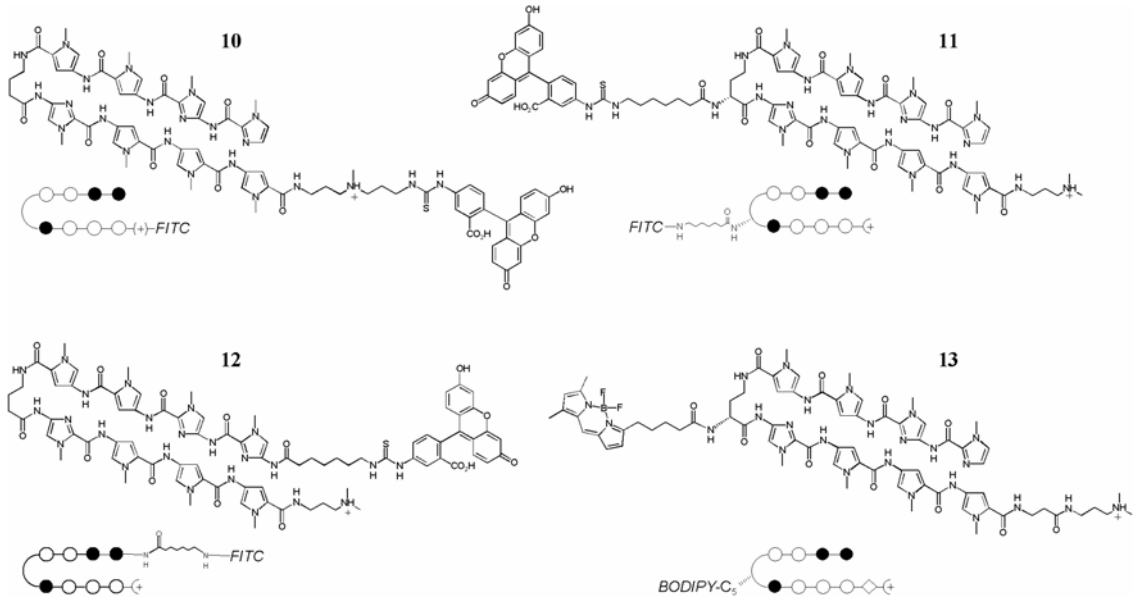


Figure 4.26 Chemical structures and ball-and-stick models of compounds **10**-**13**.

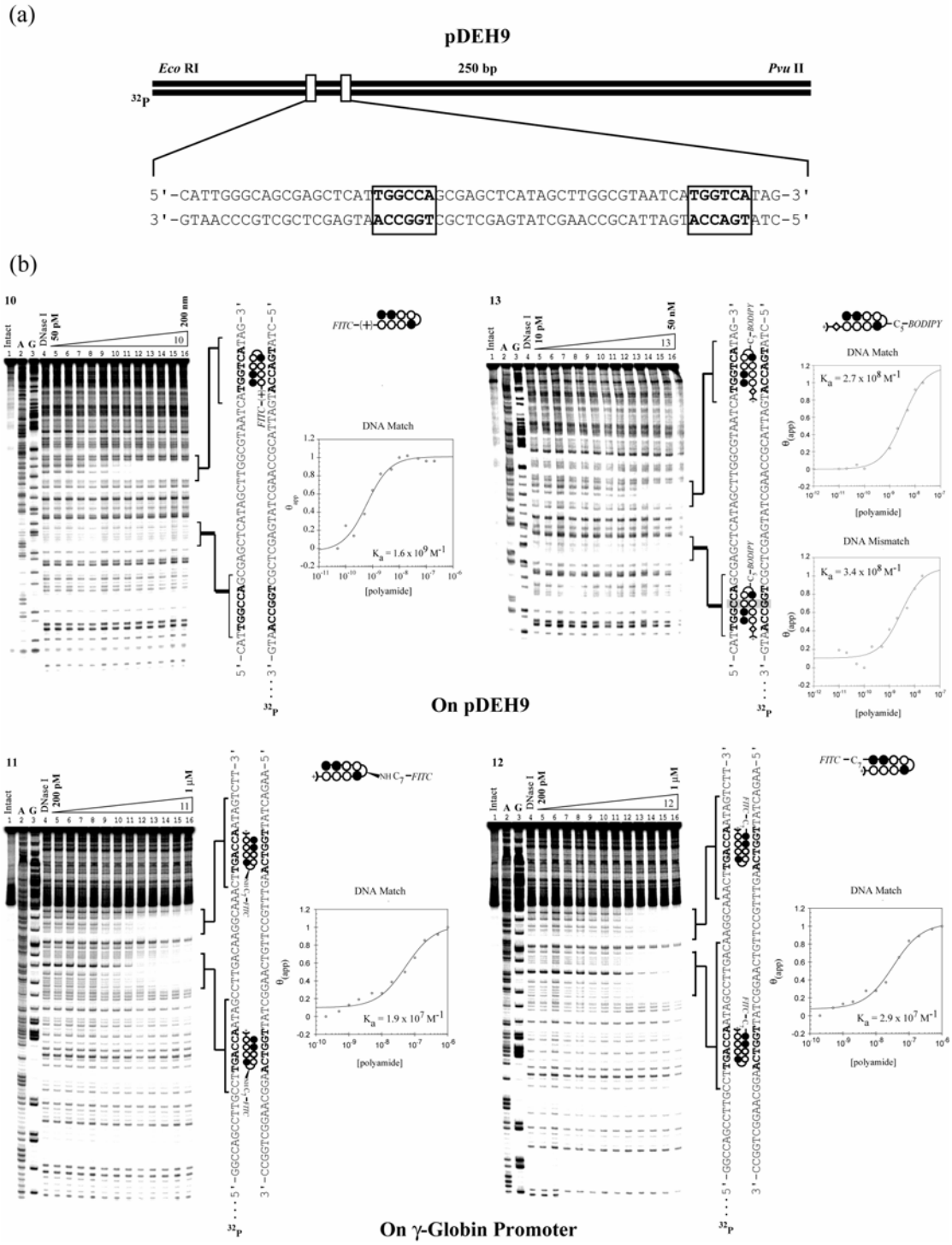


Figure 4.27 Footprinting of **10-13**. (a) pDEH9 restriction fragment used for footprinting reactions. (b) Footprinting gels and isotherms for compounds **10-13**.

Table 4.4 Comparison of alternate motifs for polyamide-FITC conjugates.

	compound	K_a (M^{-1})
7	5'-ta**TGGTCA**tg-3' 5'-at**ACCAGT**ac-3'	4.8×10^9
1	5'-at**TGGTCA**ag-3' FITC, NHAc 5'-ta**ACCAGT**tc-3'	1.2×10^9
10	5'-ca**TGGTCA**ta-3' FITC 5'-gt**ACCAGT**at-3'	1.6×10^9
11	5'-at**TGGTCA**ag-3' FITC 5'-ta**ACCAGT**tc-3'	1.9×10^7
12	5'-at**TGGTCA**ag-3' FITC 5'-ta**ACCAGT**tc-3'	2.9×10^7
13	5'-ca**TGGTCA**ta-3' BODIPY 5'-gt**ACCAGT**at-3'	2.7×10^8

An examination of the binding affinities of compounds **10-13** in comparison with those of **1** and **7** show that attachment of the fluorophore moiety from either the turn or the N-terminal imidazole 4-amino- group causes a large reduction in binding affinity. Attachment of FITC from the N-terminus results in 100- to 1000-fold reduction of affinity with reference to a parent polyamide. A comparable affinity reduction is seen when FITC is attached from the chiral turn via a C_7 aliphatic linker. It turns out that this affinity can be improved by attaching Bodipy in this position, via a C_5 aliphatic linker, for an overall 10- to 100-fold reduction in affinity with reference to polyamide. Attachment of these fluorophores from alternate positions seems to lead to a loss of sequence specificity in addition to the affinity drop. In general, this bodes poorly for using polyamide-fluorophore conjugates connected through alternate strategies as reagents to regulate gene transcription in living systems, despite the ready nuclear localization of compounds **11-13**.

