Structural Modifications to DNA-Binding Polyamides for Improved Biological Activity in Cell Culture

Thesis by

Claire Jacobs

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended March 6, 2009)

To my Family

Acknowledgements

I would like to thank my advisor, Professor Peter Dervan, and my fellow Dervanites for creating such a diverse and stimulating place to work. I count myself lucky to have shared a lab with Justin, Ryan, Jim, Christian, and Dave C., both for their scientific prowess and companionship, which made the late nights in lab a lot more fun and a little less frustrating. Katy graciously provided tireless proofreading services, helpful suggestions, and vital material support in the form of baked goods. Puckett knows a lot about almost everything, and I will miss dissecting current events with him. Dan provided invaluable suggestions on a manuscript and a cool head in Dervana. Anne's sage advice during daily coffee runs is gratefully acknowledged. I would also like to thank Mareike, Nick, Fei, Sherry, Julie, Carey, Adam, John, Mike B., Dave, Ben E., Marques, Ray, Eric, Michelle, Ben L., and Dan G. It has been a pleasure working with you all, and I wish the best of luck to the newest members of the Dervan group.

I don't have the room to thank all the people who looked out for me along the way. I do want to acknowledge Mrs. Rabin and Dr. Gordo Stonington for their early encouragement, Professor Steve Sibener for letting me into his lab as an undergraduate, and Professor Ka Yee Lee for being an inspiration. I would like to thank Kat, Russell, and Sherry for their generosity and care. I am grateful to Erika, Marten, Skumfi, and Adri for their friendship and camping companionship. PY, Tabby, Mark, Nadia, Asim; I look forward to having the time to catch up with you all.

I owe an enormous debt of gratitude to JP for his friendship. I am extraordinarily grateful for RJN's support, patience, and skills with a camping stove. Lou, thanks for the early introduction to science and enthusiasm for all things science. Finally, I want to thank my family, especially my mother, for their love, encouragement, and advice.

Abstract

Polyamides are a class of synthetic small molecules that recognize DNA in a sequence-specific fashion through a network of hydrogen bonds formed with bonding partners in the floor of the minor groove. The binding affinity of polyamides is comparable to that of numerous DNA-binding proteins, and polyamides have been shown to displace DNA-binding proteins. As such, they present a powerful opportunity to modulate expression levels of genes vital to human health. The cellular permeability and biological activity of polyamides has presented an impediment in moving from *in vitro* to *in vivo* work that was partially removed by the discovery that fluorescein dyes facilitate cell entry. The work described here details recent advances in modifications to the C-terminal polyamide linker, linkage and tail groups that improve the endogenous inducible gene regulation activity of polyamides in cell culture.

Table of Contents

List of Figures and Tables

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 1

Introduction

Introduction

The information encoding the regulation and expression of the estimated 20,000–25,000 genes in the human genome and their products is contained in the nucleotide sequence of DNA[1,2]. At the simplest level, the information encoded by nucleotide sequences is expressed through a process of transcription of DNA information to an RNA message, followed by translation into the final protein product. States of human disease are frequently the result of gene expression disregulation, and the discovery and development of technologies capable of granting control over the process of gene expression would have wide-ranging applications in human health.

Structure of DNA

DNA is comprised of two complimentary polydeoxyribonucleotide strands paired in an antiparallel fashion through formation of hydrogen-bonding contacts between the nucleotide bases: adenine pairs with thymine, and guanine with cytosine (figure 1.1).[3] Although DNA can adopt a number of conformations, the biologically relevant is B-form DNA, in which the nucleotide pairing produces a double-helical strand with a wide and shallow major groove, and a deep, narrow minor groove between the phosphate backbones. To a rough approximation, the DNA double helix is uniform, but a series of hydrogen-bonding donors and acceptors located on the edges of nucleotide bases are presented in the major and minor grooves that function as "handholds" for DNA-binding proteins and natural products.

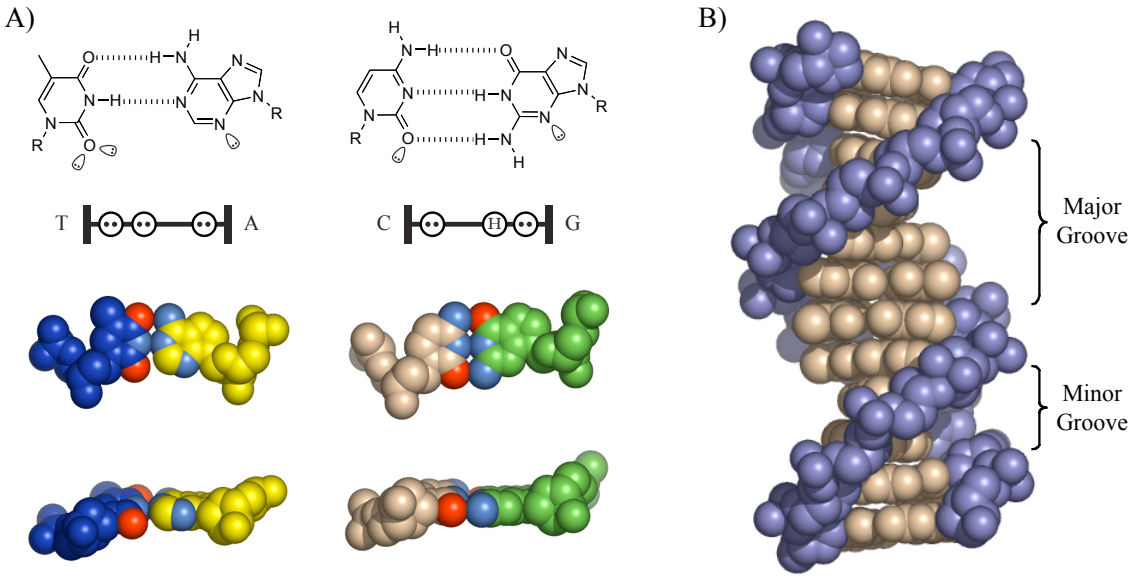


Figure 1.1. Schematic illustration of B-form DNA. A) Schematic of hydrogen-bond formation between a T•A basepair and a C•G basepair. Also shown is a schematic representation of the hydrogen-bonding donor and acceptor pattern presented by the basepair edges into the minor groove. Lone pairs are represented by two dots in an open circle, and hydrogen bonding donors represented by an "H" in an open circle. B) Structure of B-form DNA to illustrate the major and minor grooves of B-form DNA.

DNA-Binding Natural Products

Two natural products that bind DNA with some degree of sequence specificity are Hoechst 33258 and distamycin (figure 1.2).[4-7] Courtesy of their crescent shape, both these relatively simple heterocycle oligomers fit neatly into the minor groove of DNA to form hydrogen bonds with the nucleotide edges. Distamycin can form either a 1:1 complex or a 2:1

complex with DNA: the latter is illustrated in figure 1.2B). Distamycin has a preference for A,T tracts of DNA due to hydrogen-bonding contacts between the amide-bond hydrogen and minor-groove lone pairs; G•C pairs present an energetically unfavorable steric clash between the exocyclic amine of guanine and hydrogens on the pyrrole ring. Replacement of a pyrrole ring with an N-methylimidazole ring, however, enabled G•C recognition via the Nitrogen lone pair, which forms a hydrogen bond with the guanine exocyclic amine.[8]

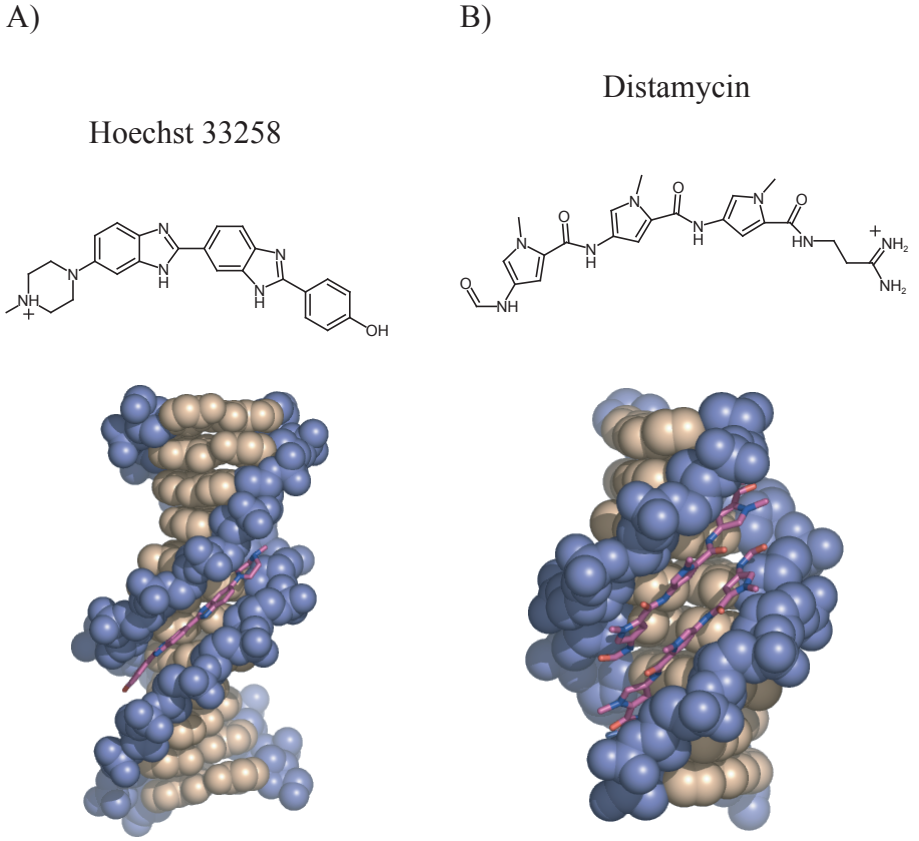


Figure 1.2. Schematic illustration of the chemical structure of Hoechst 33258 and Distamycin, and structures of DNA-Hoechst 33258 and DNA-distamycin complexes.

Polyamide Recognition of DNA

Polyamides are a class of synthetic small molecules that bind DNA in a sequence-specific manner modeled on Distamycin-DNA recognition through hydrogen-bonding interactions with hydrogen-bonding donors and acceptors in the floor of the DNA-minor groove. Polyamides are oligomers of N-methylpyrrole (Py), N-methylimidazole (Im), N-methyl-3-hydroxypyrrole (Hp), and 3-chlorothiophene (Ct) carboxamides linked through amide

bonds.[9-16] DNA binding can occur in a 1:1 or a 2:1 fashion, with polyamide strands generally aligning N → C with the 5' → 3' DNA strand direction. The sickle shape of polyamides closely matches the radius of the DNA minor groove, and allows polyamides to approach the DNA minor groove floor closely enough to engage in these hydrogen-bonding interactions.

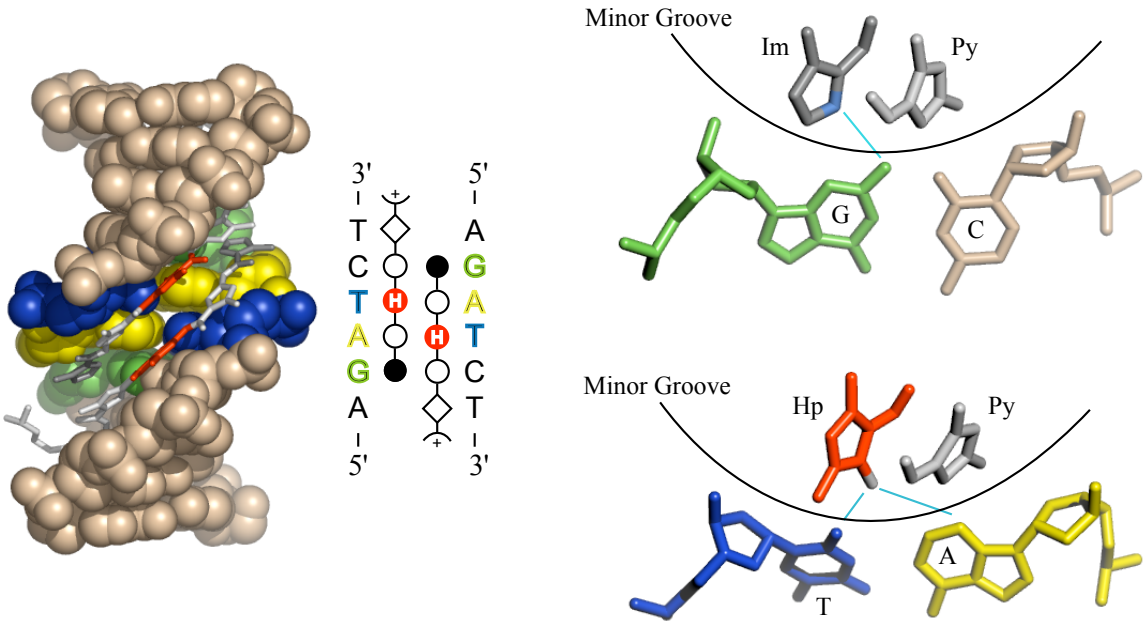


Figure 1.3. Structure of a 2:1 polyamide-DNA complex, and illustration of hydrogen-bonding interactions between the polyamide heterocycle core and the minor groove floor.

Polyamide sequence specificity is programmed through side-by-side pairings of heterocyclic amino acids in the DNA minor groove: Im/Py recognizes G•C over C•G; Py/Py is degenerate for A•T and T•A; Hp/Py and Hz/Py distinguish T•A for A•T; and Ct/Py prefers T•A over A•T in the cap position. These pairing guidelines were revealed through quantitative DNase I footprinting titrations, as well as NMR and crystallographic structural data. These data have also elucidated the peculiarities of polyamide-DNA interactions that give rise to the sequence specificity. A network of hydrogen bonding interactions is created between the polyamide amide bonds and base-pair edges, while hydrogen bonding between the nitrogen lone pair of Im and the exocyclic amine of guanine forms the basis for G•C recognition (figure 1.4).[17-19] Hp/Py recognizes T•A through hydrogen bond formation between the Hp hydroxyl and the thymine-O2 lone pair along with shape-

selective recognition of an asymmetric cleft between thymine-O2 and adenine-C2.[17, 19] The proposed recognition of T•A at the amino terminus by Ct/Py projection of the 3-chloro substituent awaits structural verification.[20]

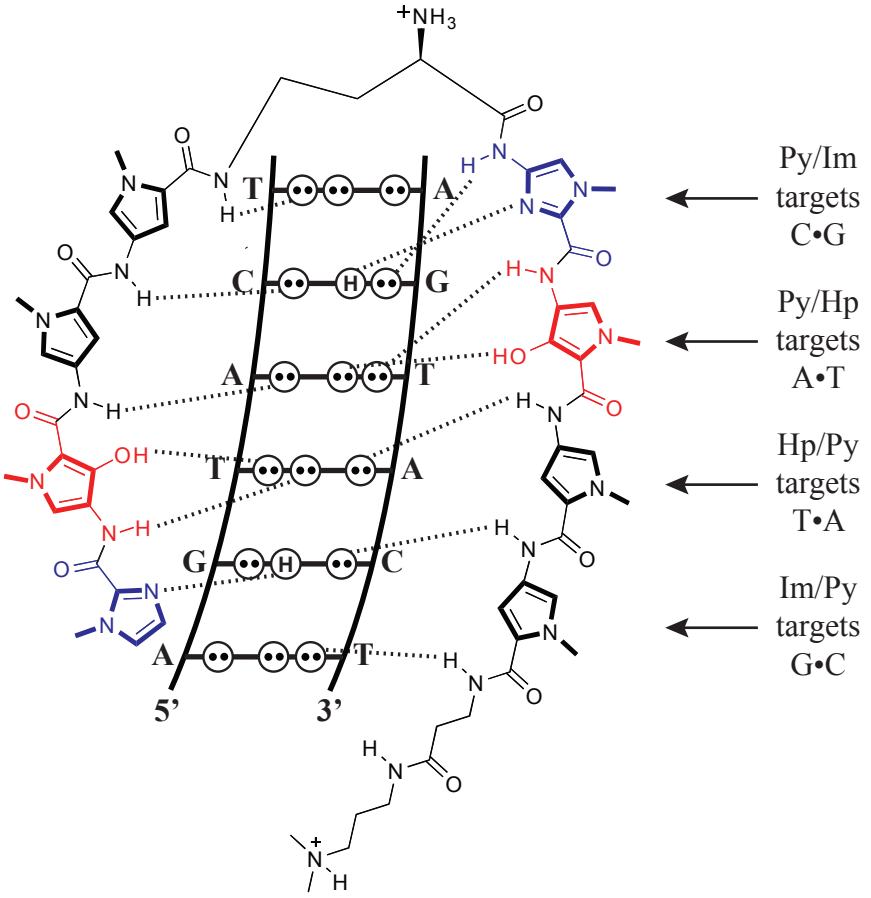


Figure 1.4. Schematic illustration of the system of hydrogen bonding pattern between the DNA nucleotide edges and an eight-ring hairpin polyamide that gives rise to DNA recognition by polyamides.

A number of polyamide architectures have been explored, and the most commonly deployed is the eight-ring hairpin polyamide with a binding site of six basepairs (figures 1.4, 1.5). Linking two antiparallel heterocycle strands through an aliphatic linker group to create a "hairpin" polyamide results in substantial gains in DNA-binding affinity ($K_a(M^{-1})$).[21] This gain in binding affinity is thought to occur as a result of reduction in the entropic cost of polyamide-DNA binding. Linking the two strands additionally enforces side-by-side heterocycle pairing rather than the slipped binding modes available to unlinked heterocycle strands.[21, 22] The most commonly employed turn group has been the γ-amino butyric acid, optionally with a chiral amine moiety at the α carbon.[23] The presence of this amine moiety

imbues the turn residue with a strong preference for T•A and A•T pairs over G•C and C•G pairs.[21]

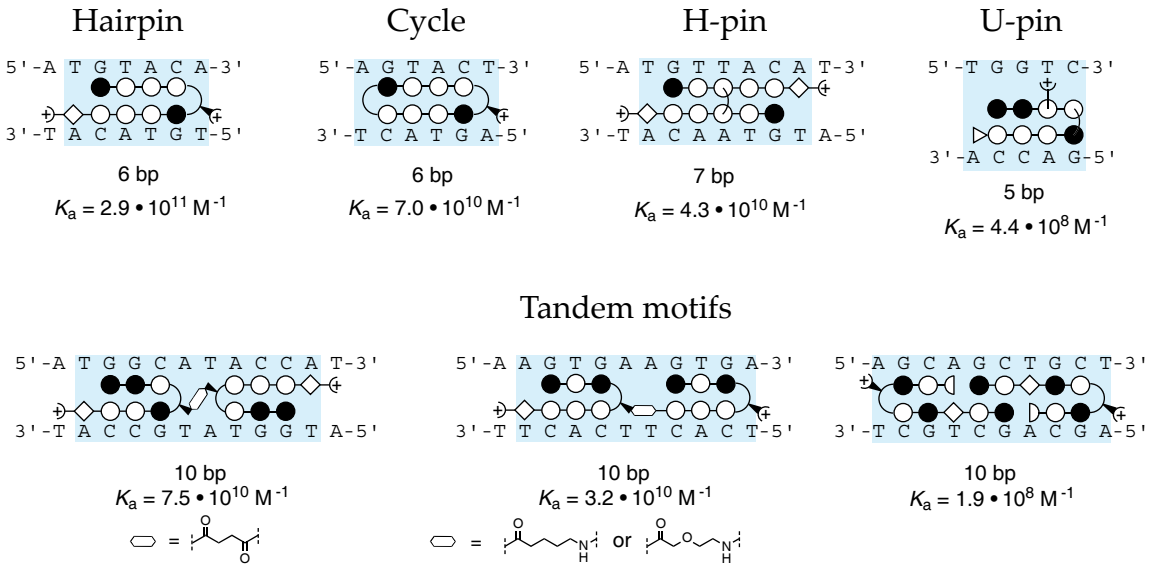


Figure 1.5. Schematic ball-and-stick representations of various polyamide structural motifs. Below each motif is given the binding site size and binding affinities (K_a (M^{-1})).

The modular nature of polyamides makes their synthesis amenable to solid-phase synthesis methods on resins including β-Ala-PAM and Kaiser oxime resin; use of β-Ala-PAM resin results in a polyamide product with a C-terminal β-alanine moiety that is specific for A•T basepairs (figure 1.6).[24, 25] Following resin loading with the initial C-terminal, standard Boc coupling chemistry is used to build the polyamide in the C→N direction. Aminolysis provides the cleaved polyamide, and judicious selection of the cleaving group allows further elaboration of the polyamide with all manner of C-terminal tail groups. Solid-phase synthesis has been successfully exploited to access rapidly a wide range of polyamide heterocycle cores in the relatively small quantities required for exploratory research, but can be adapted in a relatively straightforward fashion to produce larger quantities of polyamide product. Production of large quantities of particular cores is facilitated through synthesis of heterocycle trimers or tetramers, which can be readily introduced into the polyamide core on resin. Recently, the first solution-phase synthesis of a polyamide in our group yielded gram-scale polyamide product.

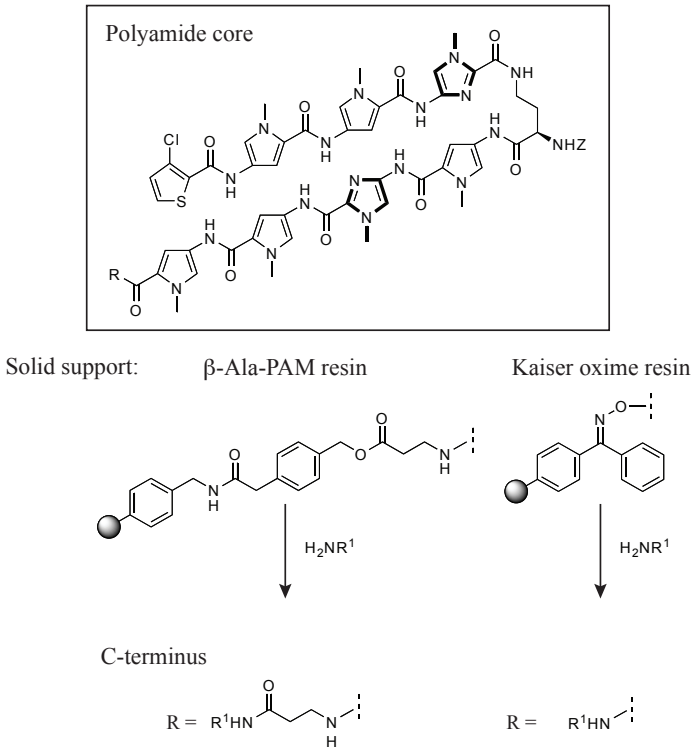


Figure 1.6. Schematic illustration of solid-phase synthesis of polyamides.

Regulation of Gene Expression with Polyamides:

The sequence-specificity of polyamides presents a potentially powerful tool for modulating gene expression in pursuit of disease treatment. Two possible mechansims by which gene regulation could be achieved are illustrated schematically below (figure 1.7). The first involves direct steric blockade of transcription factor binding in the promoter region of a gene, and the second mechanism is based on recruiting transcriptional machinery to the promoter through use of activator domains.