Experimental

Chemicals

All polyamide precursors to **1-6** and **10-13** were synthesized on either Boc-β-Ala-PAM resin or Kaiser oxime resin, as previously described.[1,2] Polyamides were labeled with fluorophores by treatment with fluorescein isothiocyanate (1.5 eq) and DIEA (>25 eq) in DMF (100-200 μL).

DNase I Footprinting Titrations

5' or 3' ^{32}P-labeled PCR fragments were generated from template plasmid p γ-globin, pDEH9, or pTPB1 in accordance with standard protocols and isolated by nondenaturing gel electrophoresis.[3] All DNase I footprinting reactions were carried out in a volume of 400 μL. A polyamide stock solution or water (for reference lanes) was added to TKMC buffer, with final concentrations of 50 mM Tris-HCl, 50 mM KCl, 50 mM $MgCl_2$, and 25 mM $CaCl_2$, pH 7.0, and 15 kcpm 5'-radiolabeled DNA. The solutions were equilibrated for 12-18 h at 22°C. Cleavage was initiated by the addition of 10 μL of a DNase I stock solution and was allowed to proceed for 7 min at 22°C. The reactions were stopped by adding 50 μL of a solution containing 2.25 M NaCl, 150 mM EDTA, 0.6 mg/mL glycogen, and 30 μM base pair calf thymus DNA and then ethanol precipitated. The cleavage products were resuspended in 100 mM Trisborate-EDTA/80% formamide loading buffer, denatured at 85°C for 10 min, and immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V for 2 h and 15 min. The gels were dried under vacuum at 80 °C and quantitated using storage phosphor technology.

References

173) Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1996**, *118*, 6141.

174) Belitsky, J.M.; Nguyen, D.H.; Wurtz, N.R.; Dervan, P.B. *Bioorg. Med. Chem*. **2002**, *10*, 2767.

175) Trauger, J.W.; Dervan, P.B. *Methods Enzymol*. **2001**, *340*, 450.

Chapter 5

Inhibition of Transcription on the Androgen Response Element with Polyamides and Polyamide Conjugates

Abstract

Upon binding an androgen molecule, androgen receptor (AR) acts as a transcriptional activator on genes bearing one of the several androgen-response elements (ARE) in the 5' upstream promoter region. The prostate specific antigen (PSA) gene is among several genes regulated through an ARE. The AR-ARE interaction is inhibited upon treatment with minor groove binding polyamides targeted to the ARE site. Inhibition of AR binding to the ARE is expected to decrease the transcription of PSA, which is shown in a transient transfection assay utilizing a luciferase gene under control of an ARE. Attempts to extend these results to native gene transcription in live cells were unsuccessful. Polyamide-fluorescein conjugates of the active polyamides were found to spontaneously enter the nuclei of LNCaP cells. These conjugates were also found to bind the ARE sequence-specifically and with high affinity. Preliminary inhibition results show only modest inhibition with the polyamide-fluorescein compounds. However, higher concentrations and/or treatment with multiple compounds may show increased PSA inhibition activity. The synthesis of higher-affinity compounds that target the same DNA sequence in the ARE would also likely effect higher transcription inhibition results.

Introduction

The androgen receptor (AR) is a ligand-activated transcription factor that is part of the steroid hormone receptor superfamily of proteins.[1-3] By binding their specific receptor proteins, steroid hormones act to coordinate many complex series of events involved in the development, differentiation, and physiological response of cells to a variety of stimuli. Upon binding a steroid hormone molecule, due to allosteric reorganization receptor proteins become able to bind short (~20 bp) *cis*-acting hormone receptor elements (HRE).[4] AR action is initiated upon binding of testosterone (T) and, especially, 5α-dihydrotestosterone (DHT). The AR regulates gene expression upon binding to androgen-response elements (ARE) located in the 5' flanking region of androgen target genes.

Many genes are regulated through ARE domains, including some found in kidney, liver, and especially prostate tissue. Chief among the latter is the prostate specific antigen (PSA), which is frequently used as a marker for the diagnosis of prostate cancer. The androgen receptor, like the similar glucocorticoid and progesterone receptors, binds its response element as homodimers.[5] The study of AR-regulated genes has produced a consensus ARE sequence: 5'-GGWACAnnnTGTTCT-3', along with other related nonconsensus sequences. The PSA ARE domain is a nonconsensus 5'-AGAACAgcaAGTGCT-3' sequence.[6] From crystal structure studies, it was shown that the capitalized hexamer sequences are key to the interaction of the AR with its DNA binding site (Figure 5.1).[7,8] Though the protein interacts mainly with the major groove, each unit of the homodimer injects a lysine residue into the minor groove. Based on the

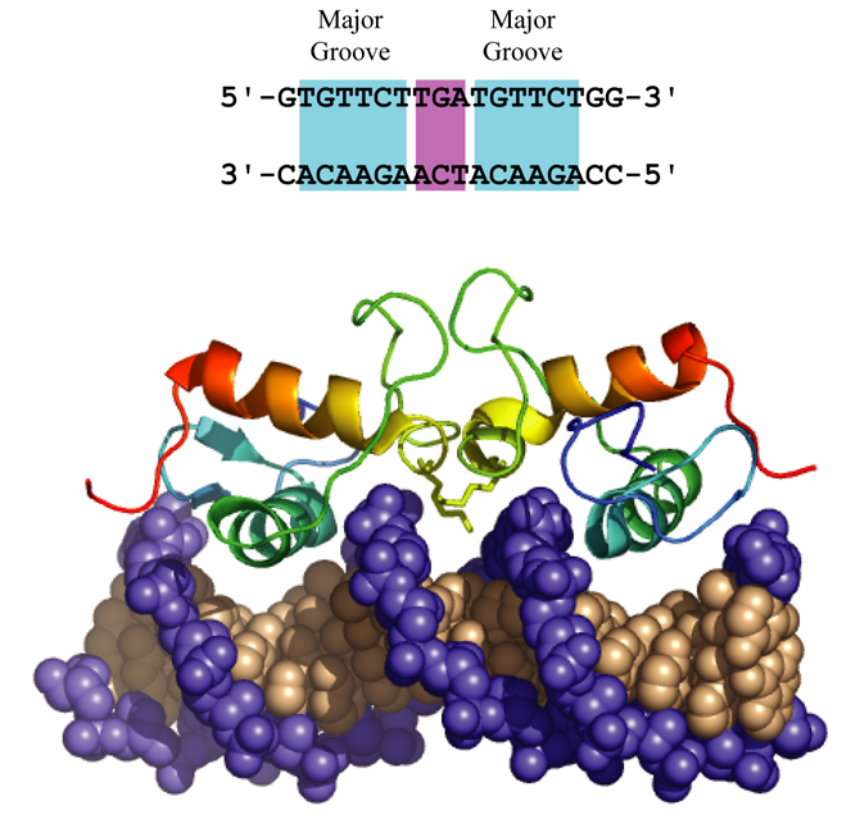


Figure 5.1 Androgen receptor binding DNA. The two homodimeric AR chains bind the consecutive hexameric half-site direct repeat 5'-ACAAGA-3' units utilizing mainly major groove contacts (blue-green helices lying in the major grooves). Each AR molecule projects a lysine (yellow) into the minor groove.[8]

structural and biochemical data, polyamides were synthesized and tested for their ability to inhibit the binding of AR to ARE.

Results

Polyamides **1-3** were synthesized based on their projected ability to bind the PSA ARE domains (Figure 5.2). These compounds were sent to Dr. Zhengxin Wang at the M.D. Anderson Cancer Center at the University of Texas Southwestern for testing as inhibitors of AR-ARE interaction. Gel mobility shift assays[9] showed promising results, with compounds **1** and **2** showing complete inhibition of mobility at concentrations $\leq$ 1

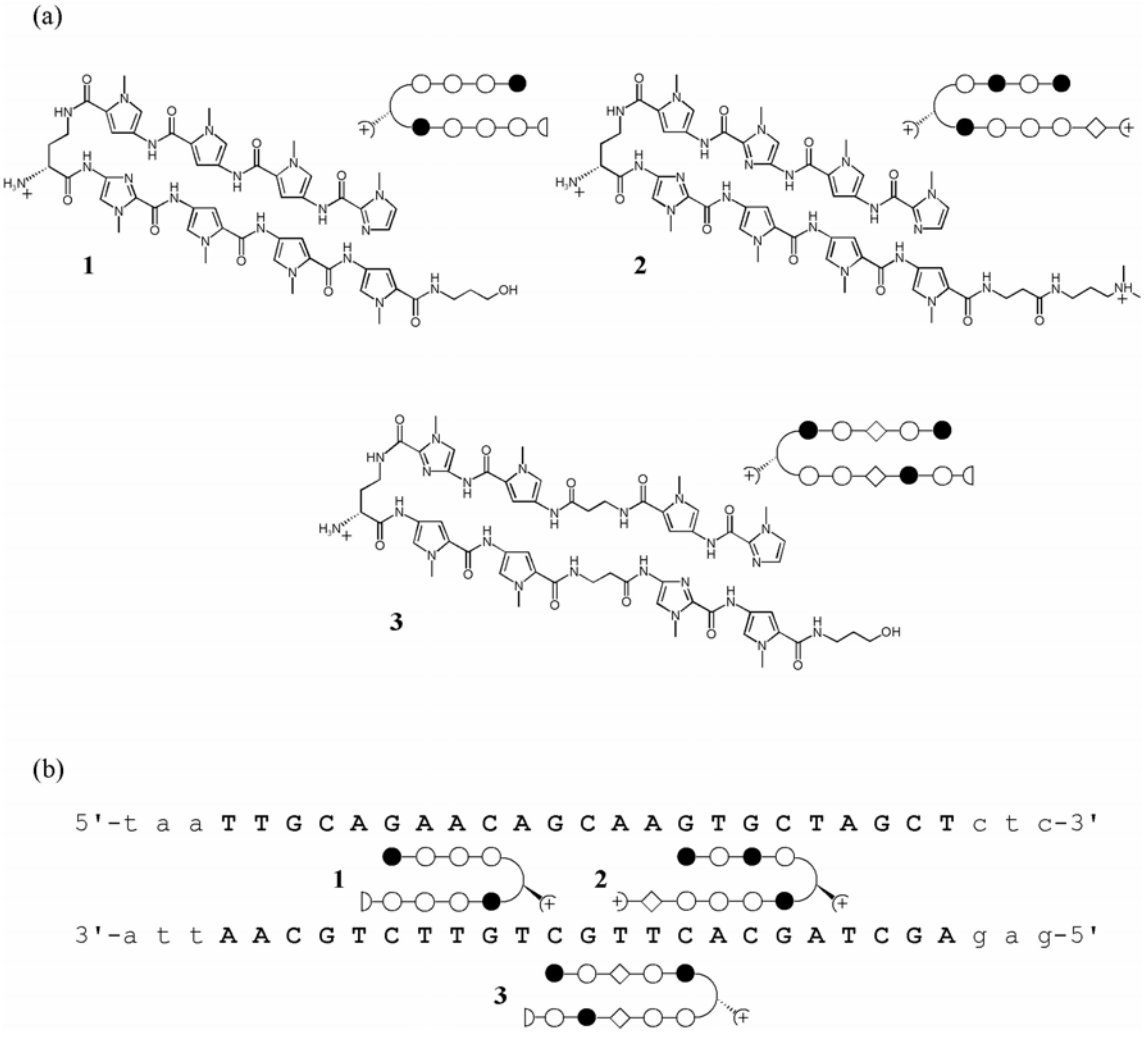


Figure 5.2 Compounds for AR inhibition. (a) Chemical structures and ball-and-stick models of compounds **1**-**3**. (b) Putative binding sites for **1**-**3** on the PSA ARE (in boldface).

nM. Compound **3** also inhibits AR-ARE interaction, but at higher concentrations (≥ 4 nM). Based on these positive results, the compounds were tested in DNase I footprinting titration assays using a plasmid containing the PSA promoter (Figure 5.3). These compounds were then tested in an *in vitro* transcription assay using a DNA construct containing a luciferase gene under transcriptional control of a PSA ARE. The polyamides **1**-**3** also showed strong inhibition in this system (Figure 5.4).

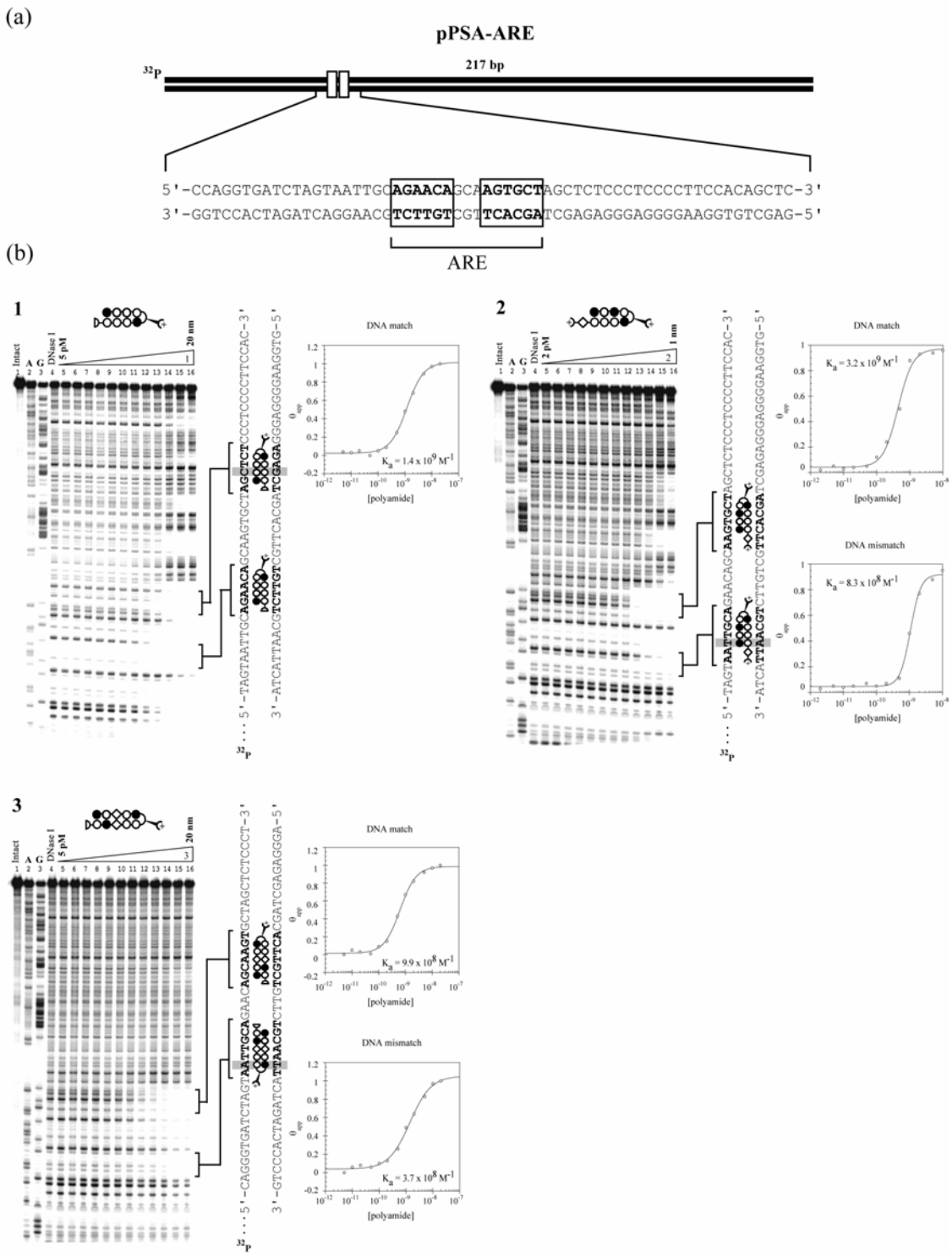
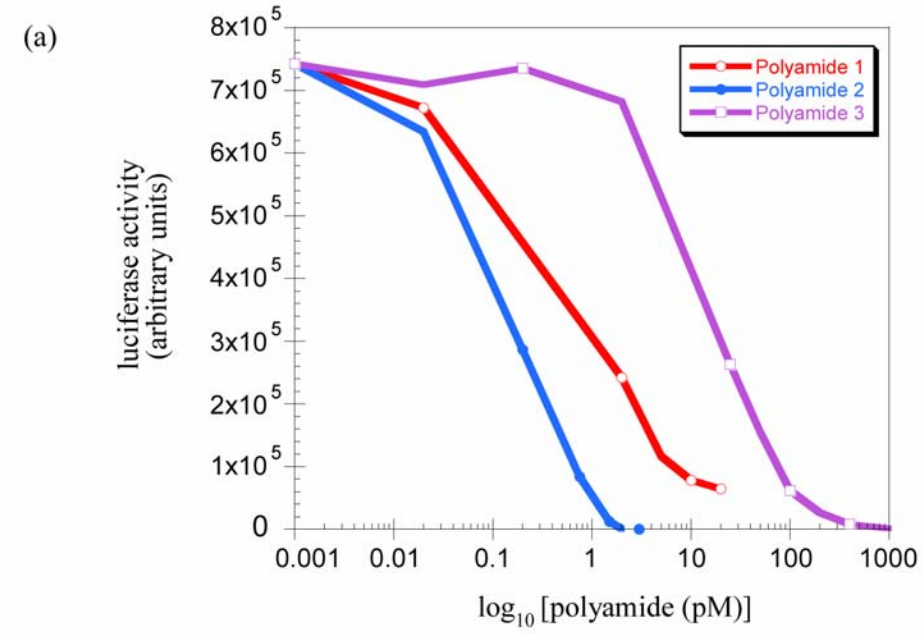