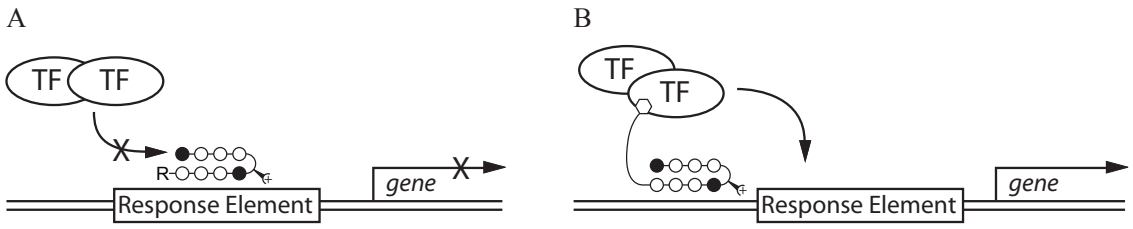


Figure 1.7. Schematic of proposed mechanisms of polyamide gene expression regulation through modulating the DNA-protein interface. A) Inhibition of transcription factor (TF) protein-DNA binding through direct steric blockade at a TF recognition element (RE). B) Polyamide elaboration with protein recruitment domains to increase local concentration of targeted DNA-binding proteins.

The binding affinity of most eight-ring hairpin polyamides is on the order of 10^9-10^{10} M^{-1}, which makes them competitive with DNA-binding proteins such as transcription factors and transcriptional activators (figure 1.8).[26] By targeting polyamides to protein binding sites or adjacent sequences, *in vitro* studies have demonstrated the ability of polyamides to inhibit DNA binding of proteins through direct steric or allosteric effects.[27] Major-groove binding proteins have been shown to co-occupy DNA sequences with minor-groove binding polyamides, but they can be displaced by intercalator-polyamide conjugates.[28-30]

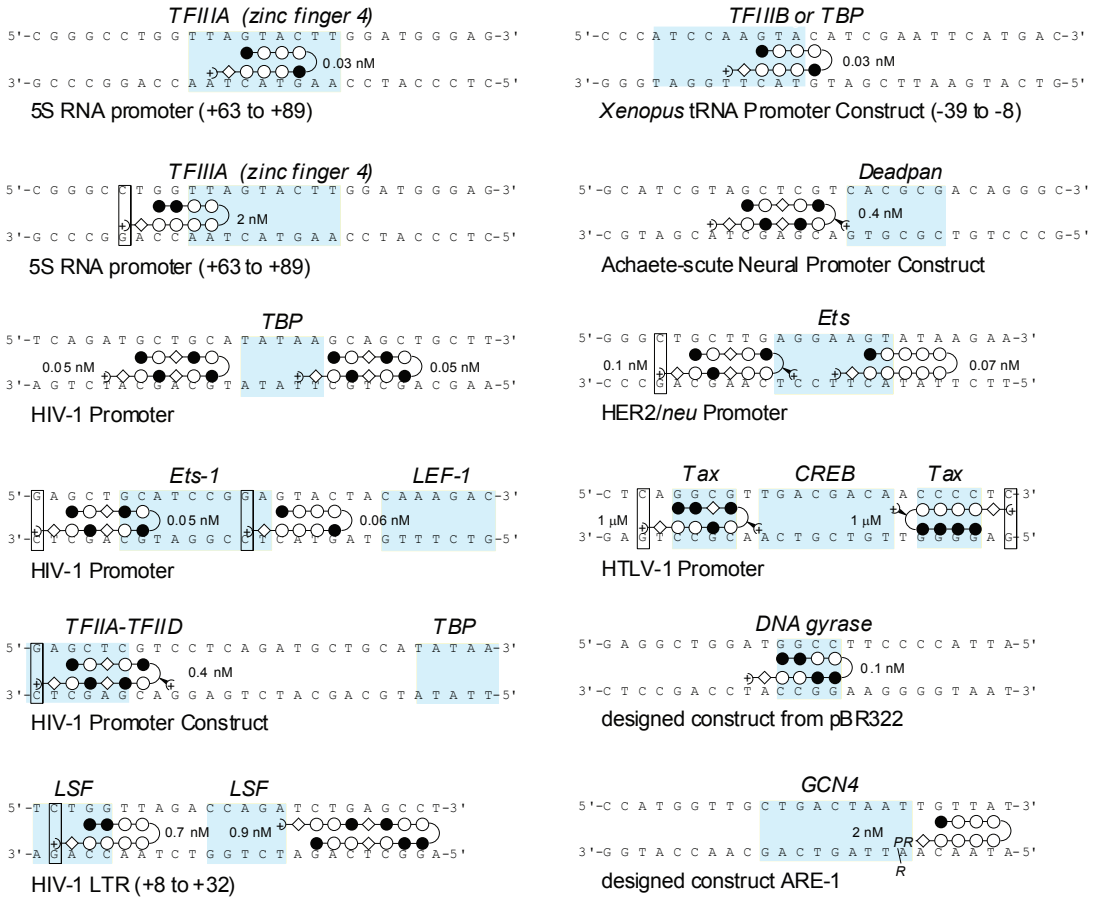


Figure 1.8: Schematic illustration in the ball-and-stick fashion of inhibition of protein binding by polyamides. Indicated by light-blue boxes are the protein binding sites to show overlap between polyamide and protein binding sites. Also given are polyamide dissociation constants (K_a (M^{-1})).

A further point of comparison between polyamides and DNA-binding proteins is their relatively limited recognition sequences. Given the size of the genome, it is clear that a much longer sequence than the six bases recognized by an eight-ring hairpin polyamide would be required to target a unique DNA binding site. The limited size of polyamide

binding sites may result in promiscuous polyamide binding thruought the genome and numerous "off-target" effects. This concern is mitigated by consideration of known DNA-protein interactions. Protein-DNA recognition, e.g. transcription factor-DNA recognition, is the result of a combination of electrostatic interactions with the polyanionic-DNA backbone and Van der Waals forces acting between the protein and the major or minor groove floors.[31-36] In the cellular context, the DNA double helix is not suspended in the nucleus in isolation, but is constantly interacting with proteins and other molecules in the nuclear millieu. These protein binding sites are generally no more than four basepairs, the same size as eight-ring hairpin polyamide binding sites.[37]

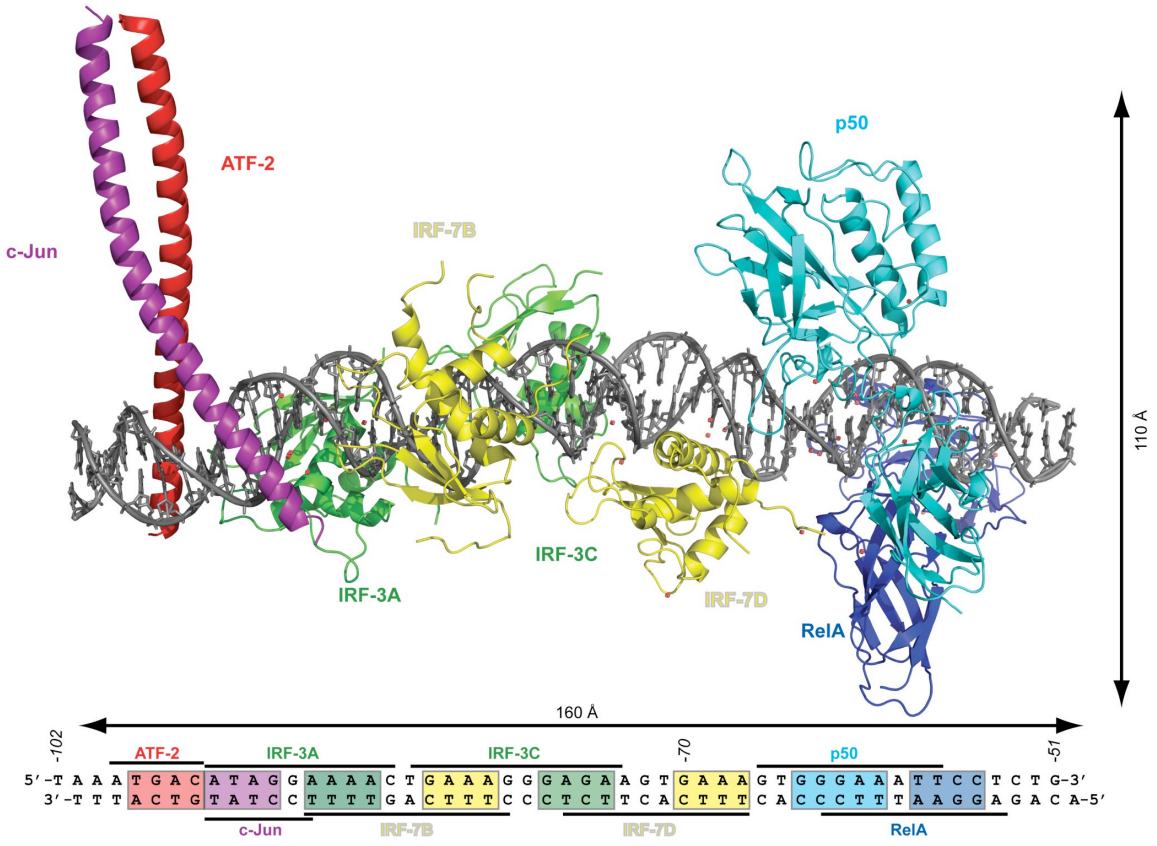


Figure 1.9. Structure of the interferon-b enhanceosome complex to demonstrate the relatively limited size of protein DNA binding sites, and protein-induced perturbations in the DNA double helix. Figure based on Figure 4 in Reference 37, and generated in PyMOL using a composite binding model obtained from the pdb file in the Supplemental Information.

Nuclear Localization of Polyamides:

A central aim is the use of polyamides in biological systems and whole organisms to regulate endogenous genes that are critical mediators of health and disease. To realize this

goal, polyamides must be able to enter the cell and accumulate in the nucleus in sufficient quantities to bind to the appropriate DNA sequence and disrupt the protein-DNA interface. The mechanism by which polyamides enter the cell and translocate to the nucleus is not well understood. Previous work in the Dervan group has shown that polyamide cell uptake is an energy-dependent process, as it is impeded by lowered temperatures and in glucose-lacking cell culture media[38]. Further, internal data have demonstrated that the biological activity of polyamides is enhanced in some cases by co-administration of verapamil, which inhibits function of drug-efflux pathways.[39]

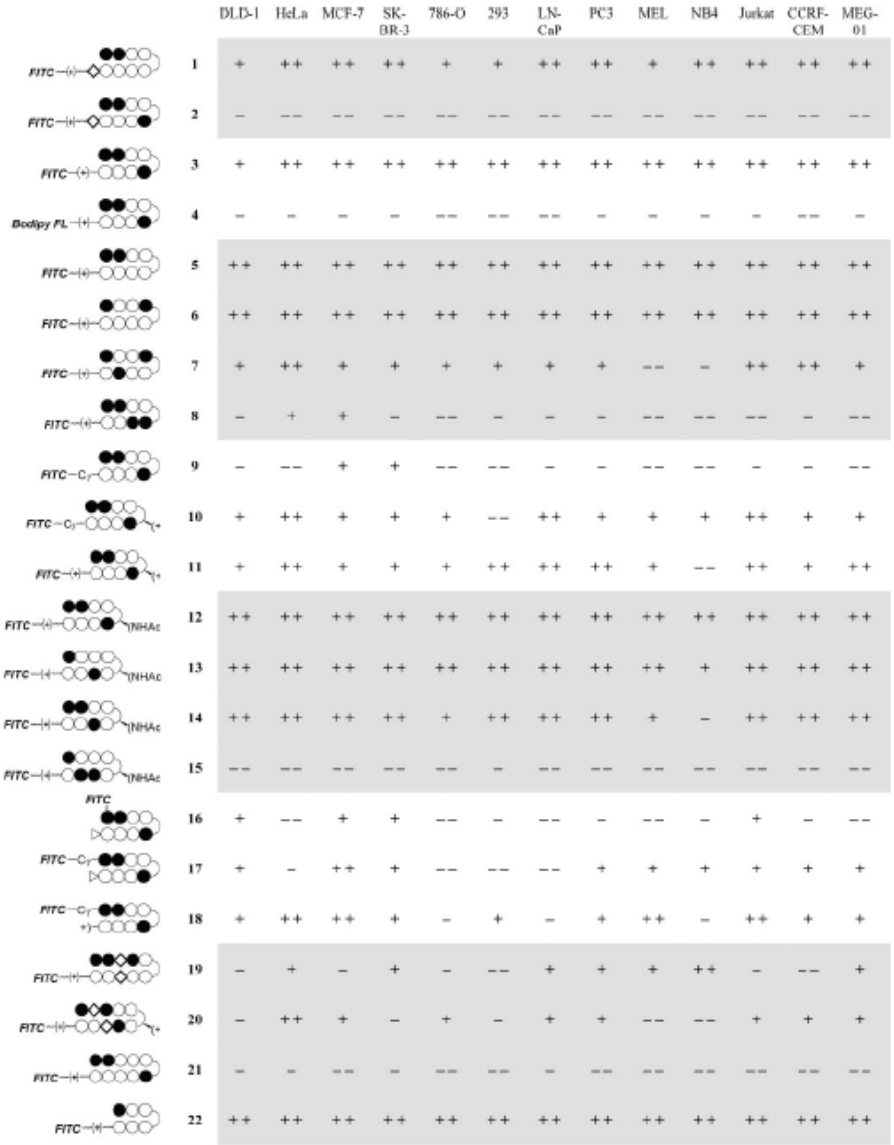


Figure 1.10. Cellular localization of 22 polyamide-fluorophore conjugates in a panel of 13 cultured cell lines. The extent of polyamide nuclear localization is indicated by a +/- nomenclature: ++, nuclear staining exceeds that of the medium: +, nuclear staining comparable to the medium, but still clear: -, very little nuclear staining predominantly cytoplasmic and medium staining; --, no nuclear staining, little cytoplasmic staining.

Extensive attention has been devoted to determining the structural features of polyamides that play a role in nuclear accumulation. Initial investigations were inspired by the work of Bashkin *et al.* demonstrating that a C-terminal fluorescein (FITC) dye enabled demonstrable nuclear localization visualized by confocal microscopy of live, unfixed cells.[39] Following this seminal piece, confocal microscopy techniques were used to examine the nuclear and subcellular localization of polyamides as a function of fluorophore identity, polyamide heterocycle core sequence and composition, and polyamide structure.[38, 40, 41] These investigations revealed that a number of factors influence polyamide cell entry, but did not lead to a set of predictive structural requirements for polyamide nuclear accumulation.

Polyamides in Biological Systems:

Although polyamides had successfully inhibited protein-DNA binding and polyamide–small molecule activator domain conjugates had recruited proteins to DNA *in vitro*, gene regulation by polyamides *in vivo* was less established.[42-48] The most notable example of polyamide gene regulation in cell culture was disruption of HIV-1 replication in human T cells subsequent to dosing with two polyamides targeted to sites adjacent to three transcription-factor binding sites.[49] A critical outgrowth of the extensive confocal microscopy studies of polyamide cell entry and subcellular localization patterns was the revelation that a wide range of heterocycle core-dye combinations are able to cross into the cell.

In collaboration with the group of Dr. Kaelin, a polyamide was designed to bind to the hypoxia-response element in the promoter region of the vascular endothelial growth factor (VEGF) gene.[50] VEGF is among a host of genes regulated by the hypoxia-inducible factor/Aryl-hydrocarbon-receptor nuclear translocator (HIF-1α/ARNT (or HIF-1β)) transcription factor heterodimer.[51] Under normal oxygen tension, the α-subunit of HIF-1α is hydroxylated at two proline residues for subsequent ubiquitination and degradation; under hypoxic conditions, the hydroxylation event does not occur, and instead the HIF-1α/ARNT heterodimer forms and binds to the hypoxia response element (HRE).[52] HIF-1a/ARNT is a master regulator of genes involved in physiological processes including

angiogenesis, glycolysis, cell survival, drug resistance and matrix remodeling.[53-57] As hypoxic conditions are frequently experienced by rapidly growing tissues, such as cancers, HIF-regulated genes and VEGF in particular are of great interest to the medical community for their involvement in cancer progression and survival.[55, 58, 59]

HeLa cells were dosed with the FITC-conjugate of a polyamide designed to bind to the HRE of VEGF (5'-TACGTG-3') and found by confocal microscopy to localize to the nucleus. VEGF mRNA levels in cells dosed with this FITC-polyamide conjugate and later treated with desferoxamine (DFO) to induce hypoxia-regulated gene expression were lower than VEGF mRNA levels, indicating that this FITC-polyamide conjugate was not only localizing in the nucleus, but was able to regulate gene expression levels of a medically relevant endogenous inducible gene. This initial work has since been validated by polyamide-regulated expression of prostate specific antigen (PSA) upon dihydrotestosterone (DHT) induction in LNCaP (prostate cancer) cells.[60]

Summary:

Although it has been shown that a wide array of polyamide core-dye conjugates are able to localize to the nucleus of a range of cells, the widespread use of dye conjugates is problematic due to the high cost and instability of fluorescent dyes. The thrust of the work presented in this thesis deals with efforts to identify structural modifications to polyamides that facilitate their biological activity, are chemically stable, and do not interfere with binding affinities and specificities. As one of the aims of this work was to avoid pricey fluorescent dyes, such as FITC, detecting cell uptake and nuclear localization by confocal microscopy was not an option. The recently published result demonstrating VEGF mRNA regulation in HeLa cells was instead adopted as a biological readout of polyamide biological activity.

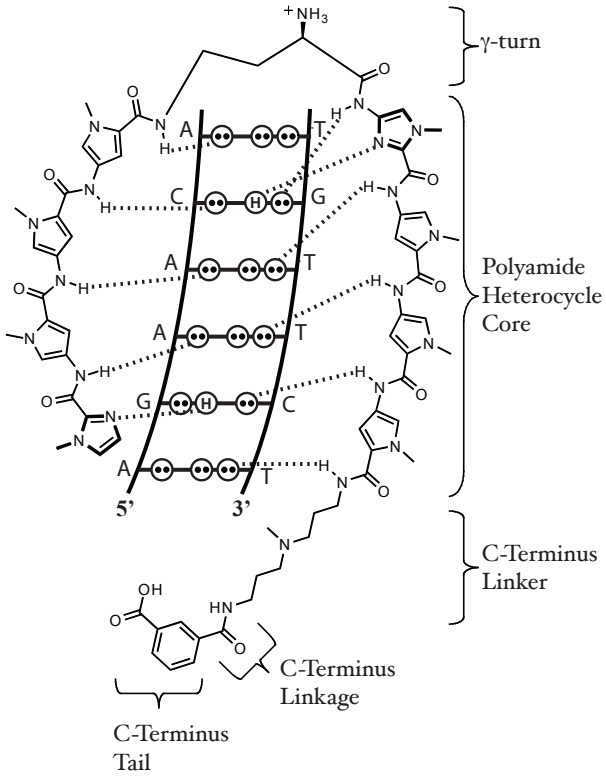


Figure 1.11. Schematic illustration of the hydrogen-bonding pattern mediating DNA recognition by a polyamide. The focus of work to identify structural modifications to improve polyamide biological activity discussed in this thesis, namely the C-Terminus linker, C-terminus linkage and C-terminus tail, are labeled.

Structural modifications to polyamides to effect cell uptake can be made at the C-terminal tail, off the amine of chiral turns, or potentially on heterocycles modified with an internal linker group. As illustrated in figure 1.11, modifications to the C-terminal portion of polyamides can be introduced at a number of positions, including the C-terminal tail

group, the C-terminal linker, or the linkage between tail and linker. Chapter 2 describes initial successes in finding a suitable FITC-replacement, while Chapter 3 deals with further applications of one of these novel tail groups, isophthalic acid (IPA). In addition, microarray data are presented that show that polyamides affect a fairly modest number of genes. Chapter 4 describes the effects of modifications made to the linker and linkage group, as well as attempts to discover the structural basis for the observed cellular uptake of FITC- and IPA-polyamide conjugates through synthesis of a focused library of tail derivatives.

References:

1. Lander, E. S.; Linton, L. M.; *et al.*, Initial sequencing and analysis of the human genome. *Nature* **2001,** 409, (6822), 860-921.

2. Venter, J. C.; Adams, M. D., *et al.*, The sequence of the human genome. *Science* **2001,** 291, (5507), 1304-+.

3. Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L., The Anatomy of A-DNA, B-DNA, and Z-DNA. *Science* **1982,** 216, (4545), 475-485.

4. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M., Structure of the side-by-side binding of distamycin to d(GTATATAC)(2). *Acta Crystallographica Section D-Biological Crystallography* **1999,** 55, 602-609.

5. Kumar, R. A.; Ikemoto, N.; Patel, D. J., Solution structure of the calicheamicin gamma(I)(1)-DNA complex. *Journal of Molecular Biology* **1997,** 265, (2), 187-201.

6. Quintana, J. R.; Lipanov, A. A.; Dickerson, R. E., Low-Temperature Crystallographic Analyses of the Binding of Hoechst-33258 to the Double-Helical DNA Dodecamer C-G-C-G-a-a-T-T-C-G-C-G. *Biochemistry* **1991,** 30, (42), 10294-10306.

7. Arcamone, F.; Nicolell.V; Penco, S.; Orezzi, P.; Pirelli, A., Structure and Synthesis of Distamycin A. *Nature* **1964,** 203, (494), 1064-&.

8. Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B., Antiparallel Side-by-Side Dimeric Motif for Sequence-Specific Recognition in the Minor Groove of DNA by the Designed Peptide 1-Methylimidazole-2-Carboxamide Netropsin. *Proceedings of the National Academy of Sciences of the United States of America* **1992,** 89, (16), 7586-7590.

9. Dervan, P. B., Molecular recognition of DNA by small molecules. *Bioorganic & Medicinal Chemistry* **2001,** 9, (9), 2215-2235.

10. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*

1998, 391, (6666), 468-471.

11. Briehn, C. A.; Weyermann, P.; Dervan, P. B., Alternative heterocycles for DNA recognition: The benzimidazole/imidazole pair. *Chemistry-a European Journal* **2003,** 9, (9), 2110-2122.

12. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Current Opinion in Structural Biology* **2003,** 13, (3), 284-299.

13. Marques, M. A.; Doss, R. M.; Foister, S.; Dervan, P. B., Expanding the repertoire of heterocycle ring pairs for programmable minor groove DNA recognition. *Journal of the American Chemical Society* **2004,** 126, (33), 10339-10349.

14. Dervan, P. B.; Poulin-Kerstien, A. T.; Fechter, E. J.; Edelson, B. S., Regulation of gene expression by synthetic DNA-binding ligands. In *DNA Binders and Related Subjects*, 2005; Vol. 253, pp 1-31.

15. Viger, A.; Dervan, P. B., Exploring the limits of benzimidazole DNA-binding oligomers for the hypoxia inducible factor (HIF) site. *Bioorganic & Medicinal Chemistry* **2006,** 14, (24), 8539-8549.

16. Chenoweth, D. M.; Poposki, J. A.; Marques, M. A.; Dervan, P. B., Programmable oligomers targeting 5' -GGGG-3' in the minor groove of DNA and NF-kappa B binding inhibition. *Bioorganic & Medicinal Chemistry* **2007,** 15, (2), 759-770.

17. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., A structural basis for recognition of A • T and T • A base pairs in the minor groove of B-DNA. *Science* **1998,** 282, (5386), 111-115.

18. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C., Structural basis for G • C recognition in the DNA minor groove. *Nature Structural Biology* **1998,** 5, (2), 104-109.

19. Kielkopf, C. L.; Bremer, R. E.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., Structural effects of DNA sequence on T • A recognition by hydroxypyrrole/pyrrole pairs in the minor groove. *Journal of Molecular Biology* **2000,** 295, (3), 557-567.

20. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., Shape selective recognition of T • A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues. *Bioorg Med Chem* **2003,** 11, (20), 4333-40.

21. Mrksich, M.; Parks, M. E.; Dervan, P. B., Hairpin Peptide Motif – a New Class of Oligopeptides for Sequence-Specific Recognition in the Minor-Groove of Double-Helical DNA. *J Am Chem Soc* **1994,** 116, (18), 7983-7988.

22. deClairac, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E., NMR characterization of hairpin polyamide complexes with the minor groove of DNA. *J Am Chem Soc* **1997,** 119, (34), 7909-7916.

23. Herman, D. M.; Baird, E. E.; Dervan, P. B., Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J Am Chem Soc* **1998,** 120, (7), 1382-1391.

24. Baird, E. E. D., P.B., Solid Phase Synthesis of polyamides containing imidazole and pyrrole amino acids. *J Am Chem Soc* **1996,** 118, 6141-6146.

25. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg Med Chem* **2002,** 10, (8), 2767-74.

26. Trauger, J. W.; Baird, E. E.; Dervan, P. B., Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **1996,** 382, (6591), 559-561.

27. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol* **2003,** 13, (3), 284-99.

28. Oakley, M. G.; Mrksich, M.; Dervan, P. B., Evidence that a Minor Groove-Binding Peptide and a Major Groove-Binding Protein can simultaneously occupy a common site on DNA. *Biochemistry* **1992,** 31, (45), 10969-10975.

29. Fechter, E. J.; Dervan, P. B., Allosteric inhibition of protein-DNA complexes by polyamide-intercalator conjugates. *J Am Chem Soc* **2003,** 125, (28), 8476-8485.

30. Nguyen-Hackley, D. H.; Ramm, E.; Taylor, C. M.; Joung, J. K.; Dervan, P. B.; Pabo, C. O., Allosteric inhibition of zinc-finger binding in the major groove of DNA by

minor-groove binding ligands. *Biochemistry* **2004,** 43, (13), 3880-3890.

31. Nikolov, D. B.; Chen, H.; Halay, E. D.; Hoffmann, A.; Roeder, R. G.; Burley, S. K., Crystal structure of a human TATA box-binding protein/TATA element complex. *Proc Natl Acad Sci U S A* **1996,** 93, (10), 4862-4867.

32. Leonard, G. A.; Hunter, W. N., Crystal and Molecular-Structure of D(CGTAGATCTACG) at 2-Center-Dot-25-Angstrom Resolution. *J Mol Biol* **1993,** 234, (1), 198-208.

33. Kim, Y. C.; Geiger, J. H.; Hahn, S.; Sigler, P. B., Crystal-Structure of a Yeast Tbp Tata-Box Complex. *Nature* **1993,** 365, (6446), 512-520.

34. Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C., The Gcn4 Basic Region Leucine Zipper Binds DNA as a Dimer of Uninterrupted Alpha-Helices – Crystal-Structure of the Protein-DNA Complex. *Cell* **1992,** 71, (7), 1223-1237.

35. Pabo, C. O.; Sauer, R. T., Transcription Factors – Structural Families and Principles of DNA Recognition. *Ann Rev of Biochem* **1992,** 61, 1053-1095.

36. Pavletich, N. P.; Pabo, C. O., Zinc Finger DNA Recognition – Crystal-Structure of a Zif268-DNA Complex at 2.1-A. *Science* **1991,** 252, (5007), 809-817.

37. Panne, D.; Maniatis, T.; Harrison, S. C., An atomic model of the interferon-beta enhanceosome. *Cell* **2007,** 129, (6), 1111-23.

38. 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, (21), 12063-8.

39. 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., Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorg Med Chem Lett* **2003,** 13, (9), 1565-70.

40. Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B., Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. *Bioorg Med*

Chem **2002,** 10, (10), 3313-8.

41. 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. *Nucleic Acids Res* **2004,** 32, (9), 2802-18.

42. Chiang, S. Y.; Burli, R. W.; Benz, C. C.; Gawron, L.; Scott, G. K.; Dervan, P. B.; Beerman, T. A., Targeting the Ets binding site of the HER2/neu promoter with pyrrole-imidazole polyamides. *J Biol Chem* **2000,** 275, (32), 24246-24254.

43. Yang, F.; Belitsky, J. M.; Villanueva, R. A.; Dervan, P. B.; Roth, M. J., Inhibition of Moloney murine leukemia virus integration using polyamides targeting the long-terminal repeat sequences. *Biochem* **2003,** 42, (20), 6249-6258.

44. Simon, H.; Kittler, L.; Baird, E.; Dervan, P.; Zimmer, C., Selective inhibition of DNA gyrase in vitro by a GC-specific eight-ring hairpin polyamide at nanomolar concentration. *Febs Letters* **2000,** 471, (2-3), 173-176.

45. Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., Activation of gene expression by small molecule transcription factors. *Proc Natl Acad Sci U S A* **2000,** 97, (8), 3930-3935.

46. Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., Towards a minimal motif for artificial transcriptional activators. *Chemistry & Biology* **2001,** 8, (6), 583-592.

47. 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, (44), 13067-13071.

48. Kwonj, Y.; Arndt, H. D.; Qian, M.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M., Small molecule transcription factor mimic. *J Am Chem Soc* **2004,** 126, (49), 15940-15941.

49. Dickinson, L. A.; Gulizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B., Inhibition of RNA polymerase II transcription in human

cells by synthetic DNA-binding ligands. *Proc Natl Acad Sci U S A* **1998,** 95, (22), 12890-12895.

50. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G., Jr.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A* **2004,** 101, (48), 16768-73.

51. Semenza, C. L., Hypoxia-inducible factor 1: Control of oxygen homeostasis in health and disease. *Pediatric Research* **2001,** 49, (5), 614-617.

52. Semenza, G. L., Hydroxylation of HIF-1: Oxygen sensing at the molecular level. *Physiology (Bethesda)* **2004,** 19, 176-82.

53. Kaelin, W. G., Jr., How oxygen makes its presence felt. *Genes Dev* **2002,** 16, (12), 1441-5.

54. Ivan, M.; Kondo, K.; Yang, H. F.; Kim, W.; Valiando, J.; Ohh, M.; Salic, A.; Asara, J. M.; Lane, W. S.; Kaelin, W. G., HIF alpha targeted for VHL-mediated destruction by proline hydroxylation: Implications for O-2 sensing. *Science* **2001,** 292, (5516), 464-468.

55. Semenza, G. L., Targeting HIF-1 for cancer therapy. *Nat Rev Cancer* **2003,** 3, (10), 721-32.

56. Semenza, G. L., HIF-1 and human disease: One highly involved factor. *Genes Dev* **2000,** 14, (16), 1983-91.

57. Melillo, G., Inhibiting hypoxia-inducible factor 1 for cancer therapy. *Mol Cancer Res* **2006,** 4, (9), 601-5.

58. Melillo, G., Targeting hypoxia cell signaling for cancer therapy. *Cancer Metastasis Rev* **2007,** 26, (2), 341-52.

59. Yancopoulos, G. D.; Davis, S.; Gale, N. W.; Rudge, J. S.; Wiegand, S. J.; Holash, J., Vascular-specific growth factors and blood vessel formation. *Nature* **2000,** 407, (6801), 242-8.

60. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci U S A*

2007, 104, (25), 10418-23.

Chapter 2

Improved Nuclear Localization of DNA-Binding Polyamides

2.1 Abstract

Regulation of endogenous genes by DNA-binding polyamides requires effective nuclear localization. Previous work employing confocal microscopy to study uptake of fluorophore-labeled polyamides has demonstrated the difficulty of predicting *a priori* the nuclear uptake of a given polyamide. The data suggest that dye identity influences uptake sufficiently such that a dye conjugate cannot be used as a proxy for unlabeled analogues. Polyamides capable of nuclear localization unaided by fluorescent dyes are desirable due to size and other limitations of fluorophores. Recently, a polyamide-fluorescein conjugate targeted to the hypoxia response element (HRE) was found to inhibit VEGF expression in cultured HeLa cells. The current study uses inhibition of VEGF expression as a biological readout for effective nuclear localization of HRE-targeted polyamides. We synthesized a focused library of nonfluorescent, HRE-targeted polyamides in which the carboxy-terminus "tail" has been systematically varied. Members of this library bind the HRE with affinities comparable or superior to that of the fluorescein-labeled analogue. While most library members demonstrate modest or no biological activity, two nonfluorescent polyamides are reported with activity rivaling that of the previously reported fluorescein-labeled polyamide. We also show evidence that promoter occupancy by HIF-1, the transcription factor that binds the HRE, is inhibited by HRE-targeted polyamides.

2.2 Introduction

Polyamides containing *N*-methylimidazole (Im) and *N*-methylpyrrole (Py) comprise a class of programmable DNA-binding ligands capable of binding to a broad repertoire of DNA sequences with affinities and specificities comparable to those of natural DNA-binding proteins (1, 2). Sequence specificity is programmed by side-by-side pairings of the heterocyclic amino acids in the minor groove of DNA: Im/Py distinguishes G•C from C•G; Py/Py binds both A•T and T•A; and 3-chlorothiophene/*N*-methylpyrrole (Ct/Py) prefers T•A over A•T at the amino terminus position (3–5). The use of polyamides to modulate the expression of selected genes through interaction with transcriptional machinery could have applications in biology and human medicine (1, 2). Regulation of endogenous genes by DNA-binding small molecules requires cellular uptake and nuclear localization (6–9), chromatin accessibility (10, 11), and site-specific interactions with gene promoters sufficient to interfere with specific transcription factor–DNA interfaces (12–15).

The use of confocal microscopy to visualize subcellular localization of fluorophore-labeled molecules is a convenient method to study uptake and trafficking of polyamides in living cells (6–9). Nuclear localization of more than one hundred hairpin polyamide-fluorophore conjugates in several human cell lines has been examined using this method (8, 9). Py/Im content, number and location of positive and negative charges, presence of a β-alanine residue at the carboxy terminus, choice of fluorophore, linker composition, and attachment point have been shown to influence nuclear localization of polyamide-fluorophore conjugates (8, 9). Although a significant number of eight-ring hairpin polyamides with fluorescein at the C-terminus were shown to have good nuclear uptake in ten cell lines, this complexity makes it difficult to predict *a priori* the subcellular distribution of a particular polyamide-fluorophore conjugate. Efficient nuclear localization of a fluorophore conjugate is not necessarily predictive for non-fluorescent analogues. Screening polyamides for a biological effect dependent on nuclear localization is a viable approach for assessment of nuclear uptake and does not require incorporation of an optical tag. As polyamides are

studied in model systems of greater biological complexity, molecules that can access the nuclei of cells without the use of fluorescent dyes are desirable.

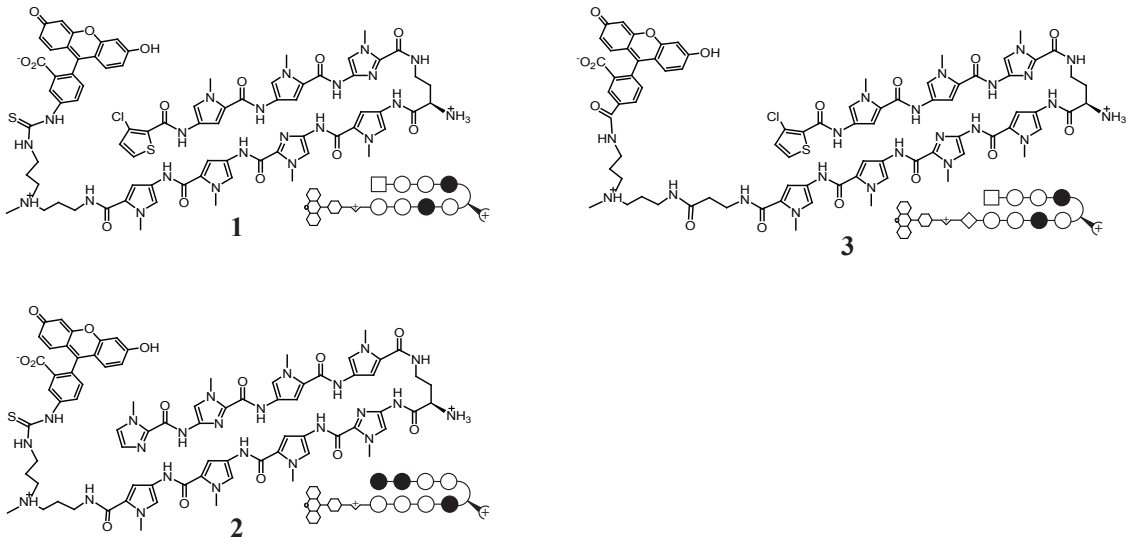


Figure 2.1. Structures of polyamide-fluorescein conjugates **1**–**3**. Polyamides **1** and **3** target the sequence 5'-WTWCGW-3' , while polyamide **2** targets the sequence 5'-WGGWCW-3'. Polyamide **3** differs from **1** and **2** in the linker region between the polyamide core sequence and the fluorescein moiety. Open circles designate *N*-methylpyrrole, closed circles designate *N*-methylimidazole, squares designate 3-chlorothiophene, and diamonds designate β-alanine.

Polyamide-fluorescein conjugate **1** (Figure 2.1) was designed to bind the sequence 5'-ATACGT-3' within the hypoxia response element (HRE) in the vascular endothelial growth factor (VEGF) enhancer. It was found to bind its target site with an affinity of 6.3×10^9 M^{-1}, disrupt hypoxia-inducible factor 1 (HIF-1) binding to the HRE *in vitro*, enter the nuclei of HeLa cells, and inhibit hypoxia-induced VEGF mRNA and protein production in cultured HeLa cells (Figure 2.2) (16). A control polyamide **2** designed to bind the DNA sequence 5'-WGGWCW-3' (where W denotes an A or T) also exhibited nuclear localization, but had a modest effect on VEGF expression. Independently, the DNA binding antibiotic echinomycin, which binds 5'-ACGT-3', was found to inhibit VEGF expression in U251 cells (17, 18). The finding that two molecules of different structure bind the same HRE sequence 5'-TACGT-3' and exhibit similar biological effects on hypoxia-driven gene expression validates the concept of using small molecules to interfere with protein-DNA interfaces.

The current study utilizes the inhibition of hypoxia-induced VEGF expression as a biological readout for nuclear localization of polyamides targeted to the HRE. As a control, we show that a polyamide-fluorescein conjugate **3** that targets the HRE but exhibits poor nuclear localization in HeLa cells due to a non-optimal linker (8) has a limited effect on hypoxia-induced VEGF expression. We have synthesized a focused library of polyamides with an identical core sequence that targets the HRE while varying carboxy terminus (tail) moieties. The carboxy terminus was chosen as the site of chemical variation in our library, as subtle modifications to this area were previously found to influence nuclear localization of polyamide-fluorophore conjugates (8, 9). The polyamides generated in this study retain high binding affinity and specificity to the target site. Synthesized as a control group was a library of "mismatch" polyamides with identical carboxy terminus modifications but a different core targeted to the DNA sequence 5'-WGGWCW-3', which is not found at the VEGF HRE. The majority of polyamides from both libraries failed to inhibit hypoxia induced VEGF expression. We identified, however, two non-fluorescent polyamides with biological activity and binding affinities and specificities comparable to those of polyamide-fluorescein conjugate **1**.

2.3 Materials and Methods

Synthesis of Polyamides

Polyamides were synthesized by solid phase methods on Kaiser oxime resin (Nova Biochem) except polyamide **3,** which was synthesized on Boc-β-alanine-PAM-resin (Peptides International) (19, 20). Fluorescein labeled polyamides **1**–**3** were prepared as previously described (16). Polyamides **4**–**8** and **16**–**20** were cleaved from resin using the appropriate amine and purified by reverse-phase HPLC. For the synthesis of **9**–**15** and **21**–**27**, the polyamide was first cleaved with 3,3'-diamino-*N*-methyl-dipropylamine and purified by reverse-phase HPLC. The designated acid was activated with PyBOP (Nova Biochem) and coupled to the polyamide in dimethylformamide and *N*,*N*-diisopropylethylamine to

yield the final product after reverse-phase HPLC purification. The purity and identity of all polyamides were verified by analytical HPLC and MALDI-TOF MS.

Confocal Microscopy

Microscopy of live, unfixed cells was performed as previously described (8, 9). Briefly, cells were plated on glass-bottom culture dishes for 24 hours prior to overnight incubation with 2 μM polyamide. Imaging was performed on a Zeiss LSM 5 Pascal inverted laser scanning microscope.

Determination of DNA-Binding Affinities and Sequence Specificities

Quantitative DNase I footprint titration experiments were used to determine polyamide binding affinities (K_a) to the 5'-ATACGT-3' sequence within the HRE of the VEGF promoter. These experiments were performed using the 5' ^{32}P-labeled 197 basepair PCR amplification product of the plasmid pGL2-VEGF-*Luc* isolated by nondenaturing gel electrophoresis (16, 21). Quantitative DNase I footprint titration experiments were conducted as previously reported (21).

Measurement of Hypoxia–Induced Relative VEGF mRNA

Cells were plated in 24-well plates at a density of 15–20 × 10^3 cells per well (30–40 × 10^3 cells/mL). Polyamides were added to adhered cells in solutions of cell media at the appropriate concentration and allowed to incubate with the cells for 48 hours. Then, hypoxic induction of gene expression was chemically induced by adding deferoxamine (DFO) to 300 μM for an additional 16 hours (22, 23). Cells were harvested, RNA isolated, cDNA synthesized, and quantitative real-time RT-PCR conducted as previously described (16). Briefly, RNA was isolated using the RNeasy kit (Qiagen), cDNA synthesized using Powerscript (BD Biosciences), and quantitative real-time RT-PCR performed using SYBR Green PCR Master Mix (Applied Biosystems) on an ABI 7300 instrument. VEGF expression was measured relative to β-glucuronidase (GUSB) as an endogenous control. Statistical analyses of results were determined using three independent biological replicates. The primer sequences used for VEGF and GUSB are available upon request. HeLa cells were

purchased from ATCC, and U251 cells were received as a generous gift from Dr. Giovanni Melillo of the National Cancer Institute.

Chromatin Immunoprecipitation

Cells were plated on 15 cm diameter culture dishes and left to attach overnight. Polyamides were incubated with the cells for 48 hours, and the cells were incubated for an additional 16 hours after addition of DFO to a final concentration of 300 μM. Cells were then treated with 1% formaldehyde for ten minutes. Chromatin was isolated and sheared. HIF-1α antibodies (Novus Biologicals) were used to immunoprecipitate HIF-1 bound DNA fragments. After crosslink reversal, PCR reactions using primers flanking the HRE of VEGF were used to assess enrichment of bound fragments as compared to mock-precipitated (no antibody) controls. PCR reactions were monitored either using SYBR Green PCR Master Mix (Applied Biosystems) on an ABI 7300 instrument, or directly visualized using gel electrophoresis.

2.4 Results

Confocal Microscopy of Polyamide-Fluorophore Conjugates 1–3

Cellular uptake of polyamide-fluorescein conjugates **1**–**3** (Figure 2.1) was examined by confocal microscopy in live, unfixed HeLa cells (Figure 2.2B). These data demonstrate that polyamide-FITC conjugates **1** and **2**, targeted to 5'-WTWCGW-3' and 5'-WGGWCW-3', respectively, both localize in the nucleus of HeLa cells. In contrast, polyamide conjugate **3**, also targeted to 5'-WTWCGW-3' but with a β-alanine residue at the carboxy terminus, is largely extracellular.

Suppression of Hypoxia-Inducible Transcription in Cell Culture by Polyamide-Fluorophore Conjugates 1–3

Induction of VEGF mRNA by the hypoxia mimetic deferoxamine (DFO) in the presence of polyamides **1**–**3** in HeLa cells was measured by quantitative real-time RT-PCR (Figure 2.2C). Polyamide **1** inhibits expression of DFO-induced VEGF mRNA production by

approximately 50%. Polyamides **2** and **3** show diminished biological activity compared to **1**. Together, the uptake and gene expression data for **1**–**3** suggest that a polyamide targeted to the HRE of VEGF must have efficient nuclear uptake in order to inhibit hypoxia–induced expression of VEGF. Chromatin immunoprecipitation experiments in cultured HeLa cells were performed to assess whether changes in VEGF promoter occupancy by HIF-1α were consistent with displacement of protein by polyamide **1**. Chromatin immunoprecipitation assays with anti-HIF1α or mock antibody treatment are consistent with inhibition of HIF-1 binding to the HRE in the presence of polyamide 1, and polyamide **2** to a modest degree (Figure 2.2D). In the absence of either polyamide **1** or **2**, HIF-1 occupies the VEGF HRE following DFO-induction. HIF-1 occupancy at the VEGF HRE is reduced in cells treated with **1** prior to DFO-induction. Polyamide **2** has a more limited effect.