Figure 5.4 DNase I Footprinting of **1**-**3**. (a) PSA promoter PCR fragment used for footprinting, bearing the ARE. (b) Footprinting gels and isotherms for compounds **1**-**3**.



(b)

	K_a (M^{-1})	IC_{50} (pM)
1	1.4 x 10^9	2
2	3.2 x 10^9	0.2
3	9.9 x 10^8	25

5′-A A T T G C A G A A C A G C A A G T G C T A G C T C T-3′
1
3′-T T A A C G T C T T G T C G T T C A C G A T C G A G A-5′

5′-A A T T G C A G A A C A G C A A G T G C T A G C T C T-3′
2
3′-T T A A C G T C T T G T C G T T C A C G A T C G A G A-5′

5′-A A T T G C A G A A C A G C A A G T G C T A G C T C T-3′
3
3′-T T A A C G T C T T G T C G T T C A C G A T C G A G A-5′

ARE

Figure 5.4 *In vitro* transcription assay for the inhibition of AR action by polyamides **1-3**. (a) The dose response curves for polyamides **1-3** are depicted showing luciferase activity as a function of polyamide concentration. Polyamide **2** shows a markedly lower IC_{50} than the other polyamides (0.2 pM, blue line.) Polyamide **1** reaches IC_{50} at 2 pM (red line). Polyamide **3** is much less potent, reaching IC_{50} at 25 pM (violet line). (b) Summary of DNA-binding and AR-inhibition characteristics of polyamides **1-3**. Though there is < 5-fold difference in binding affinity between **2** and **3**, there is a > 100-fold increase in AR inhibition, showing that polyamide position is a critical variable in successful inhibition of transcription.

In attempts to inhibit the AR-ARE interaction *in vivo*, transient transfection assays were performed in LNCaP and PC3 prostate cancer cell lines using an ARE-containing luciferase reporter. PC3 cells do not contain endogenous AR, so one may probe the activity of AR with precision by adding controlled amounts of AR. However, this also means that one cannot study the inhibition of an endogenous PSA gene via the polyamide binding of ARE. LNCaP cells contain AR and the PSA gene is actively transcribed in the cell. The advantage LNCaP cells provide is the ability to study the effect of polyamides inhibiting an endogenous gene through ARE binding by performing Western blots or RT-PCR experiments to analyze gene products (either protein or mRNA, respectively.) Unfortunately, all attempts to use polyamides **1**-**3** to inhibit gene expression *in vivo* failed to produce any positive results.

Speculating that the negative results were due to the inability of polyamides to localize to genomic DNA, either by exclusion from the cell or sequestration in cytoplasmic vesicles, polyamide conjugates were synthesized to promote uptake into live LNCaP and PC3 nuclei. The first such compounds synthesized were conjugates of polyamides with DHT (**4**-**9**, Figure 5.5). The synthesis of a representative polyamide-DHT conjugate is shown in Figure 5.6. These were designed to take advantage both of the recognition of steroid hormones by cell membrane transport proteins, to bring the polyamides in to the cell, and the DHT-recognition and nuclear transport properties of AR, to transmit the polyamides into the cell nucleus (Figure 5.7).[10-12] Compounds **4**-**7** were based on polyamides **1** and **2**, which showed the strongest inhibition *in vitro*. They incorporate the DHT moiety, linked by a short polyethylene glycol (PEG) domain, in two different positions on the polyamide, the tail and the *N*-methyl position of an internal

pyrrole residue. Compounds **8** and **9** are based on a double pairing-rule mismatch polyamide of both **1** and **2**, incorporating the DHT moiety as **4-7**.

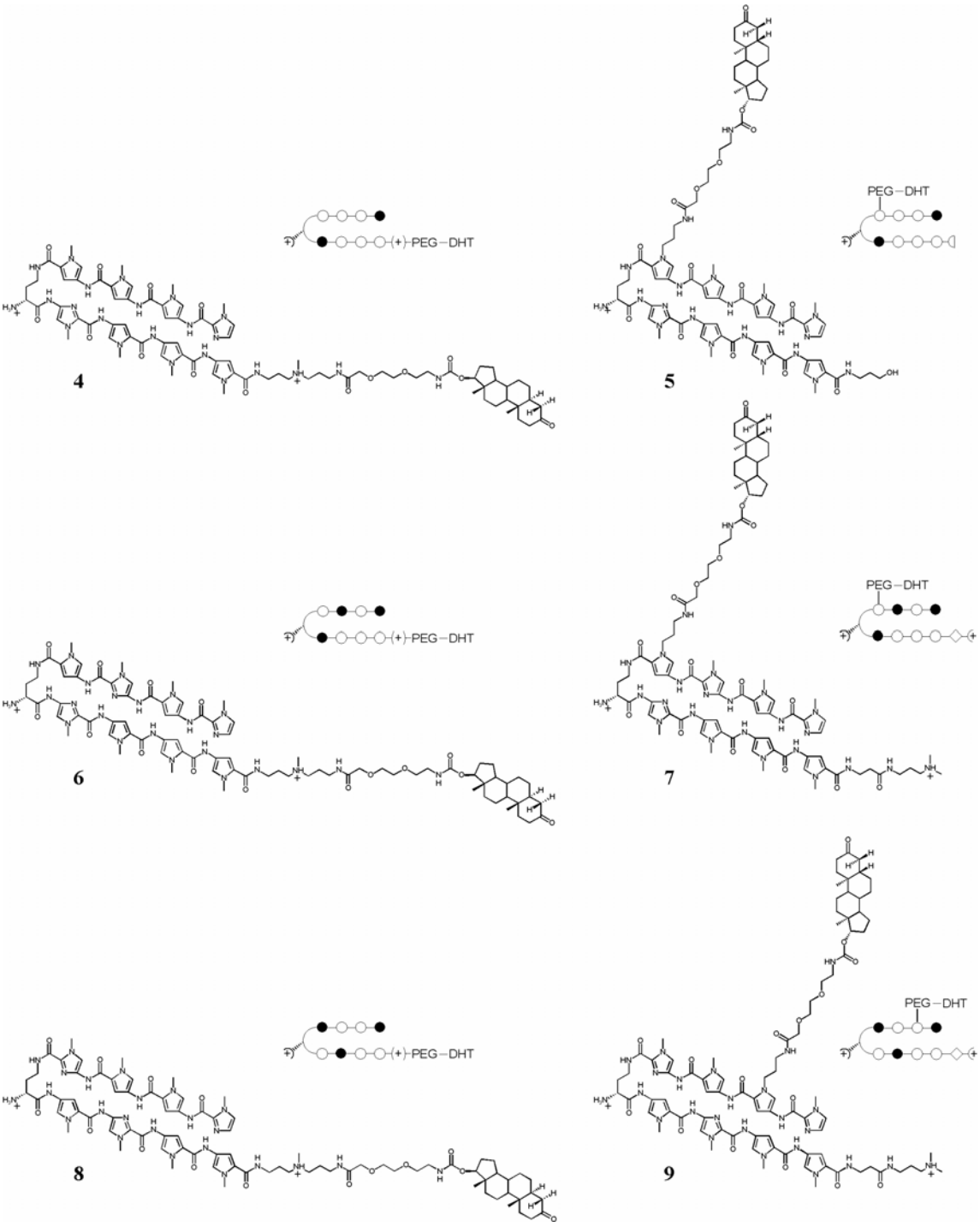


Figure 5.5 Polyamide-DHT conjugates for AR inhibition. Chemical structures and ball-and-stick models of compounds **4-9**.

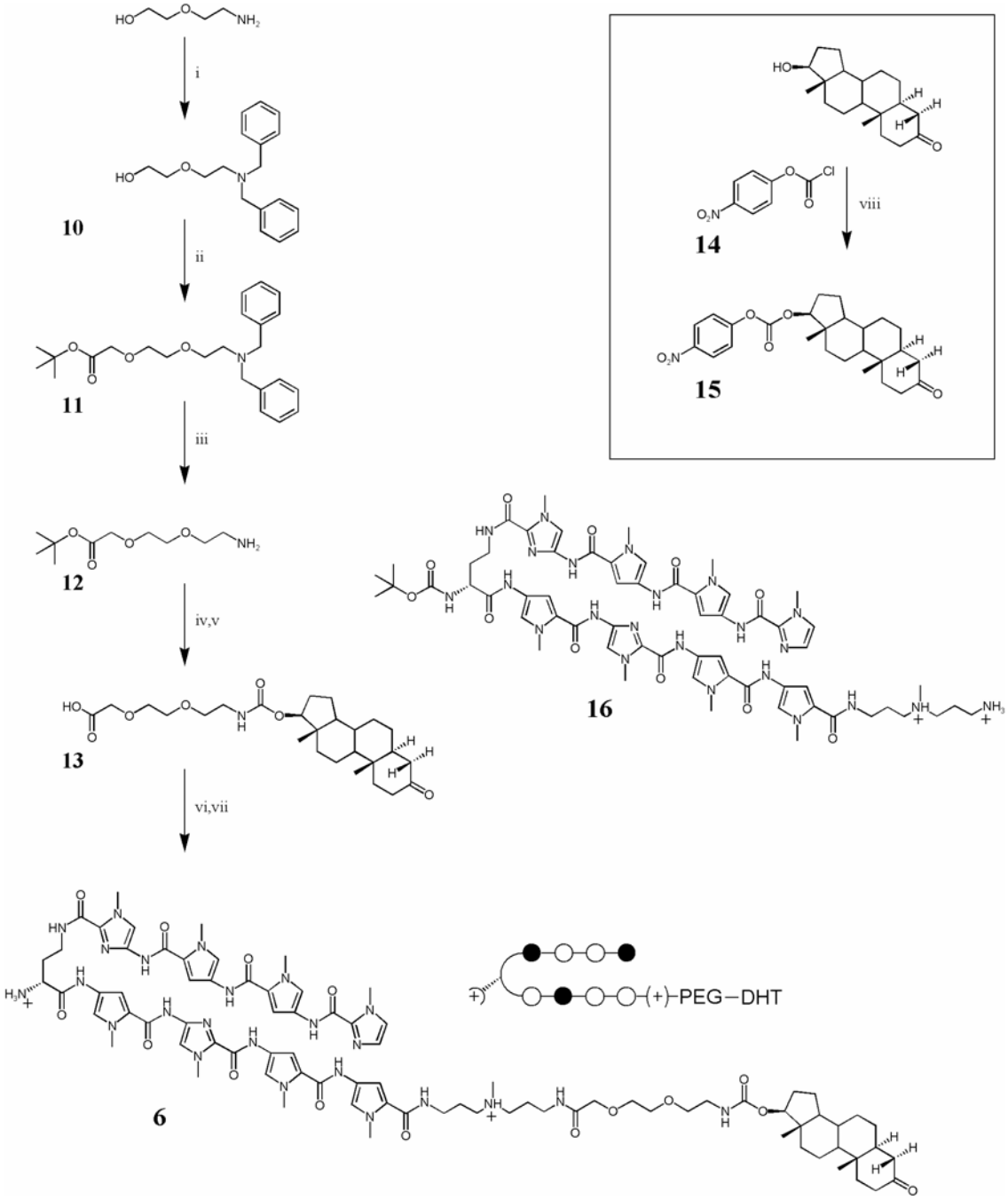


Figure 5.6 Synthesis of polyamide-DHT conjugates. (i) benzyl bromide, K_2CO_3, H_2O. (ii) *t*-butylbromoacetate, NaH, DMF, 0ºC. (iii) Pd/C (10% wt/wt), H_2, MeOH, AcOH. (iv) **15**, NEt_3, CH_2Cl_2, 60ºC. (v) 50% (v/v) TFA/CH_2Cl_2, 1hr at room temperature. (vi) **16**, DCC, HOBt, DIEA, DMF. (vii) 20% TFA/CH_2Cl_2, 1hr at room temperature. (viii) *p*-nitrophenylchloroformate (**14**), NEt_3, CH_2Cl_2.

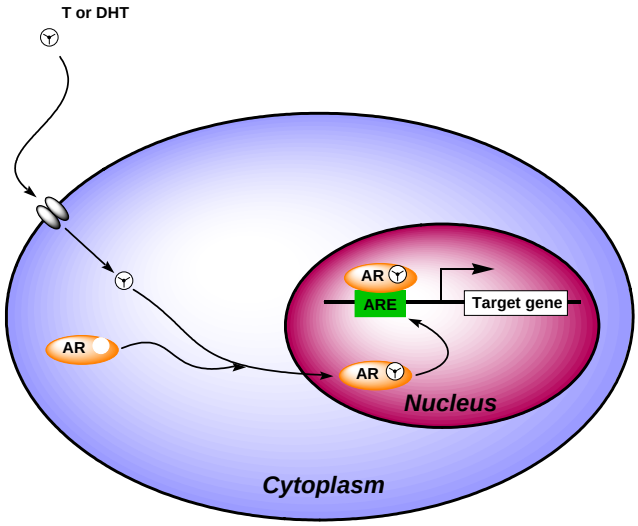


Figure 5.7 Mechanism of androgen-initiated AR-mediated transcription activation. Androgen, here T or DHT, passes through the cytoplasmic membrane either through diffusion or active transport. In the cytoplasm it binds AR, which translocates to the nucleus. Upon recognition of the ARE site, androgen-AR complex (along with other chaperone proteins) binds ARE stimulating transcription of the target gene.

In order to test the effect on DNA-binding of conjugating the DHT-linker moiety to a polyamide, conjugates **4** and **5** were footprinted on the same PSA-ARE PCR fragment as **1**-**3** (Figure 5.8a). Overall the addition of the DHT-linker moiety reduces the affinity of the polyamide for its match site ~10-fold with reference to the parent polyamide **1**. Attachment off of the tail also erodes the ability of the compound to differentiate between DNA sites, lowering its specificity. Attachment off of the *N*-methyl position of an internal pyrrole does not have the same negative effect on DNA specificity.