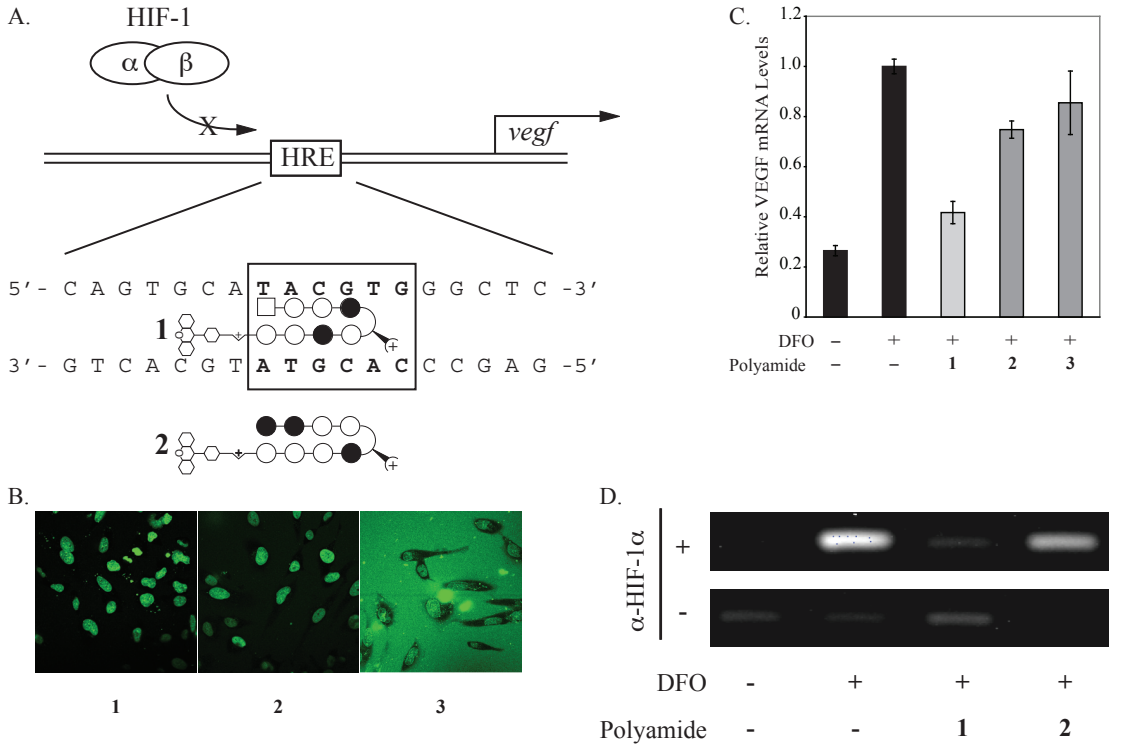


Figure 2.2. (A) Schematic illustration of polyamide **1** binding to the *vegf* HRE. (B) Uptake of polyamides **1**-**3** in HeLa cells. Polyamides **1** (left) and **2** (center) localize in the nucleus while polyamide **3** (right) localizes in the nucleus to a limited extent but is largely extracellular. (C) Induction of VEGF mRNA by the hypoxia mimetic deferoxamine (DFO) in the presence of polyamides **1**-**3** measured by quantitative real-time PCR. Polyamide **1** inhibits expression of DFO-induced VEGF expression. Polyamides **2** and **3** have a more modest effect. (D) Chromatin immunoprecipitation assays using HeLa cells with anti-HIF1α or mock antibody treatment are not inconsistent with inhibition of HIF-1 binding to the *vegf* HRE in the presence of polyamide **1**, and, to a more modest degree, **2**. Concentrations of polyamides are 1 μM.

Non-fluorescent HRE-Targeted Polyamides

We set out to synthesize a series of non-fluorescent HRE-targeted polyamides (**4**–**15**) and their mismatch analogues (**16**–**27**; Figure 2.3). Polyamides **4**–**15** can be considered in three groups: **4**–**8** were designed with minimized tail motifs; **9**–**12** were designed to probe the functional groups present in fluorescein dyes potentially responsible for the favorable nuclear uptake profile of **1**; and **13**–**15** were biotin conjugates.

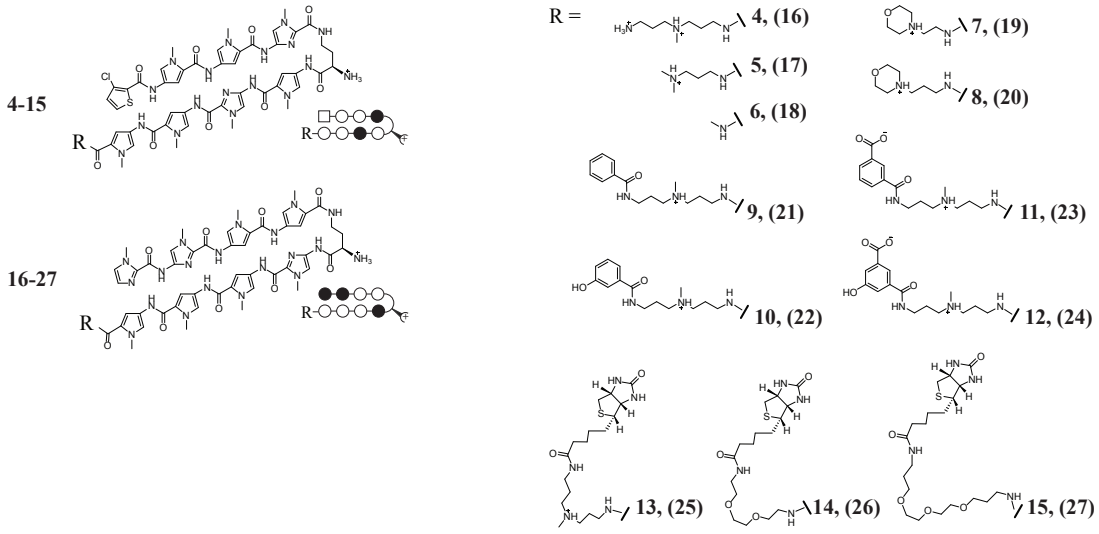


Figure 2.3. Structures of polyamides **4**–**27**. Polyamides **4**-**15** target 5'-WTWCGW-3' DNA sequences; **16**–**27** target 5'-WGGWCW-3'.

Polyamides **4** and **5** were designed to have tail moieties resembling the linker region of polyamide-fluorescein conjugate **1**. The dimethylaminopropylamine tail present in **5** is often placed at the carboxy terminus of polyamides used in biophysical studies (1, 2, 24). Polyamide **6** was designed to have a tail group of minimal size. Polyamides **7** and **8** have 2-morpholinoethanamine and 3-morpholinopropan-1-amine tail groups, respectively. A four–ring polyamide with a morpholino tail was previously shown to have antibacterial activity in a mouse peritonitis model (25). The next group, polyamides **9**–**12**, were designed to reduce molecular weight while retaining functional groups present in fluorescein and maintaining the linker moiety found in polyamide **1**. Polyamides **9**–**12** are derived from benzoic acid; **11** and **12** each contain a carboxylic acid *meta*- to the attachment site, and

10 and **12** each have a hydroxyl group *meta*- to the attachment site. Polyamides **13**–**15** are polyamide-biotin conjugates. The high affinity of the biotin moiety to streptavidin may be useful for biochemical pull-down experiments. Polyamide **13** retains the linker found in **1**–**3** while **14** and **15** contain 2- and 3-oxygen PEG linkers, respectively. Polyamides **16**–**27** were synthesized as mismatch control polyamides for **4**–**15**, respectively.

Binding Affinities and Specificities

Polyamides **4**–**15**, designed to bind the 5'-ATACGT-3' HRE binding site, share the same hairpin polyamide core as polyamides **1** and **3** but have different carboxy terminus tail groups. The DNA binding affinities of **4**–**15** for the HRE of VEGF were measured by quantitative DNase I footprint titrations using a 5' ^{32}P-labeled PCR amplification product of the plasmid pGL2-VEGF-*Luc*, which contains the VEGF HRE. Polyamide conjugate **1** was previously found to have $K_a = 6.3 \times 10^9$ M^{-1} at this site (16). Polyamides **4**–**15** were footprinted (Figure 2.4, Figure 2.5), and the K_a values obtained range from 4.4 (± 1.7) × 10^9 for **15** to 7.7 (± 0.9) × 10^{10} for **4** (Table 2.1). With the exception of **15**, the values obtained are all higher than that measured for **1**. Polyamides **4**–**15** retain good specificity for the 5'-ATACGT-3' HRE binding site, although a modest degree of non-specific binding is observed at high concentrations for polyamides **4**, **5**, **7**, **8** and **10**. Modifications to the tail moiety did not significantly abrogate high binding affinity and specificity.

Table 2.1: Equilibrium Association Constants: K_a (M^{-1})

Polyamide	5'-TACGTG-3'
4	7.7 (± 0.9) x 10^{10}
5	6.2 (± 1.2) x 10^{10}
6	2.5 (± 0.3) x 10^{10}
7	4.0 (± 0.9) x 10^{10}
8	3.8 (± 0.5) x 10^{10}
9	1.5 (± 0.6) x 10^{10}
10	2.1 (± 0.4) x 10^{10}
11	2.6 (± 0.4) x 10^{10}
12	1.1 (± 0.2) x 10^{10}
13	7.2 (± 1.1) x 10^9
14	4.8 (± 1.9) x 10^9
15	4.4 (± 1.7) x 10^9

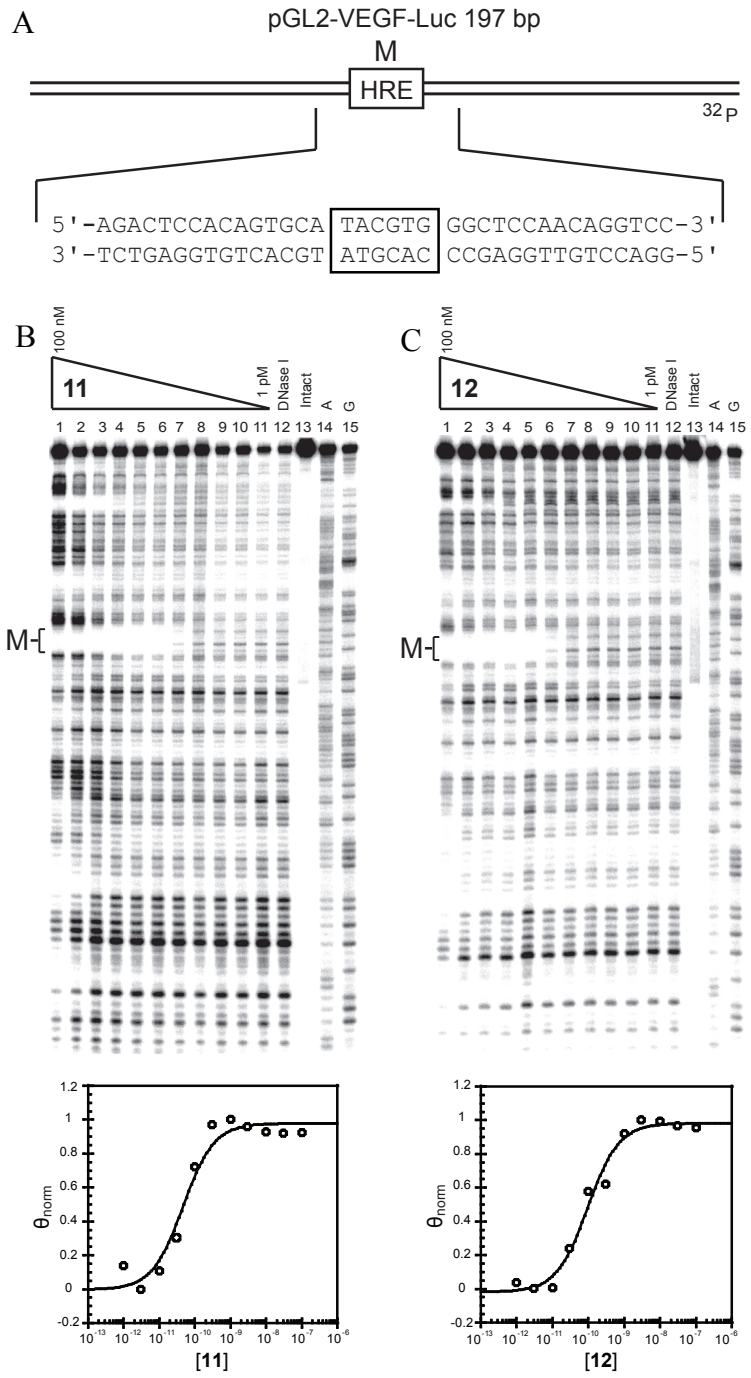


Figure 2.4. Plasmid sequence and DNA binding properties of polyamides **11** and **12**. (A) Illustration of plasmid pGL2-VEGF-*Luc*. The HRE match binding site M is indicated by the box. (B-C) Quantitative DNase I footprint titration experiments for polyamides **11** (B) and **12** (C) on the 197 bp, 5'-end-labeled PCR product of plasmid pGL2-VEGF-*Luc*: lanes 1–11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by a binding isotherm for polyamides binding the match site M (below).

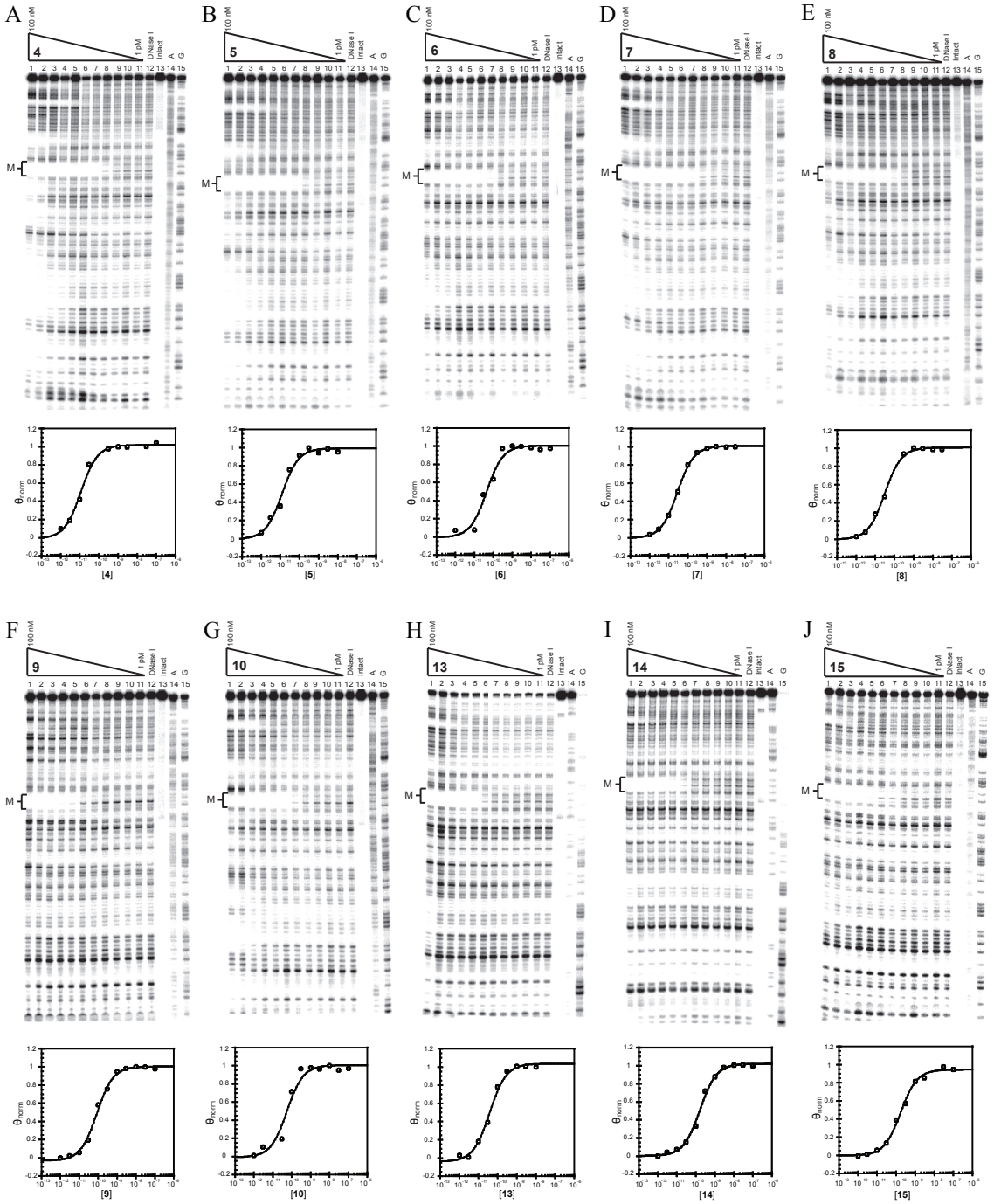


Figure 2.5. Quantitative DNase I footprint titration experiments to determine DNA binding properties and affinities of 4–10 and 13–15. These experiments were performed on the 197 bp, 5'-end-labeled PCR product of pGL2-VEGF-*Luc* plasmid, as above. The HRE match binding site M is indicated for each gel. Lanes as follows: lanes 1–11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Each footprinting gel is accompanied by a binding isotherm for polyamides binding the match site M (below).

Screening for Suppression of Hypoxia-Inducible Transcription by 4–27 in HeLa cells

Induction of VEGF mRNA by the hypoxia mimetic DFO in HeLa cells in the presence of polyamides **4**–**15** was measured by quantitative real-time RT-PCR and compared to that of polyamide **1** (Figure 2.6). Polyamides **5**, **6**, and **8** showed modest inhibition of induced VEGF expression and polyamides **4** and **7** showed no difference relative to the untreated, induced control. Mismatch polyamides **16**–**19** showed no difference in VEGF mRNA relative to the untreated, induced control. Polyamide **20** showed a modest effect.

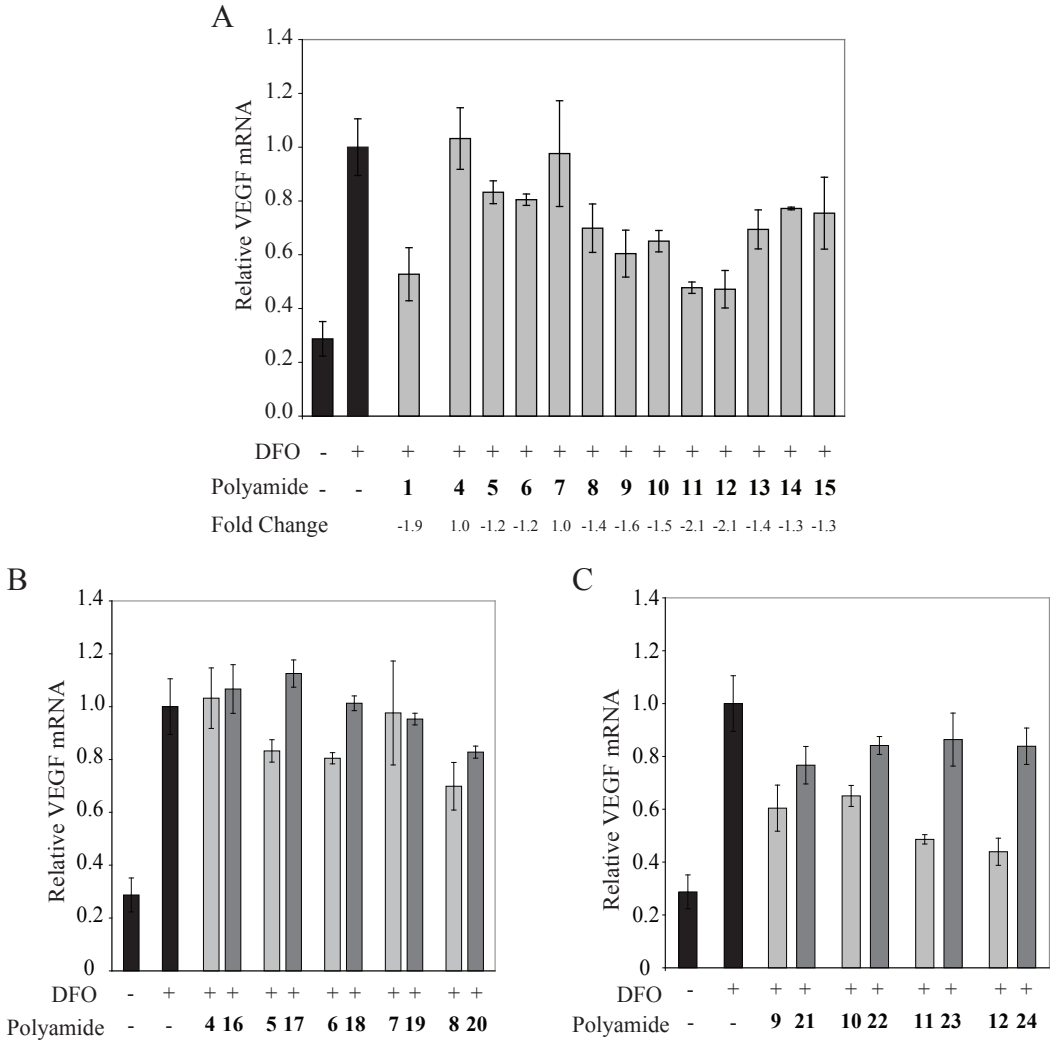


Figure 2.6. Induction of VEGF mRNA in cultured HeLa cells by the hypoxia mimetic deferoxamine (DFO) in the presence of polyamides. A) Induction of VEGF mRNA by DFO in the presence of polyamides **1** and **4**–**15** measured by quantitative real-time PCR. Concentrations of polyamides are 1 μM. B) Induction of VEGF mRNA by the hypoxia mimetic deferoxamine (DFO) in the presence of polyamides **11**, **12**, **23**, and **24** measured by quantitative real-time PCR. Polyamides **11** and **12**, which bind the 5'-ATACGT-3' sequence found in the VEGF HRE, show inhibition of DFO-induced VEGF expression. Polyamides **23** and **24**, which target 5'-WGGWCW-3', have a more modest effect. Concentrations of polyamides are 0.2 and 1 μM.

Polyamides **9** and **10** had a modest effect on induced VEGF expression, whereas **11** and **12** inhibited induced VEGF expression comparably to polyamide **1**. Control polyamides **23** and **24** had a reduced effect as compared to their match congeners **11** and **12** (Figure 2.7A). HeLa cell growth was not inhibited by 1 μM **11**, **12**, **23** or **24** (Figure 2.7B). Polyamides **13**–**15** had a modest effect on induced VEGF expression. An appreciable difference in relative VEGF mRNA levels was seen in cells treated with biotin conjugates **13**, **14** and mismatch congeners **25**, **26** (Figure 2.7C). It is noted that cells treated with polyamides **25** and **26** had VEGF expression slightly higher than that of the untreated, induced control.