Compounds **4**-**9** were used in a transient transfection assay in PC3 cells, which lack endogenous AR. The cells were treated with AR and 1μM DHT-polyamides. The effect on transcription of a luciferase gene under control of ARE sites was measured by luciferase activity (Figure 5.8b). Conjugate **4** was the most effective, showing ~45% reduction in luciferase activity. Compound **5**, also based on polyamide **1**, was less

effective, showing 25% reduction. The only other conjugate to show a significant effect was **6**, based on polyamide **2**, which showed ~20% reduction in luciferase. The higher level of inhibition of **4** and **5** over **6** and **7** is somewhat surprising, since the *in vitro* results suggested compounds based on polyamide **2** should be the most efficient inhibitors. This may be a result of higher nuclear transport of **4** and **5** over **6** and **7**.

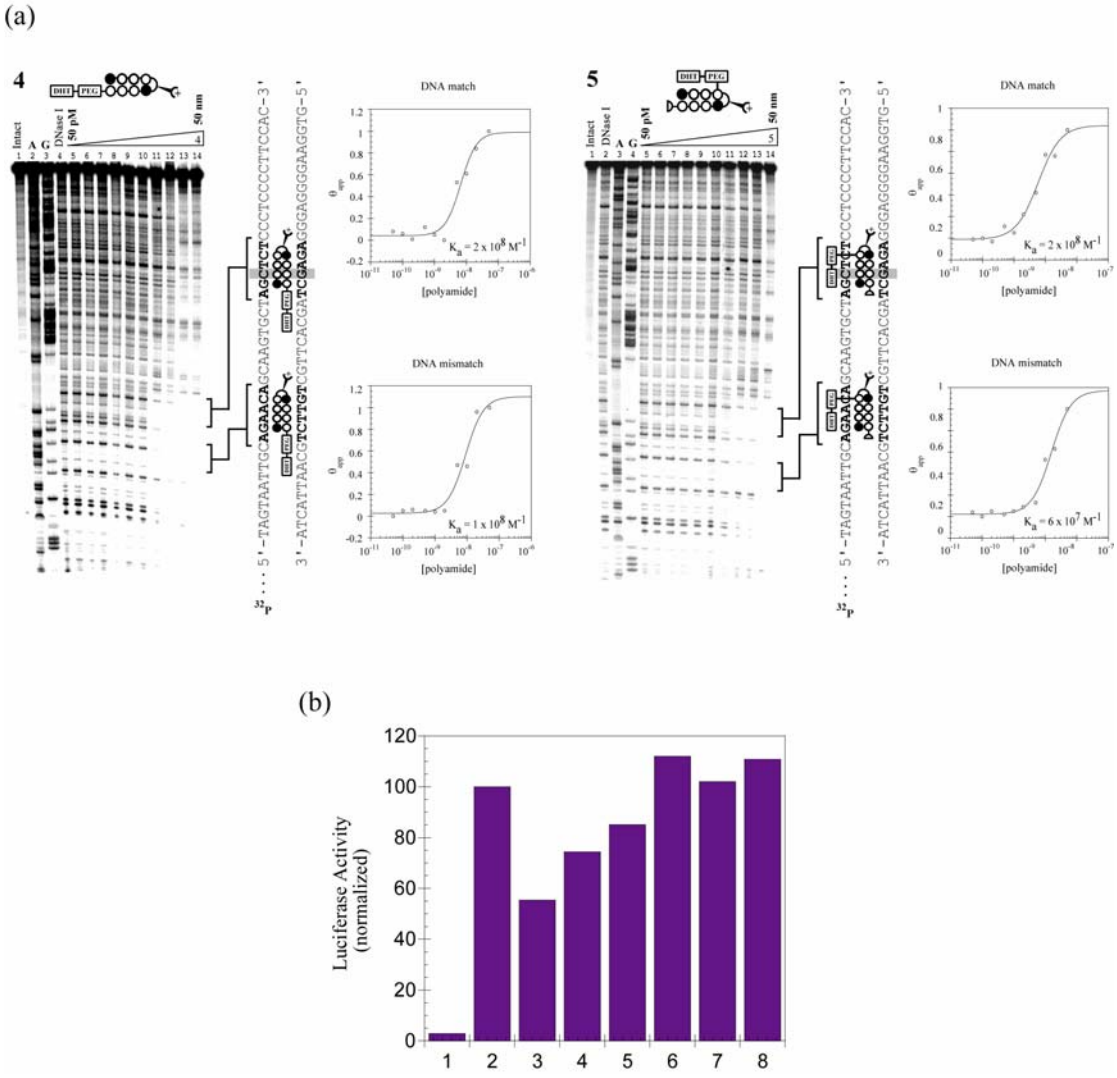


Figure 5.8 DNA-binding and transcription inhibition characteristics of DHT conjugates. (a) DNase I footprinting gels and isotherms for **4** and **5**. (b) Inhibition of luciferase transcription was greatest for compound **4**, at ~45% reduction (activity normalized to +AR –PA sample. Key: column 1: –AR, -PA; column 2: +AR, -PA; column 3: +AR, 1 μM **4**; column 4: +AR, 1 μM **5**; column 5: +AR, 1 μM **6**; column 6: +AR, 1 μM **7**; column 7: +AR, 1 μM **8**; column 8: +AR, 1 μM **9**.

After it was determined that polyamide-fluorescein conjugates tended to possess favorable uptake qualities (see Chapter 4), fluorescein derivatives **17**-**19** were prepared (Figure 5.9a). These compounds were designed based on polyamides **1** and **2**, as well as a double pairing-rule mismatch with reference to **1** and **2**. After verification that they stained the nuclei of PC3 and LNCaP cells, the conjugates were used in the same transient transfection assay that was used on compounds **4**-**9** (Figure 5.9b). Compound **17** seemed to show some inhibition at 0.1 μM. However, compound **18** did not show any significant inhibition. As expected, mismatch compound **19** also produced no inhibitory effect. Since it was shown that these three compounds all enter cell nuclei, it is surprising that the compounds were less efficient than even the DHT conjugates at inhibiting transcription. It may be that the DNA-binding affinities of **17** and **18** are low enough that higher concentrations (> 1μM) are needed. Also, a combination of treatment with both **17** and **18** might provide greater inhibition than either one alone could, in a synergistic fashion.

Outlook and Future Directions

The success in creating DNA-binding polyamides with the ability to cross the cytoplasmic membrane and localize to genomic DNA provides the basis for undertaking many *in vivo* transcription regulation projects. In the case of AR-ARE inhibition, the *in vivo* experiments performed to date have all been on model systems transiently transfected into cells. However, the ultimate goal of transfection experiments with DNA-binding polyamides is the inhibition of endogenous genes upon polyamide treatment. To this end it would be interesting to treat cells that actively transcribe the PSA gene under

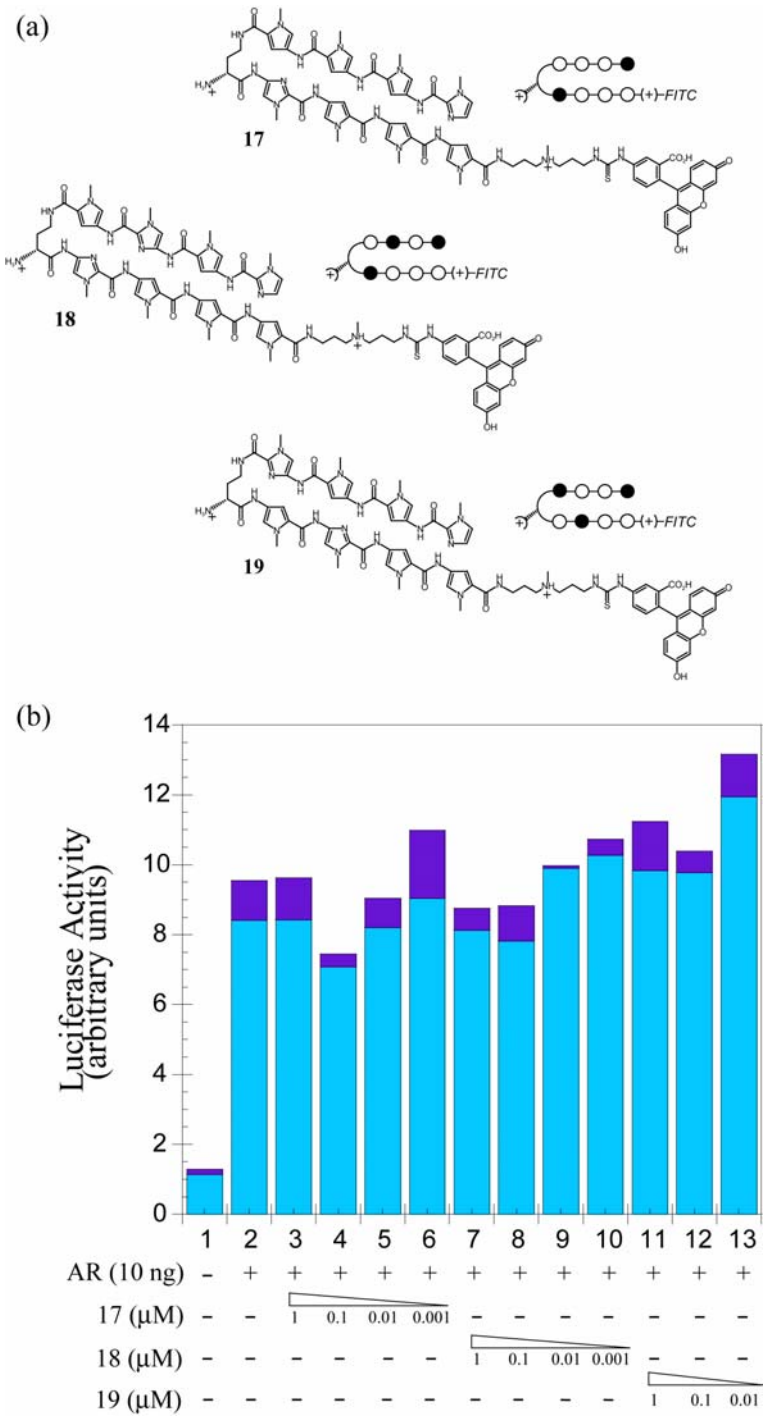


Figure 5.6 Polyamide-FITC conjugates for AR-ARE inhibition. (a) Chemical structures and ball-and-stick models for **17**-**19**. (b) Transient transfection luciferase inhibition assay. Compound **17** seems to have some small effect on luciferase transcription, though strangely not at the highest concentration (1 μM, column 3, compare to column 2). Compounds **18** and **19** show no significant effect. The light blue portion of the bars is luciferase activity, in arbitrary units. The dark blue portion is ½ of the standard deviation after three replicate experiments.

control of AR, such as LNCaP cells, and assay for changes in the level of gene products, either protein (Western blot) or mRNA (RT-PCR). The development of new aromatic heterocycles that can take part in DNA recognition will permit the targeting of DNA sequences in a more general manner. By combining new polyamides based on the lead compounds **1** and **2**, with uptake vectors such as FITC, and by assaying their effects on endogenous PSA production by RT-PCR, it seems likely that a path to successful gene inhibition in living cells by polyamides will be illuminated.

Experimental

Chemicals

Polyamides were synthesized by solid phase methods on Boc-β-ala-PAM resin (Peptides International, Louisville, KY)[13] or on Kaiser oxime resin (Nova Biochem, Laufelfingen, Switzerland).[14] All fluorescent dye reagents were from Molecular Probes. Chemicals not otherwise specified were from Sigma-Aldrich.

Synthesis of Polyamide-DHT Conjugates

2-(*N*,*N*-Dibenzyl-2-aminoethoxy)ethanol (10)[15]

2-(2-aminoethoxy)ethanol (15 mL, 150 mmol), potassium carbonate (41.5 g, 300 mmol), and 45 mL water were added to a 250 mL round bottom flask equipped with a magnetic stirrer and a dropping funnel. Benzyl bromide (35.7 mL, 300 mmol) was added dropwise over 1 hr. The solution was then stirred overnight at room temperature. Et_2O (250 mL) was added to the mixture, separated, and the aqueous layer was washed with a further 50 mL Et_2O. The combined organic portions were washed with 300 mL brine and

150 mL water, then dried with $MgSO_4$ and concentrated to yield a clear oil. This was purified on a short column of silica with 5:1 hexanes:EtOAc→100% EtOAc to yield a clear oil (22 g, 52% yield). ^{1}H-NMR ($CDCl_3$) δ 7.20-7.38 (m, 10 H), 3.65 (s, 4 H), 3.58 (t, 2 H), 3.47 (t, 2 H), 2.69 (t, 2 H), 2.39 (s, 1 H).

_N_,_N_-Dibenzyl-8-amino-3,6-dioxaoctanoic acid _t_-butyl ester (11)[16]

NaH (1.33 g of 95% wt/wt, 50 mmol) was added to a flask under argon with 100 mL DMF. Compound **10** (10 g, 35 mmol) was added and the mixture stirred at room temperature for 2 hrs. *t*-Butylbromoacetate was added to a dry flask under Ar at 0ºC. Solution from above was added via cannula and stirred at 0ºC for 30 min. The resulting solution was then allowed to warm to room temperature and stir overnight. 10 mL saturated NH_4Cl was added and the solution was concentrated to remove DMF. The remainder was partitioned between 100 mL sat. NH_4Cl and 150 mL EtOAc. The organic layer was washed with 100 mL NH_4Cl, 100 mL $NaHCO_3$ (1M), and 100 mL brine, then dried with $MgSO_4$ and concentrated to yield an amber oil. Purification on silica with 5:1 hexanes:EtOAc provided 9.5 g (68% yield) clear oil. ^{1}H-NMR ($CDCl_3$) δ 7.19-7.38 (m, 10 H), 3.94 (s, 1 H), 3.40-3.60 (m, 10 H), 2.69 (t, 2 H), 1.38 (s, 9 H).

8-amino-3,6-dioxaoctanoic acid _t_-butyl ester (12)

Compound **11** (6.4 g, 16 mmol) was dissolved in a solution of 750 μL AcOH in 150 mL MeOH. Palladium on charcoal (10% wt/wt) was added and the mixture was sealed in a Parr bomb and pressurized to ~7 atm H_2. The solution was stirred at room temperature for 18 hrs, then filtered through celite to remove Pd/C. The filtrate was concentrated to yield an amber oil, which was dissolved in 30 mL CH_2Cl_2 and washed

with 1M $NaHCO_3$ (20 mL) and concentrated. ^{1}H-NMR ($CDCl_3$) δ 4.24 (s, 2 H), 3.94 (t, 2 H), 3.71 (s, 4 H), 3.07 (t, 2 H), 1.45 (s, 9 H).

5α-dihydrotestosterone-*N*-(8-amino-3,6-dioxaoctanoic acid) carbamate (13)

12 (0.33 g, 1.5 mmol) and **15** (0.46 g, 1 mmol) were dissolved in 1 mL CH_3CN and DIEA (175 mL, 1 mmol) and stirred at 60°C for 24 hrs. The mixture was concentrated, suspended in 10 mL CH_2Cl_2 and washed with 2 x 20 mL 10% (wt/wt) citric acid and 20 mL water. The organic layer was dried over $MgSO_4$ and concentrated to yield a brown oil. This was treated with TFA (50% in CH_2Cl_2, 10 mL total volume, 2 hrs at room temperature) and concentrated to yield a brown oil. The crude product was run through a short column of silica with 2:1 hexanes:EtOAc, then eluted with MeOH and concentrated to an oily brown solid. Recrystallization from EtOH/water, resuspension in DCM, and concentration provided **13** as a light tan foam (0.154 g, 29% yield).