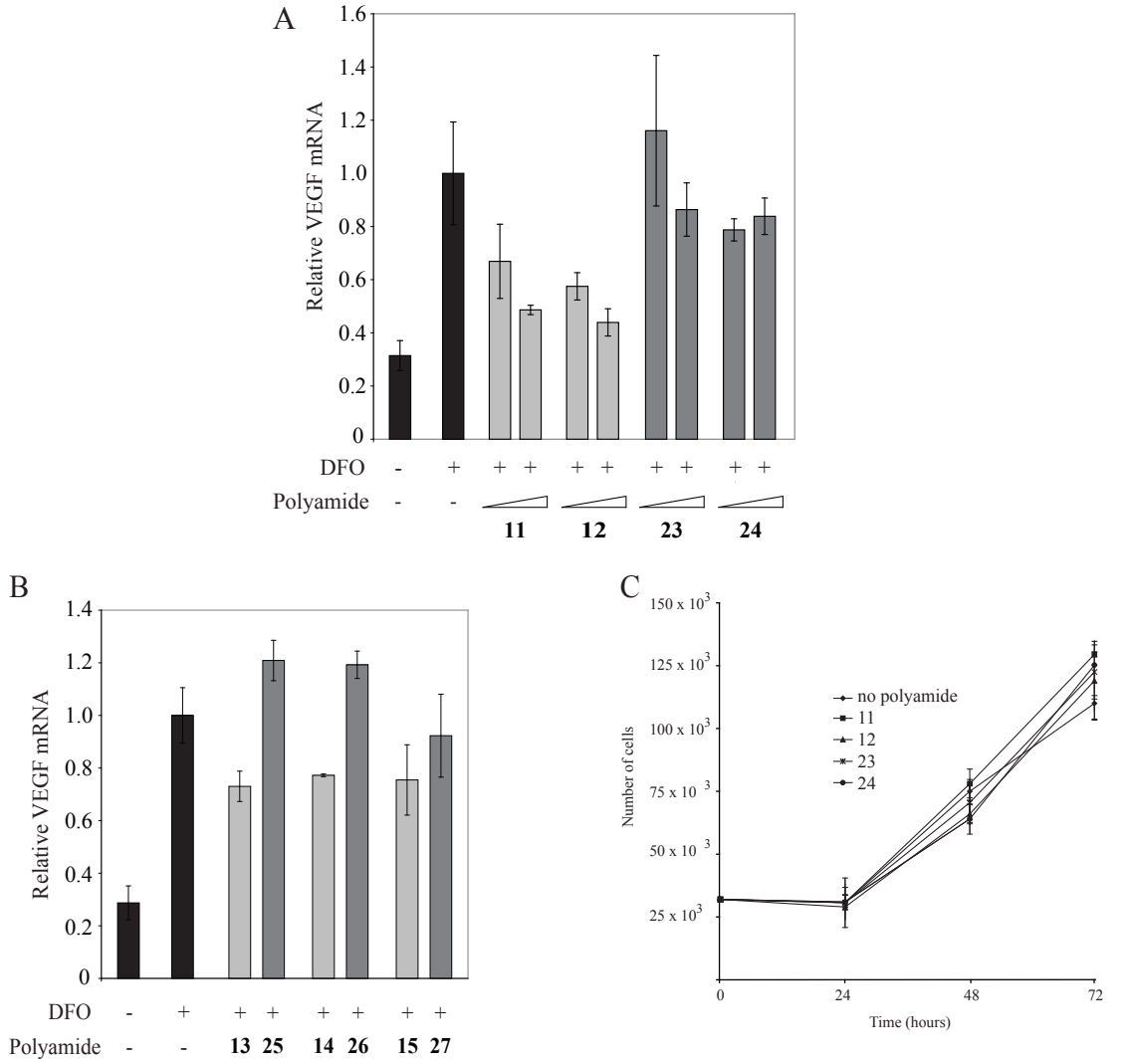
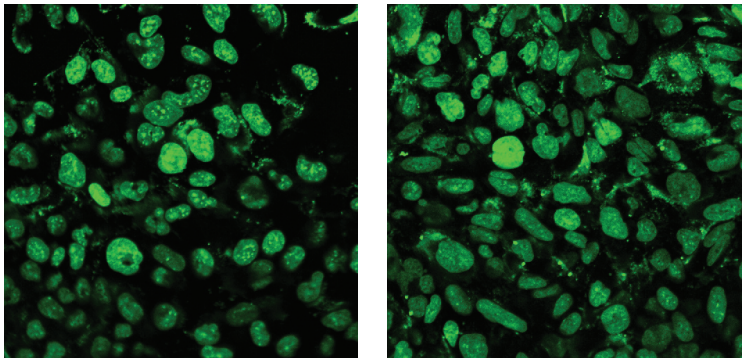


Figure 2.7. (A) Induction of VEGF mRNA by hypoxia mimetic deferoxamine (DFO) in the presence of polyamides **11**, 12, 23 and 24 (and 1 mM) measured by quantitative real-time PCR. (B) Induction of VEGF mRNA by DFO in the presence of 1 μM 13–15 and 25–27. (C) HeLa cell growth in the presence of polyamides **11**, **12**, **23**, and **24** at 1 μM as a function of time. Cells were counted at 24 hour intervals using a hemacytometer.

Suppression of Hypoxia-Inducible Transcription by Polyamides in U251 Cells

The effects of HRE targeted polyamides **1** and **11** and control polyamides **2** and **23** were tested in U251 cells, which express a higher level of VEGF than HeLa cells. Uptake of optically tagged polyamides **1** and **2** was examined by confocal microscopy in live, unfixed U251 cells dosed with 2 μM of **1** or **2** (Figure 2.8).



1 **2**

Figure 2.8. Uptake of polyamides **1** (left) and **2** (right) in U251 cells.

Induction of VEGF mRNA in the presence of polyamides **1** and **11** was inhibited dose dependently (Figure 2.9A). Polyamide **2** at had minimal effect on VEGF expression at 0.2 μM. However, at 1 μM, polyamide **2** had a moderate effect, though still less than that of its match congener **1**. Polyamide **23** had a more modest effect than its match congener **11** at both 0.2 and 1 μM. We also undertook chromatin immunoprecipitation experiments in U251 cells to assess changes in VEGF promoter occupancy by HIF-1α in the presence of **1**, **2**, **11**, and **23** (Figure 2.9B). Chromatin immunoprecipitation assays with anti-HIF1α or mock antibody treatment are consistent with decreased occupancy of HIF-1α at the VEGF HRE in the presence of polyamides **1** and **11** at 1 μM. Polyamide **23** has a more modest effect compared to that of its match congener **11**. At the concentration tested (1 μM), polyamide **2** has an effect statistically similar, though slightly less, than that of **1**. This is not unexpected; **2** also showed some inhibition of VEGF mRNA at this concentration. PCR amplification at an irrelevant locus in the coding region of GUSB showed minimal enrichment (~1–1.5-fold over background) for polyamide treated and control samples (data not shown).

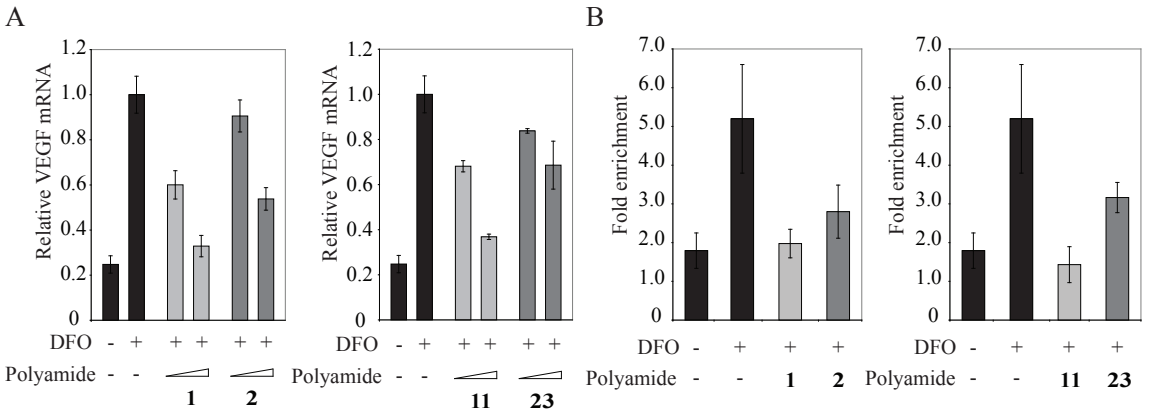


Figure 2.9. (A) Induction of VEGF mRNA by the hypoxia mimetic deferoxamine (DFO) in the presence of polyamides **1**, **2**, **11**, and **23** in U251 cells measured by quantitative real-time PCR. Concentrations of polyamides are 0.2 and 1 μM. (B) Chromatin immunoprecipitation assays with anti-HIF1α or mock antibody treatment expressed as fold-enrichment (specific/mock) of a 120 base pair sequence at the *vegf* HRE measured by real-time PCR. Concentrations of polyamides are 1 μM.

2.5 Discussion

Previous studies on fluorophore-labeled polyamides have revealed several parameters that affect cellular trafficking of these molecules, including number of heterocyclic ring pairs, relative Py/Im content and sequence, and choice of fluorophore and attachment site (8, 9). Among fluorophores considered, the protonation state of fluorescein is pH sensitive, while those of tetramethylrhodamine and Bodipy are not. In previous work, nuclear localization of polyamide **2** in HeLa cells was ablated when the cells were grown in glucose- and sodium pyruvate-free media, but subsequently restored when normal growth media was added (8). This suggests that for some polyamides, nuclear localization is energy dependent. Many fluorophore-polyamide conjugates that exhibit a punctate, cytoplasmic staining pattern co-localize with LysoTracker, a lysosomal stain. Verapamil, a calcium channel and P-glycoprotein efflux pump inhibitor, was shown to facilitate nuclear uptake of a Bodipy-polyamide conjugate otherwise sequestered within cytoplasmic vesicles (7). Some polyamides with minimal nuclear localization, e.g. **3**, localize mostly in the extracellular space. Nevertheless, the fact that many eight-ring hairpin polyamide-fluorophore conjugates do localize in the nuclei of live cells is encouraging.

The inhibition of induced VEGF expression in the presence of polyamide **1**, which binds the VEGF HRE, is an important proof-of-principle experiment for the use

of designed DNA-binding ligands to modulate the expression of specific sets of genes. Moving ahead, anticipated experiments include studies in mammalian model systems to account for biodistribution, availability, and metabolism of polyamides. Towards this end, we have attempted to identify polyamides capable of nuclear localization but of simpler structures and lower molecular weights than dye conjugates. The inhibition of induced VEGF expression with HRE-targeted polyamide **1**, but not **2**, which targets a different DNA sequence or **3**, which exhibits considerably less nuclear localization, demonstrates the viability of VEGF inhibition as a proxy for nuclear localization of polyamides that bind the HRE.

The ability to interrogate polyamide cellular uptake and nuclear localization by affecting expression of an endogenous inducible gene enables exploration of the nuclear uptake profiles of non-fluorescent tail moieties. The quantitative real-time RT-PCR data presented show that removal of the fluorescein dye (polyamide **4**) and step-wise ablation of the linker moiety (polyamides **5** and **6**) diminish the biological activity of these compounds as compared to polyamide **1** in this system. The K_a values measured for **4**–**6** are between four– and 12–fold higher than that measured for the parent polyamide **1** on the same DNA sequence. Similarly, the binding affinities of *N*-alkylaminemorpholino tail polyamides **7** and **8** are ~six-fold higher than that of the parent polyamide **1** for the same binding site, but only modest biological activities are observed.

Data generated from cultured HeLa cells treated with polyamides **9**–**12** were more encouraging. These compounds were designed to probe contributions of the alcohol and carboxylic acid groups of fluorescein to the cellular uptake and nuclear localization observed for **1**. This was done by synthesizing a focused series of polyamides comprised of an amide-bond linked benzene ring containing four possible permutations of alcohol and/or carboxylic acid substituents at the *meta*- positions. Polyamides **9**–**12** demonstrated good sequence specificities and binding affinities greater than that of **1**. Polyamide conjugates of benzoic acid and 3-hydroxy benzoic acid, **9** and **10**, respectively, show a modest degree of

biological activity, but those conjugated to either isophthalic acid or 5-hydroxy isophthalic acid, **11** and **12**, respectively, demonstrate a biological effect similar to that observed for the parent polyamide **1**. These results suggest that the presence of the carboxylic acid group on the benzene ring is a positive determinant of nuclear localization in this system.

Biotin conjugates **13**–**15** differ in the identity of the linker moiety and were of interest due to the potential utility of these conjugates for pull-down experiments. The binding affinities measured by quantitative DNase I footprint titration experiments for these conjugates were similar to that of the parent polyamide **1**. By quantitative real-time RT-PCR, polyamides **13**–**15** all demonstrated modest biological activities, and the results do not suggest strong differences in activities based on linker identity. The modest activities of the biotin conjugates, which might reflect a moderate degree of nuclear uptake, are neither encouraging nor dismissive of potential applications of these conjugates.

Our working hypothesis regarding VEGF inhibition by polyamides is through direct interactions with the DNA-HIF-1 interface at the HRE that prevents HIF-1 binding, most likely by an allosteric mechanism. To explore this model further, we used chromatin immunoprecipitation to assess HIF-1 occupancy at the VEGF HRE under hypoxia induced conditions in HeLa cells in the presence and absence of match (and mismatch) polyamides with variable C-terminus tails. Antibodies against HIF-1α were found to enrich fragments of DNA containing the VEGF HRE after DFO-induction. This enrichment was reduced in samples treated with polyamide **1** and, to a lesser extent, polyamide **2**. We also undertook experiments in U251 cells, which express VEGF more robustly and have previously been used as a model cell line for studying hypoxia-induced VEGF expression (17).

Polyamide-fluorescein conjugates **1** and **2** localize in the nucleus of U251 cells. Induction of VEGF mRNA in the presence of polyamides **1**, **2**, **11**, and **23** at 0.2 and 1 μM in U251 cells was measured by quantitative real-time RT-PCR. Polyamide **1** showed significant inhibition at both concentrations. Interestingly, at 1 μM polyamide **2** showed inhibition greater than that measured in HeLa cells but still significantly less than **1**.

Polyamide **11** exhibits dose-dependent inhibition of VEGF similar to that measured in HeLa cells, while **23** showed modest activity at both concentrations comparable to that measured in HeLa cells. Chromatin immunoprecipitation of HIF-1α from U251 cells treated with HRE-targeted polyamides **1** and **11** and mismatch polyamides **2** and **23** shows that enrichment of HIF-1α bound DNA fragments containing the VEGF HRE after DFO induction is inhibited by pre-treatment with polyamides **1** and **11**, and to a lesser extent **2** and **23**. This is generally consistent with a mechanism whereby polyamides **1** and **11** exert their effect by preventing HIF-1 binding to the cognate HRE. This is also consistent with DNase I footprint titrations and gel shift assays *in vitro* (16).

This report identifies two polyamides, **11** and **12**, with non-fluorescent moieties at the carboxy terminus that bind the VEGF HRE that have biological activity similar to that of the fluorescein-polyamide conjugate **1** known to localize in the nucleus. Both polyamides retain an amide-linked benzene ring with a carboxylic acid *meta*- to the attachment point and are similar in structure to the proximal portion of fluorescein. Work is ongoing to synthesize a second series of polyamides designed to examine the role of the aromatic ring and the carboxylic acid functional group in aiding nuclear localization. As the mechanism of polyamide uptake and nuclear localization continues to be elucidated, it is currently unclear whether the biological activity of **11** and **12** is due to improved ability to buffer the wide range of pH values that may be encountered in the cell (e.g. lysosomes), improved nuclear localization resulting in an increased effective polyamide concentration at the DNA, or decreased efflux of polyamide from the cell perhaps through one or several ATP-dependent drug efflux pumps (26, 27).

Table 2.2: MALDI-ToF MS Data for Polyamide Conjugates **3**–**27**

MALDI-TOF MS Data for Polyamide Conjugates (3–27):

(3) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-β-$NH(CH_2)_3N(CH_3)(CH_2)_3NH$(FAM): MALDI-TOF-MS Calculated [M+H^+]: 1676.17; Found: 1676.9.

(4) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3N(CH_3)(CH_2)_3NH_2$: MALDI-TOF-MS Calculated [M+H^+]: 1246.50; Found: 1246.5.

(5) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3N(CH_3)_2$: MALDI-TOF-MS Calculated [M+H^+]:1203.46 ; Found: 1203.4.

(6) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_3)$: MALDI-TOF-MS Calculated [M+H^+]: 1132.38; Found: 1132.4.

(7) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_2$(Morpholine): MALDI-TOF-MS Calculated [M+H^+]:1231.45; Found:1231.7.

(8) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3$(Morpholine): MALDI-TOF-MS Calculated [M+H^+]:1245.47 ; Found: 1245.9.

(9) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3N(CH_3)(CH_2)_3NH$(Benzoic acid): MALDI-TOF-MS Calculated [M+H^+]: 1350.53; Found: 1350.5.

(10) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3N(CH_3)(CH_2)_3NH$(3-Hydroxy benzoic acid): MALDI-TOF-MS Calculated [M+H^+]: 1366.52; Found: 1366.4.

(11) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3N(CH_3)(CH_2)_3NH$(Isophthalic acid):. MALDI-TOF-MS Calculated [M+H^+]:1394.52; Found: 1394.5.

(12) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3N(CH_3)(CH_2)_3NH$(5-Hydroxyisophthalic acid): MALDI-TOF-MS Calculated [M+H^+]:1410.51; Found: 1410.5.

(13) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3N(CH_3)(CH_2)_3NH$(Biotin): MALDI-TOF-MS Calculated [M+H^+]: 1474.10; Found: 1474.5.

(14) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_2(O(CH_2)_2)_2NH$(Biotin): MALDI-TOF-MS Calculated [M+H^+]: 1477.05; Found: 1476.8.

(15) CtPyPyIm-(R)-γ(NH_2)-PyImPyPy-$NH(CH_2)_3(O(CH_2)_2)_3CH_2NH$(Biotin): MALDI-TOF-MS Calculated [M+H^+]: 1549.16; Found: 1549.5.

(16) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-$NH(CH_2)_3N(CH_3)(CH_2)_3NH_2$: MALDI-TOF-MS Calculated [M+H^+]:1210.59; Found:1210.7.

(17) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-$NH(CH_2)_3N(CH_3)_2$: MALDI-TOF-MS Calculated [M+H^+]:1267.55; Found:1267.2.

(18) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-$NH(CH_3)$: MALDI-TOF-MS Calculated [M+H^+]:1096.47; Found:1096.8.

(19) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_2$(Morpholine): MALDI-TOF-MS Calculated [M+H^+]:1195.54; Found:1195.6.

(20) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy- NH(CH_2)$_3$(Morpholine): MALDI-TOF-MS Calculated [M+H^+]: 1210.56; Found: 1210.0.

(21) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_3$N(CH_3)(CH_2)$_3$NH(Benzoic acid): MALDI-TOF-MS Calculated [M+H^+]:1314.61; Found: 1314.7.

(22) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_3$N(CH_3)(CH_2)$_3$NH(3-Hydroxybenzoic acid): MALDI-TOF-MS Calculated [M+H^+]:1330.61; Found: 1330.5.

(23) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_3$N(CH_3)(CH_2)$_3$NH(Isophthalic acid): MALDI-TOF-MS Calculated [M+H^+]: 1358.60; Found: 1358.6.

(24) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_3$N(CH_3)(CH_2)$_3$NH(5-Hydroxyisophthalic acid): MALDI-TOF-MS Calculated [M+H^+]: 1374.60; Found: 1374.7.

(25) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_3$N(CH_3)(CH_2)$_3$NH(Biotin): MALDI-TOF-MS Calculated [M+H^+]: 1438.62; Found: 1439.0.

(26) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_2$(O(CH_2)$_2$)$_2$NH(Biotin): MALDI-TOF-MS Calculated [M+H^+]: 1440.57; Found: 1440.4.

(27) ImImPyPy-(R)-γ(NH_2)-ImPyPyPy-NH(CH_2)$_3$(O(CH_2)$_2$)$_3$$CH_2$NH(Biotin): MALDI-TOF-MS Calculated [M+H^+]: 1512.68; Found: 1512.8.

References

1. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.*, **13**, 284-299.
2. Dervan, P.B., Poulin-Kerstien, A.T., Fechter, E.J. and Edelson, B.S. (2005) Regulation of gene expression by synthetic DNA-binding ligands. *Top. Curr. Chem*, **253**, 1-31.
3. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
4. Kielkopf, C.L., Baird, E.E., Dervan, P.D. and Rees, D.C. (1998) Structural basis for G•C recognition in the DNA minor groove. *Nat. Struct. Biol.*, **5**, 104-109.
5. Foister, S., Marques, M.A., Doss, R.M. and Dervan, P.B. (2003) Shape selective recognition of T•A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues. *Bioorg. Med. Chem.*, **11**, 4333-4340.
6. Belitsky, J.M., Leslie, S.J., Arora, P.S., Beerman, T.A. and Dervan, P.B. (2002) Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. *Bioorg. Med. Chem.*, **10**, 3313-3318.
7. Crowley, K.S., Phillion, D.P., Woodard, S.S., Schweitzer, B.A., Singh, M., Shabany, H., Burnette, B., Hippenmeyer, P., Heitmeier, M. and Bashkin, J.K. (2003) Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorganic & Medicinal Chemistry Letters*, **13**, 1565-1570.
8. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. U. S. A.*, **100**, 12063-12068.

9. Edelson, B.S., Best, T.P., Olenyuk, B., Nickols, N.G., Doss, R.M., Foister, S., Heckel, A. and Dervan, P.B. (2004) Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.*, **32**, 2802-2818.

10. Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B. and Luger, K. (2003) Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.*, **326**, 371-380.

11. Dudouet, B., Burnett, R., Dickinson, L.A., Wood, M.R., Melander, C., Belitsky, J.M., Edelson, B., Wurtz, N., Briehn, C., Dervan, P.B., and Gottesfeld, J.M. (2003) Accessibility of nuclear chromatin by DNA binding polyamides. *Chem. Biol.*, **10**, 859-867.

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

13. Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M. and Dervan, P.B. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 12890-12895.

14. Nguyen-Hackley, D.H., Ramm, E., Taylor, C.M., Joung, J.K., Dervan, P.B. and Pabo, C.O. (2004) Allosteric inhibition of zinc-finger binding in the major groove of DNA by minor-groove binding ligands. *Biochemistry*, **43**, 3880-3890.

15. Wurtz, N.R., Pomerantz, J.L., Baltimore, D. and Dervan, P.B. (2002) Inhibition of DNA binding by NF-kappa B with pyrrole-imidazole polyamides. *Biochemistry*, **41**, 7604-7609.

16. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 16768-16773.

17. Kong, D.H., Park, E.J., Stephen, A.G., Calvani, M., Cardellina, J.H., Monks, A., Fisher, R.J., Shoemaker, R.H. and Melillo, G. (2005) Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity. *Cancer Research*, **65**, 9047-9055.

18. Vandyke, M.M. and Dervan, P.B. (1984) Echinomycin Binding-Sites on DNA. *Science*, **225**, 1122-1127.

19. Baird, E.E. and Dervan, P.B. (1996) Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.*, **118**, 6141-6146.

20. Belitsky, J.M., Nguyen, D.H., Wurtz, N.R. and Dervan, P.B. (2002) Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.*, **10**, 2767-2774.

21. Trauger, J.W. and Dervan, P.B. (2001), Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. *Methods Enzymol.*, **340**, 450-466.

22. Wood, S.M. and Ratcliffe, P.J. (1997) Mammalian oxygen sensing and hypoxia inducible factor-1. *International Journal of Biochemistry & Cell Biology*, **29**, 1419-1432.

23. Bianchi, L., Tacchini, L. and Cairo, G. (1999) HIF-1-mediated activation of transferrin receptor gene transcription by iron chelation. *Nucleic Acids Res.*, **27**, 4223-4227.

24. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.*, **9**, 2215-2235.

25. Kaizerman, J.A., Gross, M.L., Ge, Y.G., White, S., Hu, W.H., Duan, J.X., Baird, E.E., Johnson, K.W., Tanaka, R.D., Moser, H.E., and Burli, R.W. (2003) DNA binding ligands targeting drug-resistant bacteria: Structure, activity, and pharmacology. *Journal of Medicinal Chemistry*, **46**, 3914-3929.

26. Juliano, R.L. and Ling, V. (1976) Surface Glycoprotein Modulating Drug

Permeability in Chinese-Hamster Ovary Cell Mutants. *Biochimica et Biophysica Acta*, **455**, 152-162.

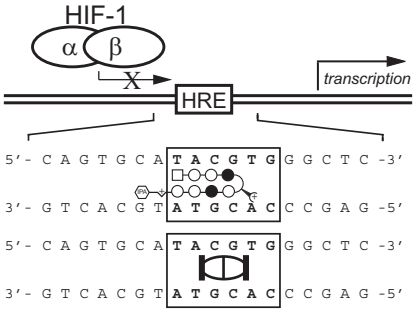
27. Ueda, K., Cardarelli, C., Gottesman, M.M. and Pastan, I. (1987) Expression of a Full-Length cDNA for the Human Mdr1 Gene Confers Resistance to Colchicine, Doxorubicin, and Vinblastine. *Proc. Natl. Acad. Sci. U. S. A.*, **84**, 3004-3008.