5α-dihydrotestosterone-4-nitrophenyl carbonate (15)

5α-dihydrotestosterone (2.0 g, 6.9 mmol) was dissolved in 40 mL CH_2Cl_2. *p*-nitrophenyl chloroformate (**14**) (1.7 g, 8.3 mmol) and triethylamine (1.15 mL, 8.3 mmol) were then added and the resulting solution was stirred at 35ºC for 12 hrs. The solution was concentrated, resuspended in 50 mL CH_2Cl_2, washed with 40 mL 1M $NaHCO_3$, 40 mL 10% (wt/wt) citric acid, and 40 mL brine. The organic portion was dried over $MgSO_4$ and concentrated to yield 2.83 g **15** (91% yield) as a fluffy white solid. ^{1}H-NMR ($CDCl_3$) δ 3.99 (s, 2 H), 3.65 (m, 4 H), 3.55 (t, 2 H), 3.33 (t, 2 H), 2.45-0.75 (series of m, 29 H), 1.47 (s, 9 H).

DHT-polyamides 4-9

13 (4.8 mg, 10 μmol) was activated as the –OBt ester with DCC (10 μmol) and HOBt (10 μmol) in DMF (60 μL) for 30 min at room temperatue. This was added to a solution of polyamide (each as **16**, possessing a free amine in the appropriate position and a *t*-Boc-protected (*R*) ^{H_2N}γ–turn, 2 μmol) and DIEA (1 μL, 6 μmol) in 50 μL DMF and the resulting reaction allowed to react for 1 hr at room temperature. Dp (2.5 μL) was added to quench remaining activated acid and the solution was concentrated to a brown oil in a speedvac. This was treated with 80% (v/v) TFA/CH_2Cl_2 for 2 hrs at room temperature, purified by C_{18} reverse-phase preparatory HPLC, and lyophilized to provide **4-9** as white powders. Characterization: **4**, MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1670.9, obsd 1671.0; **5**, MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1643.8, obsd 1643.9; **6**, MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1671.9, obsd 1672.0; **7**, MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1742.9, obsd 1743.1; **8**, MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1671.9, obsd 1672.0; **9**, MALDI-TOF $[M+H]^+$ (monoisotopic mass) calcd 1742.9, obsd 1743.0.

FITC-polyamides 17-19

Polyamides were synthesized on Kaiser oxime resin as previously detailed, each possessing a chiral (*R*) ^{H_2N}γ–turn protected as the *t*-butyl carbamate. Each polyamide (2 μmol) was dissolved in 100 μL DMF and 10 μL DIEA. Fluorescein isothiocyanate (3 μmol) was added and the reactions allowed to proceed in the dark at room temperature for 1 hr. The polyamides were precipitated by addition of Et_2O and treated with 100 μL 20% (v/v) TFA/CH_2Cl_2 for 1 hr at room temperature. The solutions were then

concentrated, suspended in 10 mL 0.1% (v/v) TFA/water, purified by C_{18} reverse-phase preparatory HPLC, and lyophilized to provide **17**-**19** as orange powders.

Gel Mobility Shift Assays

Gel shift assays were performed by Dr. Zhengxin Wang at the M.D. Anderson Cancer Center at the University of Texas Southwestern.

In vitro Transcription Reactions

In vitro transcription reactions were performed by Dr. Zhengxin Wang at the M.D. Anderson Cancer Center at the University of Texas Southwestern.

Transient Transfection Assays

Transient transfection assays were performed by Dr. Zhengxin Wang at the M.D. Anderson Cancer Center at the University of Texas Southwestern. Briefly, the human prostate cancer cell lines PC3 and LNCaP were cultured in a 5% CO_2 atmosphere at 37°C in RPMI medium 1640 supplemented with 10% fetal bovine serum (Irvine Scientific) and 1% penicillin/streptomycin solution (Sigma). The report vector was pGL3-ARE, containing 4 ARE's and the E4 core promoter fused to the luciferase gene. The expression vector is pcDNA-AR containing human AR under control of the CMV promoter. Assays were performed on 0.25 million cells, using 500 μg pcDNA-AR, 500 μg pGL3-ARE, 20 μg pRL-LUC as an internal control, and 10 μL LipofectAmine (Invitrogen). After 48 hrs, the cells were harvested for luciferase assay.

DNase I Footprinting Titrations

A 217 bp 5' ^{32}P-labeled PCR fragment was generated from template plasmid pPSA-ARE in accordance with standard protocols and isolated by nondenaturing gel electrophoresis.[17] All DNase I footprinting reactions were carried out in a volume of 400 μL. A polyamide stock solution or water (for reference lanes) was added to TKMC buffer, with final concentrations of 50 mM Tris-HCl, 50 mM KCl, 50 mM $MgCl_2$, and 25 mM $CaCl_2$, pH 7.0, and 15 kcpm 5'-radiolabeled DNA. The solutions were equilibrated for 12-18 h at 22°C. Cleavage was initiated by the addition of 10 μL of a DNase I stock solution and was allowed to proceed for 7 min at 22°C. The reactions were stopped by adding 50 μL of a solution containing 2.25 M NaCl, 150 mM EDTA, 0.6 mg/mL glycogen, and 30 μM base pair calf thymus DNA and then ethanol precipitated. The cleavage products were resuspended in 100 mM Trisborate-EDTA/80% formamide loading buffer, denatured at 85°C for 10 min, and immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V for 2 h and 15 min. The gels were dried under vacuum at 80 °C and quantitated using storage phosphor technology.

References

176) Chang, C.; Kokontis, J.; Liao, S.T. *Science* **1988**, *240*, 324.

177) Chang, C.; Kokontis, J.; Liao, S.T. *Proc. Natl. Acad. Sci. U.S.A*. **1988**, *85*, 7211.

178) Chang, C.; Saltzman, A.; Yeh, S.; Young, W.; Keller, E.; Lee, H.-J.; Wang, C.; Mizokami, A. *Crit. Rev. Eukaryotic Gene Expression* **1995**, *5*, 97.

179) Evans, R.M. *Science* **1988**, *240*, 889 and references therein.

180) Truss, M.; Beato, M *Endocr. Rev*. **1993**, *14*, 459.

181) Cleutjens, K.B.J.M.; van Eekelen, C.C.E.M.; van der Korput, H.A.G.M.; Brinkmann, A.O.; Trapman, J. *J. Biol. Chem*. **1996**, *271*, 6379.

182) Luisi, B.F.; Xu, W.X.; Otwinowski, Z.; Freedman, L.P.; Yamamoto, K.R.; Sigler, P.B. *Nature* **1991**, *352*, 457.

183) Shaffer, P.L.; Jivan, A.; Dollins, D.E.; Claessens, F.; Gewirth, D.T. *Proc. Natl. Acad. Sci. U.S.A*. **2004**, *101*, 4758.

184) Zhang, Y.L.; Wang, C.; Wilding, G.; Chang, C. *Biochem. Biophys. Res. Comm*. **1993**, *195*, 710.

185) Lee, H.-J.; Chang, C. Cell. Mol. Life Sci. 2003, 60, 1613.

186) Boffa, L.C.; Scarfi, S.; Mariani, M.R.; Damonte, G.; Allfrey, V.G.; Benatti, U.; Morris, P.L. *Cancer Res*. **2000**, *60*, 2258.

187) Sack, J.S.; Kish, .R.; Wang, C.; Attar, R.M.; Kiefer, S.E.; An, Y.; Wu, G.Y.; Scheffler, J.E.; Salvati, M.E.; Krystek, S.R., Jr.; Weinmann, R.; Einspahr, H.M. *Proc. Natl. Acad. Sci. U.S.A*. **2001**, *98*, 4904.

188) Baird, E.E.; Dervan, P.B. *J. Am. Chem. Soc*. **1996**, *118*, 6141.

189) Belitsky, J.M.; Nguyen, D.H.; Wurtz, N.R.; Dervan, P.B. *Bioorg. Med. Chem*. **2002**, *10*, 2767.

190) Koskinen, A.M.P.; Valo, T.; Vihavainen, S.; Hakala, J.M.L. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 573.

191) Wittmann, V.; Takayama, S.; Gong, K.W.; Weitz-Schmidt, G.; Wong, C.-H. *J. Org. Chem*. **1998**, *63*, 5137.

192) Trauger, J.W.; Dervan, P.B. *Methods Enzymol*. **2001**, *340*, 450.