Chapter 3

Modulating Hypoxia-Inducible Transcription by Disrupting the HIF-1-DNA Interface

3.1 Abstract

Transcription mediated by hypoxia inducible factor (HIF-1) contributes to tumor angiogenesis and metastasis but is also involved in activation of cell-death pathways and normal physiological processes. Given the complexity of HIF-1 signaling, it could be advantageous to target a subset of HIF-1 effectors rather than the entire pathway. We compare the genome-wide effects of three molecules that each interfere with the HIF-1-DNA interaction: a polyamide targeted to the hypoxia response element (HRE), siRNA targeted to HIF-1α, and echinomycin, a DNA-binding natural product with a similar but less specific sequence preference than the polyamide. The polyamide affects a subset of hypoxia-induced genes consistent with its binding site preferences. For comparison, HIF-1α siRNA and echinomycin each affect the expression of nearly every gene induced by hypoxia. Remarkably, the total number of genes affected by either polyamide or HIF-1α siRNA over a range of thresholds is comparable. The data show that polyamides can be used to affect a subset of a pathway regulated by a transcription factor. In addition, this study offers a unique comparison of three complementary approaches towards exogenous control of endogenous gene expression.



3.2 Introduction

Exogenous chemical and biological methods to control directly expression of selected endogenous genes could have broad implications for human medicine. Towards this goal, a number of technological approaches are currently being investigated. Polydactyl zinc finger proteins are a programmable class of DNA binding proteins capable of sequence-specific binding (*1, 2*). These designed proteins have been used to inhibit expression of target genes (*3*), and transcriptional activator domain-zinc finger conjugates have been used to activate expression of target genes (*4*). The RNA-interference pathway can be used to regulate gene expression at the post-transcriptional level (*5*). siRNA and shRNA molecules enlist cellular machinery to degrade selected mRNA targets (*6, 7*). RNAi technology has been highly effective in achieving potent and specific knock-down of target mRNAs and is now widely used to probe target gene function (*8*). However, bioavailability and delivery of zinc finger proteins and siRNA to targets in humans could be an obstacle to their therapeutic application and continues to receive considerable attention (*8*). Cell-permeable small molecules that modulate protein-protein or protein-DNA interactions offer another approach to the control of endogenous gene regulation. Screening small molecule and natural product libraries for a desired effect can identify candidate molecules with high likelihoods of possessing drug-like bioavailability; drawbacks include the need to screen anew for each target protein-protein or protein-DNA interaction. Polyamides containing *N*-methylimidazole (Im) and *N*-methylpyrrole (Py) are a class of programmable DNA-binding small molecules previously used to disrupt protein-DNA interactions in a sequence-specific manner in cell culture (*9, 10*).

Controlling the transcriptional activity of hypoxia inducible factor (HIF-1), a heterodimer of HIF-1α and HIF-1β (ARNT), is a clinically relevant goal (*11–14*). HIF-1 is the principal mediator of the adaptive cellular response to hypoxia (*15*). Under normoxic conditions, HIF-1α is specifically hydroxylated by an iron-dependent proline hydroxylase, ubiquitinated by the von Hippel-Lindau (VHL) tumor suppressor protein-ubiquitin ligase

protein complex, and degraded by the proteosome (*16*). Iron chelators, such as deferoxamine (DFO), can be used to mimic hypoxia in cell culture (*16*). Through interaction with co-activators p300/CBP, HIF-1 directly activates the expression of at least 100 genes involved in cellular and tissue adaptation to hypoxia (*13*), including pro-angiogenic factors such as vascular endothelial growth factor (VEGF), glycolytic enzymes, extra-cellular matrix remodeling enzymes, and genes involved in both pro-survival and death pathways (*11*). HIF-1 activation by the hypoxic microenvironment of solid tumors, or by deactivating mutations in VHL, contribute to an aggressive phenotype of increased cell proliferation, invasion, metastasis, and drug resistance (*11*). Given the complexity of HIF-1 signaling in cellular survival and death pathways, and its critical role in physiological processes in normal tissues, it could be advantageous to target a subset of HIF-1 effectors rather than the entire pathway (*13*).

In important proof-of-principle experiments, introduction of siRNA against HIF-2α to $VHL^{-/-}$ renal carcinoma cells was sufficient to abrogate tumor formation by these cells in mice (*17*). Screening for small molecules capable of disrupting the HIF-1-p300 interaction identified chetomin, a natural product that binds p300 that was shown to inhibit expression of HIF-1 regulated genes and exhibit anti-tumor activity in a mouse model (*18*). In an effort to inhibit directly HIF-1-DNA binding, a hairpin polyamide was designed to bind the sequence 5'-ATACGT-3' found in the VEGF HRE. This polyamide bound its target site and prevented HIF-1-DNA binding in a sequence specific manner, and inhibited hypoxia-induced expression of VEGF and several other HIF-1 regulated genes in cultured cells without the use of transfection agents (*19, 20*). Melillo and colleagues screened a library of small molecules for inhibition of HIF-1 mediated transcription in a cell-based assay and identified the natural product echinomycin, a DNA-binding bisintercalator (*21*). Echinomycin binds the four basepair sequence 5'-NCGN-3' found in the consensus HRE 5'-NACGTG-3' and inhibited expression of VEGF in cultured cells (*22*).

To establish a benchmark of comparison for the global effects of polyamides,

we compare the genome-wide effects on U251 cells induced with DFO of a polyamide targeted to the HRE sequence 5'-WTWCGW-3', echinomycin, and siRNA targeted against HIF-1α. siRNA-mediated destruction of HIF-1α mRNA establishes a maximum level of inhibition that can be achieved for HIF-1 target genes through disruption of the HIF-1-HRE interaction. Nearly all transcripts induced by DFO are inhibited by both HIF-1α siRNA and echinomycin. Polyamide **1** inhibits only a subset of these genes, and shows a preference for genes containing HREs of the sequence 5'-(T/A)ACGTG-3', consistent with this molecule's predicted binding preferences. Remarkably, the total number of genes affected by either polyamide **1** or HIF-1α siRNA over a range of thresholds is comparable. We show that HIF-1 occupancy at the HREs of two genes affected by polyamide **1** is reduced in the presence of the polyamide, while HIF-1 occupancy at the HREs of two unaffected genes is unchanged. We also show that a polyamide that binds a site immediately 5' to the VEGF HRE inhibits induced expression of VEGF but not of FLT1, a HIF-1 target gene lacking this flanking site.

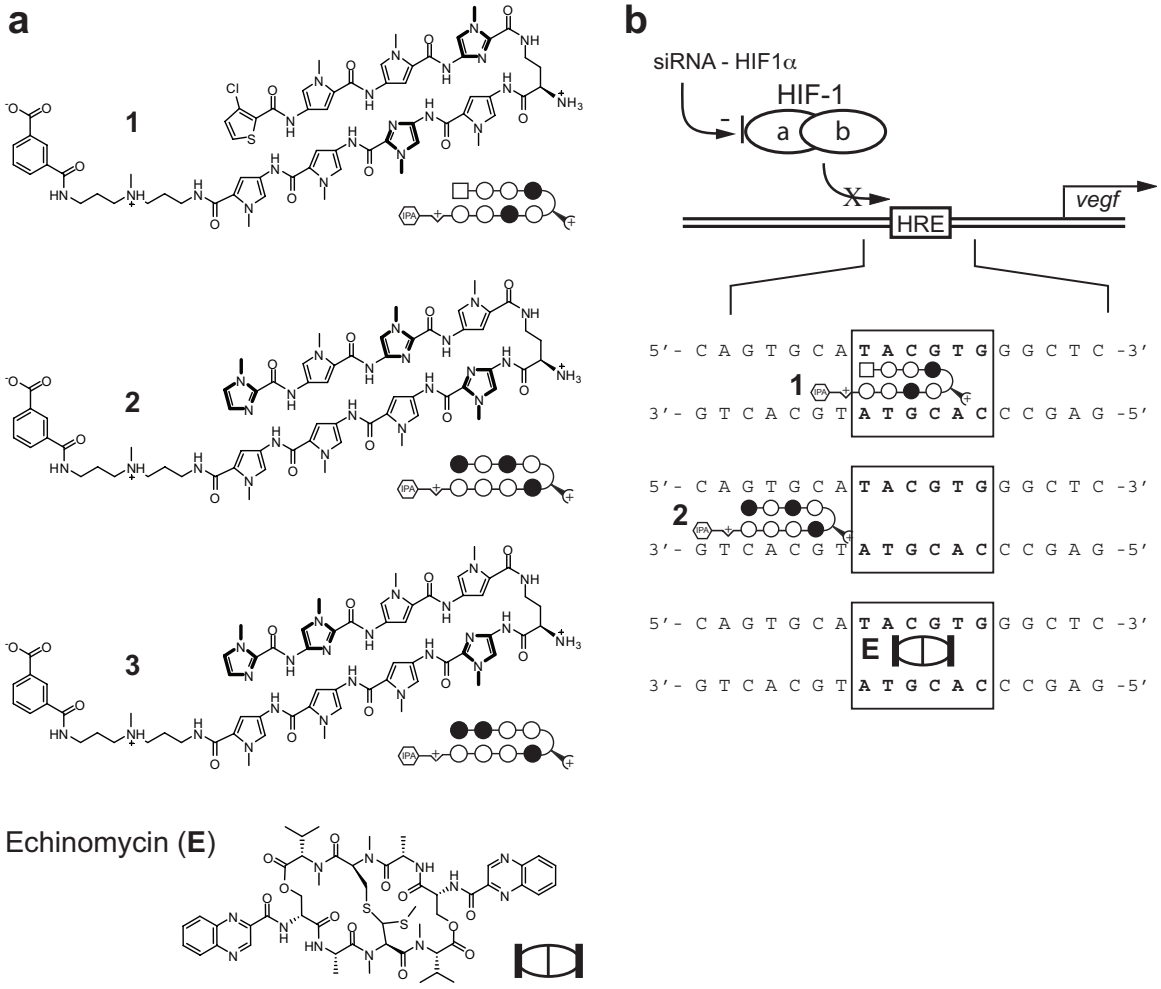


Figure 3.1. Structures of molecules used in this study. a) Structures of polyamides **1**–**3** and echinomycin. Imidazole, pyrrole and chlorothiophene monomer units are represented, respectively, by closed circles, open circles and squares. b) Three approaches to inhibiting HIF-1 induced gene expression: sequence-specific small molecule binding to the HRE by a polyamide or echinomycin, and reduction in HIF-1α mRNA using siRNA.

3.3 Results and Discussion

Binding of Polyamide 1, 2, 3 and Echinomycin at the VEGF and FLT1 HREs. Polyamide sequence specificity is programmed by side-by-side pairings of heterocyclic amino acids in the minor groove of DNA: Im/Py distinguishes G•C from C•G; Py/Py binds both A•T and T•A; 3-chlorothiophene/N-methylpyrrole (Ct/Py) prefers T•A at the N-terminus position (*23–25*). Polyamide **1** and echinomycin are expected to bind at the VEGF HRE sequence 5'-TACGTG-3'. Polyamide **2** is expected to bind the sequence 5'-AGTGCA-3' immediately 5' to the VEGF promoter HRE. HRE-mismatch control polyamide **3** targets the sequence 5'-WGGWCW-3', which is not found near the VEGF HRE. The DNA-binding affinities of

2, **3** and echinomycin for the VEGF HRE were measured by quantitative DNase I footprint titrations using a 5' ^{32}P-labeled PCR amplification product of the plasmid pGL2-VEGF-*Luc* containing the VEGF HRE (Figure 3.2A and 3.2B). Polyamide **1** was previously found to bind the VEGF HRE (K_a = 2.6 (±0.4) × 10^{10} M^{-1}) (*20*). For ease of comparison, a footprinting gel of **1** is included. Polyamide **2** binds the site 5'-AGTGCA-3' immediately 5' to the VEGF HRE (K_a = 3.2 (±0.6) × 10^9 M^{-1}). Echinomycin binds the VEGF HRE (K_a = 8.4 (±2.1) × 10^6 M^{-1}). Polyamide **3** binds the VEGF HRE (K_a = 8.0 (±1.0) × 10^8 M^{-1}), and approximately 35 basepairs 3' of the HRE, most likely two 5'-WGWCW-3' sites (K_a = 7.6 (±1.0) × 10^8 M^{-1}**)**.

The DNA-binding affinities of **1**, **2**, **3** and echinomycin for the HRE of FLT1 were measured by quantitative DNase I footprint titrations using a 5' ^{32}P-labeled PCR amplification product of the plasmid pCSJ-FLT1 (Figure 3.2C and 3.2D). Although formally targeted to the sequence 5'-WTWCGW-3', polyamide **1** would be expected to bind 5'-CAACGT-3', albeit with a moderate decrease in affinity (*25*). The sequence preference of a Ct/Py pair for T•A is approximately 4-fold over A•T, but 50-fold over a G•C (*25*). Polyamide **1** binds the FLT1 HRE (K_a = 2.7 (±0.2) × 10^9 M^{-1}) . Polyamide **2** binds the FLT1 HRE (K_a = 2.2 (±0.8) × 10^8 M^{-1}). Echinomycin binds the FLT1 HRE (K_a = 2.9 (±0.7) × 10^7 M^{-1}). Polyamide **3** does not bind the FLT1 HRE with a measurable K_a, but was observed to bind a 5'-AGACA-3' site 16 basepairs 5' to the FLT1 HRE (K_a = 2.7 (±0.4) × 10^9 M^{-1}). These K_a data are summarized in Table 3.1.

Table 3.1. K_a values (M^{-1}) for **1**, **2**, **3**, and Echinomycin (**E**) at HRE on pGL2-VEGF-*Luc* and pCSJ-FLT1

pGL2-VEGF-*Luc*	Ligand	Sequence	K_a (M^{-1})
	1	5'-ATACGT-3'	2.6 (±0.4) x 10^{10}
	2	5'-AGTGCA-3'	3.2 (±0.6) x 10^9
	3	5'-TGGGCT-3'	8.0 (±1.0) x 10^8
	E	5'-ATACGT-3'	8.4 (±2.1) x 10^6
pCSJ-FLT1	Ligand	Sequence	K_a (M^{-1})
	1	5'-CAACGT-3'	2.7 (±0.2) x 10^9
	2	5'-TGAGGA-3'	2.2 (±0.8) x 10^8
	3	5'-CAACGT-3'	--------
	E	5'-CAACGT-3'	2.9 (±0.7) x 10^7

K_a values reported are mean values of three DNase 1 footprint titration experiments; standard deviations in parentheses.

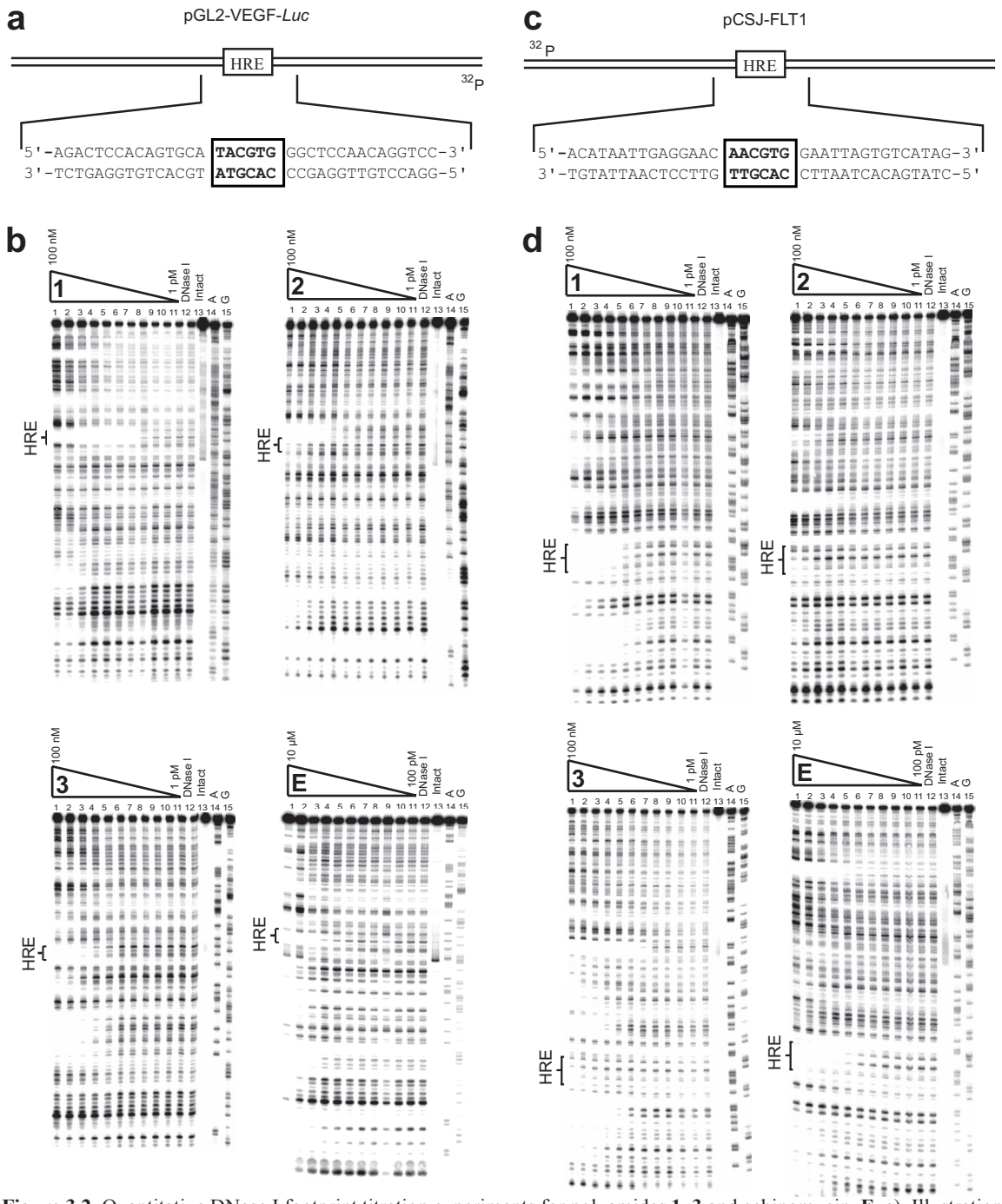


Figure 3.2. Quantitative DNase I footprint titration experiments for polyamides **1**–**3** and echinomycin, **E**. a) Illustration of pGL2-VEGF-*Luc* and partial sequence containing VEGF HRE and polyamides **1**, **2** and echinomycin putative binding sites. b) Quantitative DNase I footprint titration experiments for polyamides **1, 2**, **3**, and **E**, on 5'-end-labeled PCR product of plasmid pGL2-VEGF-*Luc*. For **1**, **2**, and **3**: lanes 1–11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM and 1 pM polyamide, respectively; lane 12, DNAse I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. For **E**: lanes 1–11, 10 µM, 3 µM, 1 µM, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM echinomycin, respectively; lanes 12-15 as above. Polyamide **1** and **E** $K_a = 2.6\ (\pm 0.4) \times 10^{10}\ M^{-1}$ and $K_a = 8.4\ (\pm 2.1) \times 10^{6}\ M^{-1}$, respectively, at VEGF HRE. Polyamide **2** $K_a = 3.2\ (\pm 0.6) \times 10^{9}\ M^{-1}$ for 5'-AGTGCA-3' site immediately 5' to VEGF HRE. Polyamide **3** $K_a = 8.0\ (\pm 1.0) \times 10^{8}\ M^{-1}$ for VEGF HRE. c) Illustration of pCSJ-FLT1 and partial sequence containing FLT1 HRE and putative binding sites for **1** and **E**. d) Quantitative DNase I footprint titration experiments for **1**, **2**, **3**, and **E** on 5' end-labeled PCR product of plasmid pCSJ-FLT1. Lane assignments for gels in d) are as for b). Polyamide **1** and **E** $K_a = 2.7\ (\pm 0.2) \times 10^{9}\ M^{-1}$ and $K_a = 2.9\ (\pm 0.7) \times 10^{7}\ M^{-1}$, respectively, at FLT1 HRE. Polyamide **2** $K_a = 2.2\ (\pm 0.8) \times 10^{8}$ at this site. Polyamide **3** does not bind FLT1 HRE with a measurable K_a.

Suppression of induced VEGF and FLT1 expression. HIF-1 induces VEGF expression by binding to the 5'-TACGTG-3' HRE located approximately 950 base pairs upstream of the transcription start site (*26, 27*). The effect on induced VEGF expression by siRNA silencing of HIF-1α mRNA establishes a theoretical maximum level of inhibition through disruption of HIF-1-DNA binding. HIF-1α mRNA was reduced by approximately 95% in the presence of HIF-1α siRNA, but was minimally affected by polyamides **1**, **2** or echinomycin under induced conditions (Figure 3.3). A mismatch control siRNA did not reduce HIF-1α mRNA. Polyamides **1** and **2** (1 μM) and HIF-1α siRNA had similar effects on induced VEGF expression; treatment inhibited most of the increase in VEGF mRNA following DFO treatment, but not to levels below that observed for non-induced controls (Figure 3.3A). As previously reported, 100 nM echinomycin potently inhibits VEGF expression to levels below the non-induced control (*21*). HRE-mismatch control polyamide **3**, which binds the HRE with a much reduced affinity relative to **1**, had a more modest effect on VEGF mRNA levels. It is also possible that the modest but measurable effect of polyamide **3** on VEGF expression could be due to interference with other protein-DNA interactions elsewhere in the promoter or enhancer of VEGF, or is due to secondary effects from direct effects on other genes. A mismatch control siRNA had a limited effect on VEGF mRNA levels.

Induction of FLT1 (VEFG receptor type 1) is mediated by HIF-1 binding to the 5'-AACGTG-3' HRE in the FLT1 promoter (*28*). Polyamide **1** and HIF-1α siRNA both inhibited FLT1 expression following DFO induction (Figure 3.3B). Echinomycin reduced FLT1 expression to below that of the non-induced control. Polyamides **2** and **3** had minimal effect. A mismatch control siRNA also had a limited effect on FLT1 mRNA levels. Given the relative binding affinities of polyamide **1** and echinomycin, it may be surprising that 1 μM of polyamide **1** is necessary to inhibit VEGF and FLT1 expression comparably to HIF-1α siRNA, while 100 nM echinomycin reduces their expression to sub-basal levels. The structure of double-helical DNA is not greatly perturbed by minor groove-binding hairpin

polyamides (*23*); echinomycin-DNA binding causes local unwinding and lengthening of the DNA helix, which might account for its greater potency in these experiments (*29, 30*). Polyamide-intercalator conjugates have been shown to unwind DNA in a sequence-specific fashion and to improve the ability of a polyamide to inhibit binding of several DNA-binding proteins *in vitro* (*31, 32*). Attempts to use polyamide-intercalator conjugates to target the VEGF HRE have not been successful due to poor nuclear uptake.

The ability to target DNA sequences flanking critical protein-DNA binding sites while maintaining productive inhibition of protein-DNA binding expands the repertoire of such interactions amenable to inhibition by polyamides. In a similar approach, Kageyama *et al*. showed that polyamides targeted to sequences flanking the VEGF HRE could inhibit VEGF expression (*33*). Polyamides targeted to flanking sites have previously successfully inhibited protein-DNA binding in the cases of TATA-binding protein and LEF-1 (*34*). It should be noted that minor groove-binding polyamides and some major groove-binding proteins co-occupy DNA sequences in some cases (*35*).

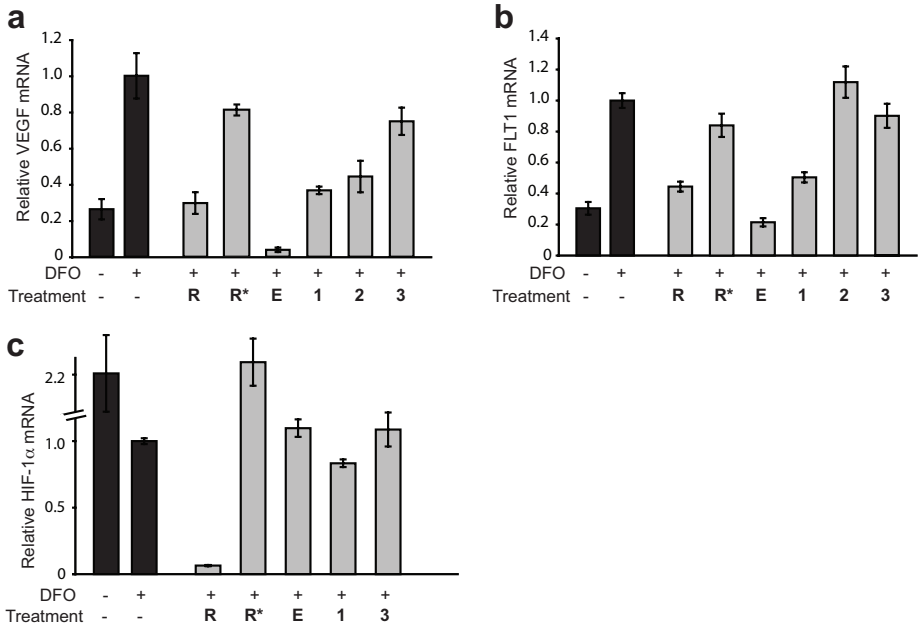


Figure 3.3. Quantitative real-time PCR measurements. a) Induction of VEGF mRNA by deferoxamine (DFO) measured by quantitative real-time PCR: HIF-1α siRNA, **R**; mismatch control siRNA, **R***; echinomycin (100 nM), **E**; and polyamides **1**, **2**, and **3** (1 μM). Treatment with siRNA, **1**, or **2** decrease VEGF mRNA levels to near non-induced levels. Echinomycin potently inhibits VEGF mRNA to below non-induced levels. Polyamide **3** has a more modest effect. b) Induction of FLT1 mRNA by DFO measured by quantitative real-time PCR: HIF-1α siRNA, **R**; mismatch control siRNA, **R***; echinomycin (100 nM), **E**; and polyamides **1**, **2**, and **3** (1 μM). The siRNA, **E**, and **1** decrease FLT1 mRNA levels. Polyamides **2** and **3** have minimal or no effect. c) Measurement of HIF-1α mRNA by quantitative real-time PCR: HIF-1α siRNA, **R**; mismatch control siRNA, **R***; echinomycin (100 nM), **E**; and polyamides **1**, **3** (1 μM). Treatment with siRNA decreases HIF-1α mRNA by more than 95%.

Microarray analysis of gene expression. One potential limitation to the use of hairpin-polyamides for gene regulation is binding site size and specificity for match versus mismatch sites, which may result in prohibitively large numbers of affected genes. To examine this, the global effects of polyamide treatment on hypoxia-induced gene expression were measured using Affymetrix Human Genome U133 Plus 2.0 Arrays containing oligonucleotide sequences representing more than 50,000 transcripts. To establish a benchmark for comparison, the effects of HIF-1α siRNA and echinomycin were also measured. Experiments were conducted in triplicate, and gene expression levels normalized to DFO-treated controls. Cells not treated with DFO were normalized to DFO-treated controls.

Polyamide **1** (1 μM) affected expression of 2,284 transcripts by more than 2-fold ($p \leq 0.01$) (Table 3.2). At the same threshold, HIF-1α siRNA affected 3,190 transcripts and echinomycin (100 nM) affected 10,906. In all cases, a majority of affected genes were down-regulated. For comparison, DFO treatment alone affected expression of 2,142 transcripts (4.6% of interrogated transcripts), with a majority up-regulated.

Table 3.2. Number of transcripts affected ($p \leq 0.01$). **R**, HIF-1α siRNA; **E**, echinomycin; **1**, polyamide.

Treatment	**-**	**R**	**E**	**1**
DFO	**-**	+	+	+
up-regulated (fold change ≥ 2.0)	662	1,380	3,480	709
down-regulated (fold change ≤ -2.0)	1,480	1,810	7,426	1,575
up-regulatd (fold change ≥ 4.0)	62	122	413	57
down-regulated (fold change ≤ -4.0)	296	356	4,133	336

Clustering analysis was performed to identify similarities in the expression profiles between the different treatments (Figure 3.4). The expression profile of cells treated with HIF-1α siRNA is similar to that of cells not treated with DFO under the conditions of the analysis; the expression profiles of echinomycin-treated and polyamide-treated cells are less similar to each other and to the other treatments. Analysis of transcripts affected by both **1** and HIF-1α siRNA shows that 395 and 150 transcripts are commonly down-

and up-regulated, respectively, at least 2-fold ($p \leq 0.01$). A similar analysis of transcripts affected by both **1** and echinomycin shows that 731 and 112 transcripts are commonly down- and up-regulated, respectively. Analysis of transcripts affected by both siRNA and echinomycin shows that 1140 and 443 transcripts are commonly down- and up-regulated, respectively. A side-by-side, genome-wide expression analysis of fluorescein-tagged analogs of polyamides **1** and **3** in DFO-induced cells was previously reported (*19*), and found that a majority of genes were uniquely affected by each polyamide, with a number of genes commonly affected, under the conditions of the experiments. It is not entirely unsurprising that there is some overlap in genes affected by polyamides targeted to different DNA sequences, given that we do not have a full understanding of all DNA sequences involved in the direct or indirect regulation of a given gene.

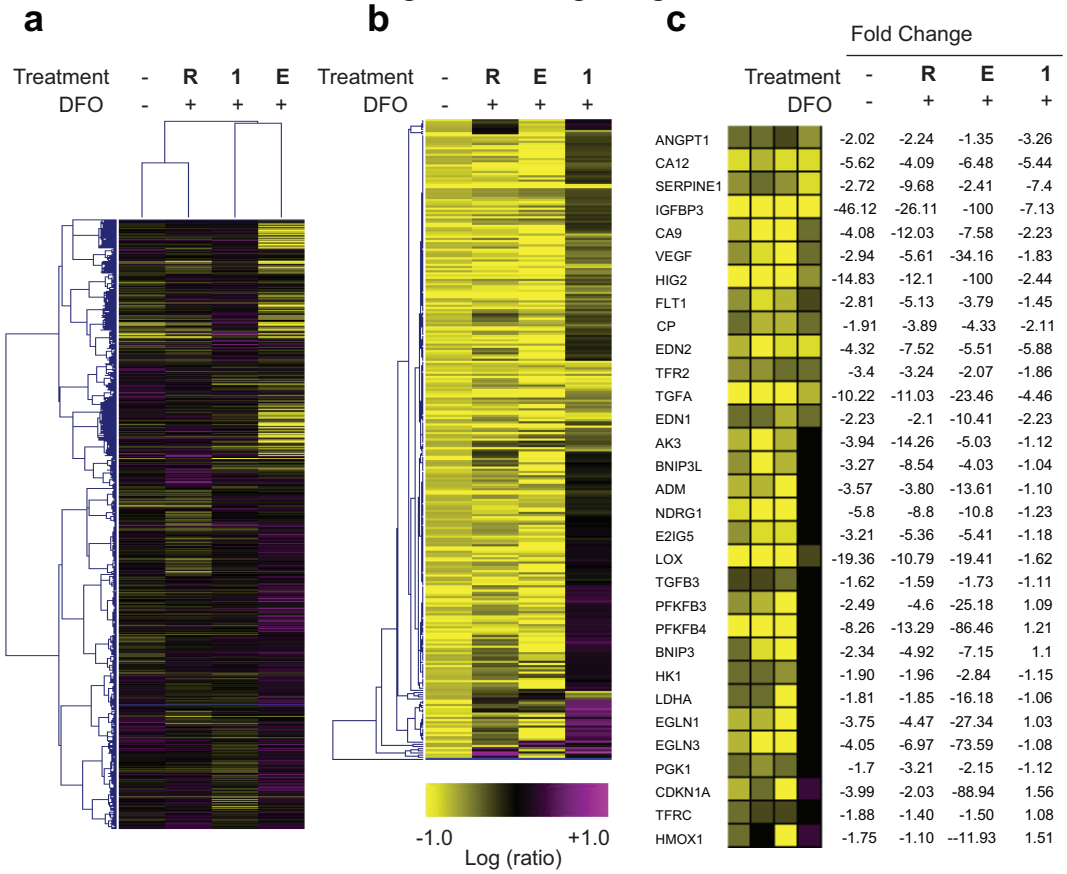


Figure 3.4. Microarray analysis of gene expression. a) Divisive clustering analysis over all interrogated transcripts for DFO-induced cells: HIF-1α siRNA, **R**; echinomycin (100 nM), **E**; and polyamides **1** (1 μM). b) Agglomerative clustering analysis over all 297 transcripts induced by DFO at least 4-fold ($p \leq 0.01$). Of this transcript set, HIF-1α siRNA inhibited 244, echinomycin inhibited 263, and polyamide **1** inhibited 69 by ≥ 2-fold ($p \leq 0.01$). c) Effects of the indicated treatments on a panel of genes previously characterized as direct targets of HIF-1 and also induced by DFO at least 1.5-fold ($p \leq 0.01$) in this experiment. Treatments reported are an error-weighted average from three experiments.

We find that DFO induced the expression of a set of 297 transcripts by at least 4-fold ($p \leq 0.01$) (Figure 3.4B). Of this set, 69 were inhibited by polyamide **1** by at least 2-fold ($p \leq 0.01$). For comparison, HIF-1α siRNA inhibited 244, and echinomycin 263 of the 297 DFO-induced transcripts. It is not known what proportion of these affected transcripts are direct HIF-1 targets. To examine more closely the effects of polyamide **1**, HIF-1α siRNA, and echinomycin on transcripts induced directly by HIF-1, we examined a limited set of 31 transcripts consisting of previously identified direct HIF-1 targets that were induced at least 1.5-fold ($p \leq 0.01$) by DFO in this experiment (Figure 3.4C) (*28, 36–45*) Nearly all 31 transcripts in this set were down-regulated by HIF-1α siRNA. In most cases, the expression was reduced to levels observed in cells untreated with DFO. Echinomycin treatment resulted in down-regulation of all 31 transcripts. For some genes, including VEGF, expression was reduced to levels far below those of the siRNA-treated cells and non-induced controls. Polyamide **1** inhibited the expression of 14 of these but displayed minimal effect on the others. Shown in Figure 3.5 is a Venn diagram representing transcripts commonly down- and up-regulated ($|\text{fold-change}| \geq 2.0$, $p \leq 0.01$) by **1** and HIF-1α siRNA, by **1** and echinomycin, and by HIF-1α siRNA and echinomycin.

The HRE sequences for these genes, where known, are displayed in Table 3.3. Quantitative real-time RT–PCR was used to confirm the effects of polyamides **1**, **3**, echinomycin, and siRNA treatments on these 11 genes. RT-PCR confirms that siRNA and echinomycin reduced expression of all genes in this set. Polyamide **1** significantly affected four genes in this set. Polyamide **3** had a modest but measurable effect on VEGF expression, but little effect on the expression of all the others. Chromatin immunoprecipitation was used to measure HIF-1 occupancy at the VEGF and carbonic anhydrase IX (CA9) HRE, which were both affected by polyamide **1**, and PGK1, which was unaffected (Figure 3.6). HIF-1α occupancy at the VEGF HRE was decreased by HIF-1α siRNA, echinomycin, and polyamide **1**, but less so by HRE-mismatch polyamide **3**. HIF-1 occupancy at the CA9 locus was also decreased by HIF-1α siRNA, echinomycin, and polyamide **1**, but

was unaffected by **3**. HIF-1 occupancy at PGK1 was markedly decreased by siRNA, but minimally affected by polyamides **1** and **3**. Surprisingly, echinomycin did not appear to affect HIF-1 occupancy at this locus. It is interesting to note that all of the genes affected by polyamide **1** (Table 3.2) have HREs that fall within the sequence 5'-(T/A)ACGTG-3', consistent with the expected DNA binding preferences for **1**.

Table 3.3. HIF-1 induced genes affected by HIF-1α siRNA (**R**), echinomycin (**E**, 100 nm), **1** and **3** (1 μM).

	HRE(s) (5' to 3')	R	E	1	3
TFRC	agcg**TACGT**Gcctc	-2.0	-2.3	1.1	1.4
PKFB3	gcgg**GACGTG**acgc gacg**CACGTG**ggca	-5.5	-82.0	1.0	1.0
LDHA	ggcg**GACGTG**cggg ctca**CACGTG**ggtt	-1.7	-4.4	1.3	1.1
BNIP3	gccg**CACGTG**ccac	-9.4	-6.0	1.3	1.0
EGLN3	gggc**TACGTG**caga	-5.3	-33.6	1.0	-1.1
EGLN1	ggtg**TACGTG**caga	-3.4	-19.8	1.1	1.0
PGK1	gtga**GACGTG**cggc tgcc**GACGTG**cgct	-5.4	-3.3	-1.2	1.0
CA9	gctg**TACGTG**catt	-89.0	-9.4	-2.1	-1.1
VEGF	tgca**TACGTG**ggct	-3.4	-34.0	-2.0	-1.3
FLT1	gaac**AACGTG**gaat	-2.2	-4.7	-2.0	-1.1
EDN1	aggc**AACGTG**cagc	-3.5	-31.0	-2.5	1.3

3.4 Discussion

In this experiment, polyamide **1** (1 μM) affected expression of 2,284 transcripts by more than 2-fold ($p \leq 0.01$), which represents less than 5% of transcripts assayed. A search of the publicly available human genome for the sequence 5'-WTWCGW-3' finds 1,876,480 potential match sites for polyamide **1.** This corresponds to an average of one binding site every 1,600 base pairs. The proportion of these sites accessible to polyamide binding in

the context of heterochromatin *in vivo* is currently unknown. Additionally, data from *in vitro* transcription experiments suggest that polyamides non-covalently bound within the coding region of a gene would not interfere with RNA polymerase activity at that locus (*34*). It would thus not be surprising if a significant fraction of polyamide-DNA binding events in a cell do not directly affect gene expression. In parallel to this, global analysis of transcription factor binding to chromatin *in vivo* has shown occupancy at promoters and enhancers without associated changes in gene expression at that locus (*46*).

Interestingly, polyamide **1** (1 μM) affected the expression of fewer genes than HIF-1α siRNA under the conditions of the experiment. A direct comparison in genomic specificity between polyamide and siRNA cannot be drawn from these limited data because a large proportion of the genes affected by siRNA are likely a result of silencing the target gene, HIF-1α, rather than off-target effects involving post-transcriptional silencing of mRNA using the RNA interference pathway (*47*). If we eliminate from the total number of transcripts affected by the HIF-1α siRNA (2-fold, $p \leq 0.01$) all transcripts affected by treatment with DFO alone (1.5-fold, $p \leq 0.01$) we are left with 1,523 affected transcripts. A similar treatment of the data for polyamide **1** results in 1,626 affected transcripts.

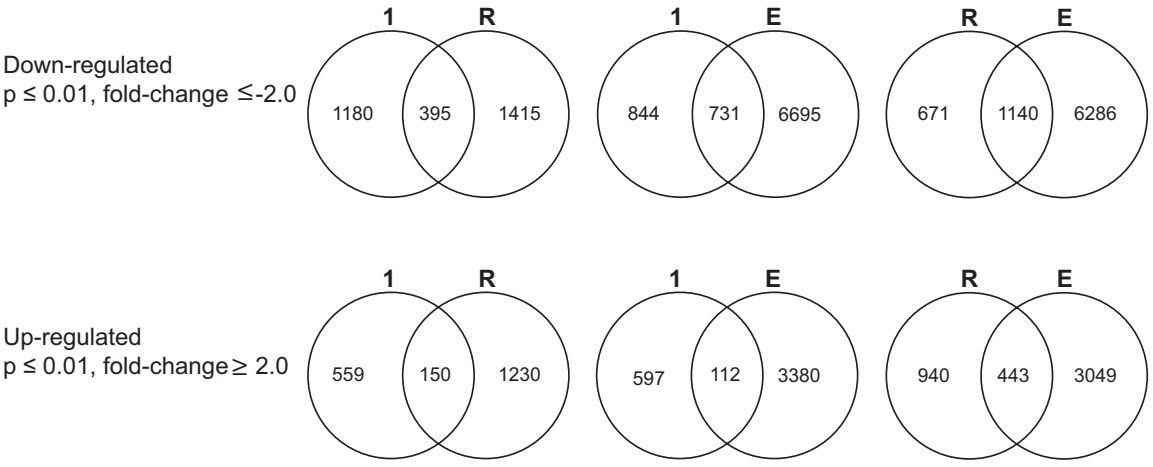


Figure 3.5: Venn diagrams representing transcripts commonly down- and up-regulated ($|\text{fold-change}| \geq 2.0$, $p \leq 0.01$) by 1 and HIF-1α siRNA, by 1 and echinomycin, and by HIF-1α siRNA and echinomycin. Numbers inside intersections represent transcripts affected by both treatments.

It should also be noted that for most HIF-1 regulated genes affected by both polyamide and siRNA, inhibition by the polyamide was more modest than siRNA, suggesting incomplete abrogation of HIF-1 DNA binding by the polyamide. Approximately 23% of the 297 transcripts induced by DFO were inhibited by polyamide **1**. For genes where the functional HRE is known, the effects of treatment with polyamide **1** or echinomycin are, thus far, consistent with the expected binding preferences of these molecules.

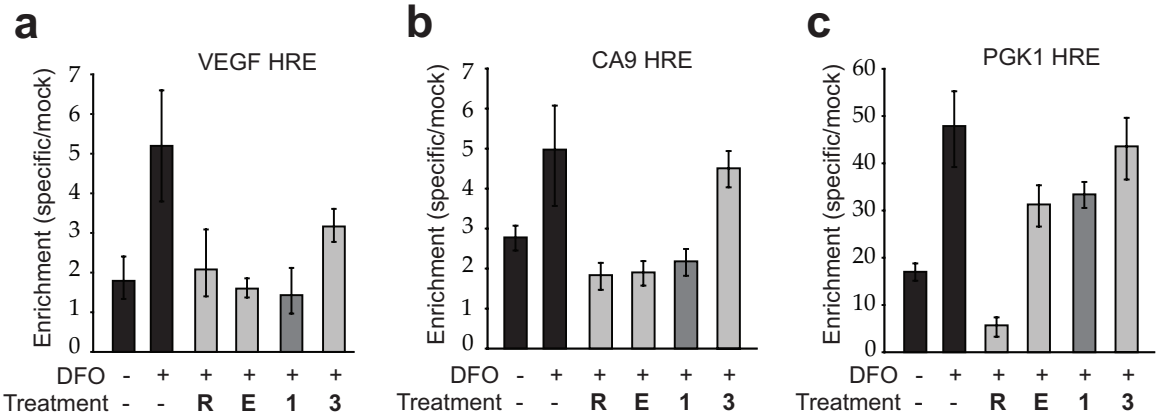


Figure 3.6. Chromatin immunoprecipitation at three HREs. a) Chromatin immunoprecipitation of HIF-1α at the VEGF HRE following DFO treatment: HIF-1α siRNA, **R**; echinomycin (100 nM), **E**; and polyamides **1** and **3** (1 μM). Occupancy is decreased in the presence of **R**, **E**, and **1**, but only modestly affected by **3**. Chromatin immunoprecipitation of HIF-1α at the CA9 HRE (b), and PGK1 HRE (c).

HIF-1 is frequently overactive in tumors, and a number of direct targets in the HIF-1 pathway have become points of clinical intervention (*48*). Bevacizumab, an anti-VEGF antibody, and sorafenib and sunitinib, tyrosine kinase inhibitors with activity against the VEGF receptors, have shown some promise in clinical trials as cancer therapeutics (*49–51*). Echinomycin had been previously brought to the clinic as a cancer therapeutic in phase I and II trials (*52*), based on observations that echinomycin exhibits potent anti-proliferative effects on several tumor-derived cell lines (*52, 53*). However, survival benefit was not established (*52*). In light of recent work by Melillo and colleagues, re-examination of the clinical use of echinomycin in tumor types expected to be highly sensitive to HIF-1 activity may be justified (*21*).

The induction of pro-angiogenic, proliferative, metastatic, and glycolytic genes by HIF-1 are established as contributing to the cancer phenotype (*11*). Genes that promote

cell death, such as BNIP3 and NIX (BNIP3L), are also induced by hypoxia through HIF-1 (*54*). In this sense, HIF-1 plays dual roles in the survival and death pathways of tumor cells (*12*). A functional separation of these targets of HIF-1 at the level of HIF-1-DNA binding might have clinical relevance (*12*). Given sufficient knowledge of the particular regulatory sequences involved, one could, in principle, design a polyamide or cocktail of polyamides to affect a selected subset of target genes in the HIF-1 pathway, making use of the programmability of polyamide recognition for targeting particular HREs and flanking sequences. The utility of polyamides as regulators of hypoxia-induced gene expression awaits continued study in small animal models of HIF-1 activity.

3.5 Methods

Synthesis of Polyamides

Polyamides were synthesized by solid-phase methods on Kaiser oxime resin (Nova Biochem), cleaved from resin with 3,3'-diamino-*N*-methyl-dipropylamine and purified by reverse-phase HPLC (*55*). Isophthalic acid was activated with PyBOP (Nova Biochem) and conjugated to polyamides as previously described (*20*). Polyamide purity and identity assessed by HPLC, UV-visible spectroscopy, and MALDI-ToF MS (Figure 3.7).

Determination of DNA-binding affinities and sequence specificities. Quantitative DNase I footprint titration experiments were used to measure the binding affinities of polyamides **1**, **2**, **3**, and echinomycin on 5' ^{32}P-labeled fragments of pGL2-VEGF-*Luc and pCSJ-FLT1* containing promoter sequences containing the HREs of VEGF and FLT1, respectively. Quantitative DNase I footprint titration experiments were conducted as reported previously (*56*).

Measurement of hypoxia-induced gene expression. U251 cells were plated in 24-well plates at a density of 20–30 x 10^3 cells per well (40–60 x 10^3 cells/ml) in RPMI (ATCC) supplemented with 5% FBS (Irvine Scientific). After 24 hours, polyamides were added to the adhered cells in cell media solutions at the appropriate concentration and incubated

with the cells for 48 hours. Hypoxic induction of gene expression was chemically induced by adding DFO to 300 μM for an additional 16 hours. When appropriate, echinomycin was added two hours prior to DFO stimulation. Isolation of RNA and subsequent cDNA synthesis were as previously described (*19*). When appropriate, HIF-1α siRNA (HIF-1α validated stealth duplex, Invitrogen) or mismatch control siRNA with similar GC content (Invitrogen) was transfected 48 hours prior to RNA isolation. Transfection of siRNA was achieved using Lipofectamine 2000 (Invitrogen) according to manufacturer's protocols. Quantitative real-time RT–PCR was performed using SYBR Green PCR Master Mix (Applied Biosystems) on an ABI 7300 instrument. Target gene mRNA was measured relative to ß-glucuronidase as an endogenous control. Primer employed were: VEGF, L 5'-AGGGCAGAATCATCACGAAG-3', R 5'-GGGTACTCC TGGAAGATGTCC-3'; ß-glucuronidase, L 5'-CTCATTTGGAATTTTGCCGATT-3', R 5'-CCGAGTGAAGATCCCCTTTTTA-3'; FLT1, L 5'-CAGCAACATGGGAAACAGAAT-3', R 5'-TAGAGTCAGCCACAACCAAGG-3'.

Chromatin Immunoprecipitation. U251 cells were plated in 15 cm diameter plates and left to attach overnight. Media, time course, DFO, polyamide, echinomycin, and siRNA treatments were as described above. After the 16 hour DFO treatment, cells were treated with 1% formaldehyde for 10 minutes. Chromatin was isolated and sheared. HIF-1α antibodies (Novus Biologicals) were used to immunoprecipitate HIF-1-bound DNA fragments. After crosslink reversal, PCRs using primers targeted to the regions of interest were used to assess enrichment of bound fragments as compared to mock-precipitated (no antibody) controls. PCRs were monitored either using SYBR Green PCR Master Mix (Applied Biosystems) on an ABI 7300 instrument, or directly visualized using gel electrophoresis. The following primers were used: VEGF, L 5'-CCTTTGGGTTTTGCCAGA-3', R 5'-CCAAGTTTGTGGAGCTGA-3'; CA9, L 5'-AAAAGGGCGCTCTGTGAGT-3', R 5'-GCTGACTGTGGGGTGTCC-3'; PGK1, L 5'-CCCCTAAGTCGGGAAGGTT-3', R 5'- GTCCGTCTGCGAGGGTACTA-3'.

Analysis of Gene Expression with Oligonucleotide Microarrays. U251 cells were plated in 12-well plates at a density of 40–60 x 10^3 cells per well. Media, time course, DFO, polyamide, echinomycin, and siRNA treatments were as described above. RNA was isolated as previously described. Further sample preparation for microarray experiments was carried out at the Millard and Muriel Jacobs Gene Expression Facility at Caltech. Labeled mRNA was hybridized to Affymetrix Human 133 arrays according to established protocols. Gene expression was analyzed using Resolver (Rosetta Biosoftware, Seattle).

Accession Codes. Data have been deposited in NCBIs Gene Expression Omnibus (GEO, http://www.ncbi.nlm.gov/geo/), accessible through GEO Series accession number GSE7535.

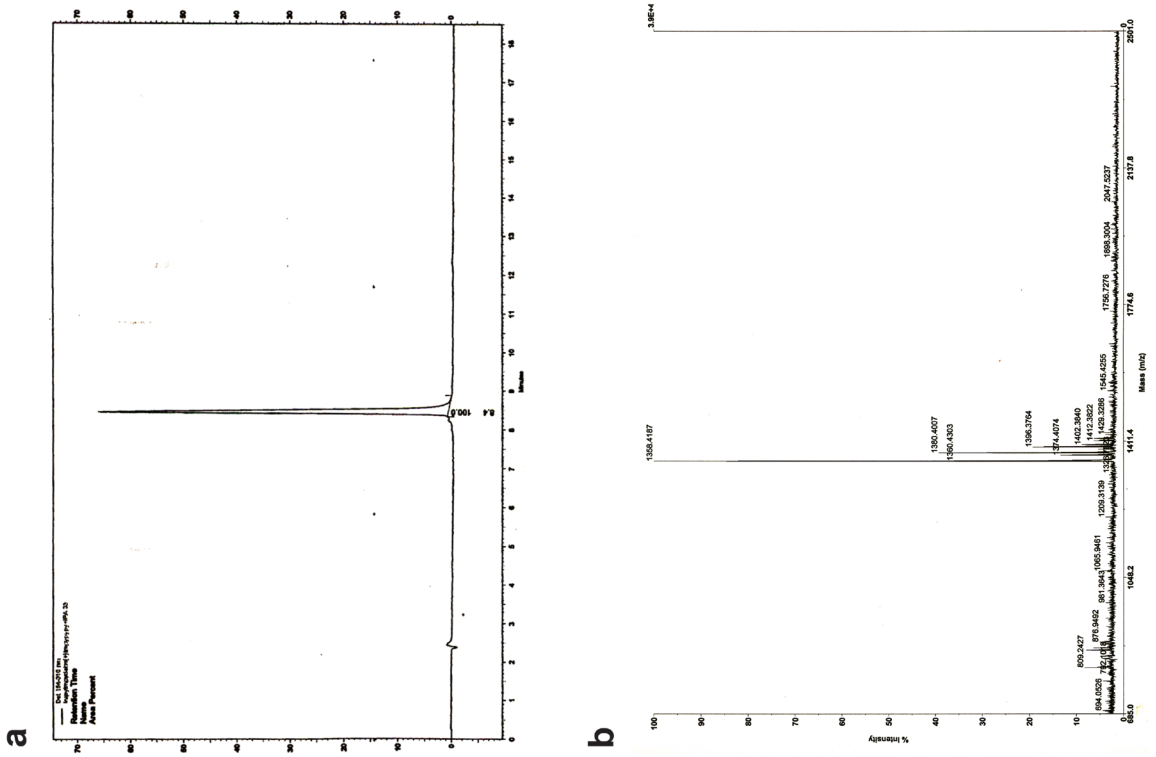


Figure 3.7: Characterization data for **3** A) Analytical HPLC trace and B) MALDI-ToF data.

References

1. Blancafort, P., Segal, D. J., and Barbas, C. F. (2004) Designing transcription factor architectures for drug discovery, *Mol. Pharmacol. 66*, 1361–1371.

2. Beerli, R. R., and Barbas, C. F. (2002) Engineering polydactyl zinc-finger transcription factors, *Nat. Biotechnol. 20*, 135–141.

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

4. Liu, P. Q., Rebar, E. J., Zhang, L., Liu, Q., Jamieson, A. C., Liang, Y. X., Qi, H., Li, P. X., Chen, B. L., et al. (2001) Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions - Activation of vascular endothelial growth factor A, *J. Biol. Chem. 276*, 11323–11334.

5. Fire, A., Xu, S. Q., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*, *Nature 391*, 806–811.

6. Meister, G., and Tuschl, T. (2004) Mechanisms of gene silencing by double-stranded RNA, *Nature 431*, 343–349.

7. Mello, C. C., and Conte, D. (2004) Revealing the world of RNA interference, *Nature 431*, 338–342.

8. Hannon, G. J., and Rossi, J. J. (2004) Unlocking the potential of the human genome with RNA interference, *Nature 431*, 371–378.

9. Dervan, P. B., and Edelson, B. S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides, *Curr. Opin. Struct. Biol. 13*, 284–299.

10. Nickols, N. G., and Dervan, P. B. (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide, *Proc. Natl. Acad. Sci. U. S. A. 104*, 10418–10423.

11. Semenza, G. L. (2003) Targeting HIF-1 for cancer therapy, *Nat. Rev. Cancer 3*, 721–732.

12. Zhou, J., Schmid, T., Schnitzer, S., and Brune, B. (2006) Tumor hypoxia and cancer progression, *Cancer Lett. 237*, 10–21.

13. Pouyssegur, J., Dayan, F., and Mazure, N. M. (2006) Hypoxia signalling in cancer and approaches to enforce tumour regression, *Nature 441*, 437–443.

14. Melillo, G. (2006) Inhibiting hypoxia-inducible factor 1 for cancer therapy, *Molecular Cancer Research 4*, 601–605.

15. Semenza, G. L. (2000) HIF-1 and human disease: one highly involved factor, *Genes Dev. 14*, 1983–1991.

16. Ivan, M., Kondo, K., Yang, H. F., Kim, W., Valiando, J., Ohh, M., Salic, A., Asara, J. M., Lane, W. S., et al. (2001) HIF alpha targeted for VHL-mediated destruction by proline hydroxylation: Implications for O-2 sensing, *Science 292*, 464–468.

17. Zimmer, M., Doucette, D., Siddiqui, N., and Iliopoulos, O. (2004) Inhibition of hypoxia-inducible factor is sufficient for growth suppression of VHL-/- tumors, *Molecular Cancer Research 2*, 89–95.

18. Kung, A. L., Zabludoff, S. D., France, D. S., Freedman, S. J., Tanner, E. A., Vieira, A., Cornell-Kennon, S., Lee, J., Wang, B. Q., et al. (2004) Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway, *Cancer Cell 6*, 33–43.

19. Olenyuk, B. Z., Zhang, G. J., Klco, J. M., Nickols, N. G., Kaelin, W. G., and Dervan, P. B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist, *Proc. Natl. Acad. Sci. U. S. A. 101*, 16768–16773.

20. Nickols, N. G., Jacobs, C. S., Farkas, M. E., and Dervan, P. B. (2007) Improved nuclear localization of DNA-binding polyamides, *Nucleic Acids Res. 35*, 363–370.

21. Kong, D. H., Park, E. J., Stephen, A. G., Calvani, M., Cardellina, J. H., Monks, A., Fisher, R. J., Shoemaker, R. H., and Melillo, G. (2005) Echinomycin, a small-molecule

inhibitor of hypoxia-inducible factor-1 DNA-binding activity, *Cancer Res. 65*, 9047–9055.

22. Van Dyke, M. M., and Dervan, P. B. (1984) Echinomycin Binding-Sites on DNA, *Science 225*, 1122–1127.

23. Kielkopf, C. L., Baird, E. E., Dervan, P. D., and Rees, D. C. (1998) Structural basis for G•C recognition in the DNA minor groove, *Nat. Struct. Biol. 5*, 104–109.

24. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands, *Nature 391*, 468–471.

25. Foister, S., Marques, M. A., Doss, R. M., and Dervan, P. B. (2003) Shape selective recognition of T•A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues, *Bioorg. Med. Chem. 11*, 4333–4340.

26. Liu, Y. X., Cox, S. R., Morita, T., and Kourembanas, S. (1995) Hypoxia Regulates Vascular Endothelial Growth-Factor Gene Expression in Endothelial Cells – Identification of a 5'-Enhancer, *Circ.Res. 77*, 638–643.

27. Forsythe, J. A., Jiang, B. H., Iyer, N. V., Agani, F., Leung, S. W., Koos, R. D., and Semenza, G. L. (1996) Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1, *Mol. Cell. Biol. 16*, 4604–4613.

28. Gerber, H. P., Condorelli, F., Park, J., and Ferrara, N. (1997) Differential transcriptional regulation of the two vascular endothelial growth factor receptor genes –Flt-1, but not Flk-1/KDR, is up-regulated by hypoxia, *J. Biol. Chem. 272*, 23659–23667.

29. Waring, M. J., and Wakelin, L. P. G. (1974) Echinomycin – Bifunctional Intercalating Antibiotic, *Nature 252*, 653–657.

30. Low, C. M. L., Drew, H. R., and Waring, M. J. (1984) Sequence-Specific Binding of Echinomycin to DNA – Evidence for Conformational-Changes Affecting Flanking Sequences, *Nucleic Acids Res. 12*, 4865–4879.

31. Fechter, E. J., and Dervan, P. B. (2003) Allosteric inhibition of protein-DNA

complexes by polyamide-intercalator conjugates, *J. Am. Chem. Soc. 125*, 8476–8485.

32. Fechter, E. J., Olenyuk, B., and Dervan, P. B. (2004) Design of a sequence-specific DNA bisintercalator, *Angew. Chem.-Int. Edit. 43*, 3591–3594.

33. Kageyama, Y., Sugiyama, H., Ayame, H., Iwai, A., Fujii, Y., Huang, L. E., Kizaka-Kondoh, S., Hiraoka, M., and Kihara, K. (2006) Suppression of VEGF transcription in renal cell carcinoma cells by pyrrole-imidazole hairpin polyamides targeting the hypoxia responsive element, *Acta Oncol. 45*, 317–324.

34. Dickinson, L. A., Gulizia, R. J., Trauger, J. W., Baird, E. E., Mosier, D. E., Gottesfeld, J. M., and Dervan, P. B. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands, *Proc. Natl. Acad. Sci. U. S. A. 95*, 12890–12895.

35. Oakley, M. G., Mrksich, M., and Dervan, P. B. (1992) Evidence That a Minor Groove-Binding Peptide and a Major Groove-Binding Protein Can Simultaneously Occupy a Common Site on DNA, *Biochemistry 31*, 10969–10975.

36. Tacchini, L., Bianchi, L., Bernelli-Zazzera, A., and Cairo, G. (1999) Transferrin receptor induction by hypoxia – HIF-1-mediated transcriptional activation and cell-specific post-transcriptional regulation, *J. Biol. Chem. 274*, 24142–24146.

37. Lok, C. N., and Ponka, P. (1999) Identification of a functional hypoxia-response element in the transferrin-receptor gene, *Exp. Hematol. 27*, 91–91.

38. Fukasawa, M., Tsuchiya, T., Takayama, E., Shinomiya, N., Uyeda, K., Sakakibara, R., and Seki, S. (2004) Identification and characterization of the hypoxia-responsive element of the human placental 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase gene, *J. Biochem. 136*, 273–277.

39. Obach, M., Navarro-Sabate, A., Caro, J., Kong, X. G., Duran, J., Gomez, M., Perales, J. C., Ventura, F., Rosa, J. L., et al. (2004) 6-phosphofructo-2-kinase (pfkfb3) gene promoter contains hypoxia-inducible factor-1 binding sites necessary for transactivation in response to hypoxia, *J. Biol. Chem. 279*, 53562–53570.

40. Semenza, G. L., Jiang, B. H., Leung, S. W., Passantino, R., Concordet, J. P., Maire,

P., and Giallongo, A. (1996) Hypoxia response elements in the aldolase A, enolase 1, and lactate dehydrogenase A gene promoters contain essential binding sites for hypoxia-inducible factor 1, *J. Biol. Chem. 271*, 32529–32537.

41. Semenza, G. L., Roth, P. H., Fang, H. M., and Wang, G. L. (1994) Transcriptional Regulation of Genes Encoding Glycolytic-Enzymes by Hypoxia-Inducible Factor-1, *J. Biol. Chem. 269*, 23757–23763.

42. Pescador, N., Cuevas, Y., Naranjo, S., Alcaide, M., Villar, D., Landazuri, M., and Del Peso, L. (2005) Identification of a functional hypoxia-responsive element that regulates the expression of the egl nine homologue 3 (egln3/phd3) gene, *Biochem. J. 390*, 189–197.

43. Willam, C., Nicholls, L. G., Ratcliffe, P. J., Pugh, C. W., and Maxwell, P. H.(2004) The prolyl hydroxylase enzymes that act as oxygen sensors regulating destruction of hypoxia-inducible factor alpha, *Advances in Enzyme Regulation, 44*, 75–92.

44. Grabmaier, K., de Weijert, M. C. A., Verhaegh, G. W., Schalken, J. A., and Oosterwijk, E. (2004) Strict regulation of CAIX(G250/MN) by HIF-1 alpha in clear cell renal cell carcinoma, *Oncogene 23*, 5624–5631.

45. Hu, J., Discher, D. J., Bishopric, N. H., and Webster, K. A. (1998) Hypoxia regulates expression of the endothelin-1 gene through a proximal hypoxia-inducible factor-1 binding site on the antisense strand, *Biochem. Biophys. Res. Commun. 245*, 894–899.

46. Zhang, X. M., Odom, D. T., Koo, S. H., Conkright, M. D., Canettieri, G., Best, J., Chen, H. M., Jenner, R., Herbolsheimer, E., et al. (2005) Genome-wide analysis of cAMP-response element binding protein occupancy, phosphorylation, and target gene activation in human tissues, *Proc. Natl. Acad. Sci. U. S. A. 102*, 4459–4464.

47. Jackson, A. L., Bartz, S. R., Schelter, J., Kobayashi, S. V., Burchard, J., Mao, M., Li, B., Cavet, G., and Linsley, P. S. (2003) Expression profiling reveals off-target gene regulation by RNAi, *Nat. Biotechnol. 21*, 635–637.

48. Cohen, H. T., and McGovern, F. J. (2005) Renal-cell carcinoma, *N. Engl. J. Med. 353*, 2477–2490.

49. Escudier, B., Eisen, T., Stadler, W. M., Szczylik, C., Oudard, S., Siebels, M., Negrier, S., Chevreau, C., Solska, E., et al. (2007) Sorafenib in advanced clear-cell renal-cell carcinoma, *N. Engl. J. Med. 356*, 125–134.

50. Motzer, R. J., Hutson, T. E., Tomczak, P., Michaelson, M. D., Bukowski, R. M., Rixe, O., Oudard, S., Negrier, S., Szczylik, C., et al. (2007) Sunitinib versus interferon alfa in metastatic renal-cell carcinoma, *N. Engl. J. Med. 356*, 115–124.

51. Yang, J. C., Haworth, L., Sherry, R. M., Hwu, P., Schwartzentruber, D. J., Topalian, S. L., Steinberg, S. M., Chen, H. X., and Rosenberg, S. A. (2003) A randomized trial of bevacizumab, an anti-vascular endothelial growth factor antibody, for metastatic renal cancer, *N. Engl. J. Med. 349*, 427–434.

52. Foster, B. J., Clagettcarr, K., Shoemaker, D. D., Suffness, M., Plowman, J., Trissel, L. A., Grieshaber, C. K., and Leylandjones, B. (1985) Echinomycin – the 1st Bifunctional Intercalating Agent in Clinical Trials, *Invest. New Drugs 3*, 403–410.

53. May, L. G., Madine, M. A., and Waring, M. J. (2004) Echinomycin inhibits chromosomal DNA replication and embryonic development in vertebrates, *Nucleic Acids Res. 32*, 65–72.

54. Kothari, S., Cizeau, J., McMillan-Ward, E., Israels, S. J., Bailes, M., Ens, K., Kirshenbaum, L. A., and Gibson, S. B. (2003) BNIP3 plays a role in hypoxic cell death in human epithelial cells that is inhibited by growth factors EGF and IGF, *Oncogene 22*, 4734–4744.

55. Belitsky, J. M., Nguyen, D. H., Wurtz, N. R., and Dervan, P. B. (2002) Solid-phase synthesis of DNA binding polyamides on oxime resin, *Bioorg. Med. Chem. 10*, 2767–2774.

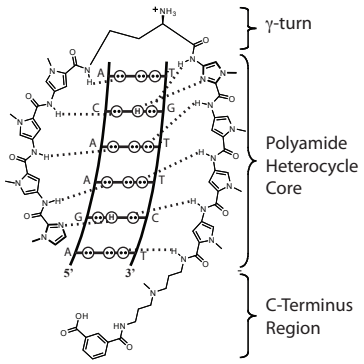
56. Trauger, J. W., and Dervan, P. B. (2001) Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes, *Methods Enzymol. 340*, 450–466.

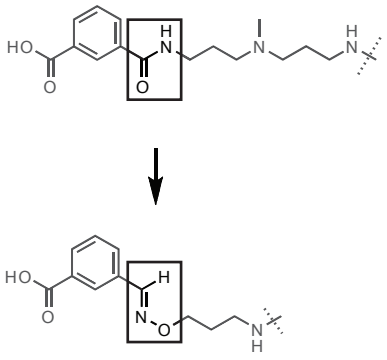
Chapter 4

Effect of Linker, Linkage, and Tail Modifications on Biological Activity of Pyrrole/Imidazole Polyamides

Abstract

Pyrrole/imidazole (Py/Im) hairpin polyamides are a class of small molecule DNA-minor groove binding compounds that have been shown to regulate endogenous gene expression *in vivo*. Gene regulation by polyamides requires efficient cellular uptake and nuclear localization properties for candidate compounds. To further optimize Py-Im polyamides for enhanced potency *in vivo*, a library of polyamides possessing various tail, linker and linkage moieties was synthesized and tested. Comparison of polyamide biological activity in two cell lines revealed tolerance for structural modifications and agreement in activity trends between cell lines. The use of an oxime linkage between the polyamide terminus and an aromatic functionality on the tail resulted in a ~20-fold increase in the potency of polyamides targeted to the Androgen Response Element (ARE) in LNCaP cells by measuring AR-activated PSA expression.





Introduction

Small molecules capable of regulating endogenous gene expression in a specific fashion could find numerous applications in molecular biology and human medicine.[1, 2] DNA-binding polyamides comprised of N-methylimidazole (Im) and N-methylpyrrole (Py) are a class of programmable small molecules that bind the minor groove of DNA in a sequence specific fashion and have been employed as regulators of gene expression through DNA-binding and modulation of the transcription factor-DNA interface.[3] Hairpin polyamides are comprised of two Im/Py strands linked via an aliphatic linker, with sequence specificity resulting in a programmed fashion from side-by-side heterocyclic amino acid pairings: Im/Py recognizes G•C over C•G; Py/Py is degenerate for A•T and T•A, while the 3-chlorothiophene N-terminus cap/Py recognizes T•A over A•T.[4-6] Eight-ring Im/Py polyamides have been shown to exhibit specific binding with binding affinities comparable to natural DNA-binding transcription factors.[7] Polyamide regulation of gene expression presumably occurs through either direct steric blockade of transcription factor binding sites or allosteric modification of DNA topology. While the sequence specificity and DNA-binding affinity of these small molecules is well studied, investigations into their potential therapeutic applications are active areas of investigation. We have recently shown that the biological activity of Im/Py polyamides can be improved via the incorporation of an isophthalic acid (IPA) group at the polyamide C-terminus, and we present here further structural modifications to polyamide tail and linker groups that increase even further their biological activity.[8]

Early *in vivo* gene regulation studies relied on fluorescein-polyamide conjugates as cell permeable compounds.[9-13] Subsequent studies in this arena aimed at eliminating the fluorescent tag utilized quantitative real-time RT-PCR to measure mRNA levels of an endogenous inducible gene as a biological readout of polyamide nuclear entry and DNA-binding. A focused library of minimized FITC analogues identified a stable, inexpensive replacement for fluorescein in the form of the IPA tail, which rivaled the activity of the

original FITC-labeled polyamide in HeLa (cervical cancer) and U251 (glioma) cells.[8] Subsequent work in other cell lines aimed at regulating different target genes has validated this C-terminus modification as a positive contributor to biological activity with negligible detrimental impact on polyamide sequence specificity or binding affinity.[14, 15]

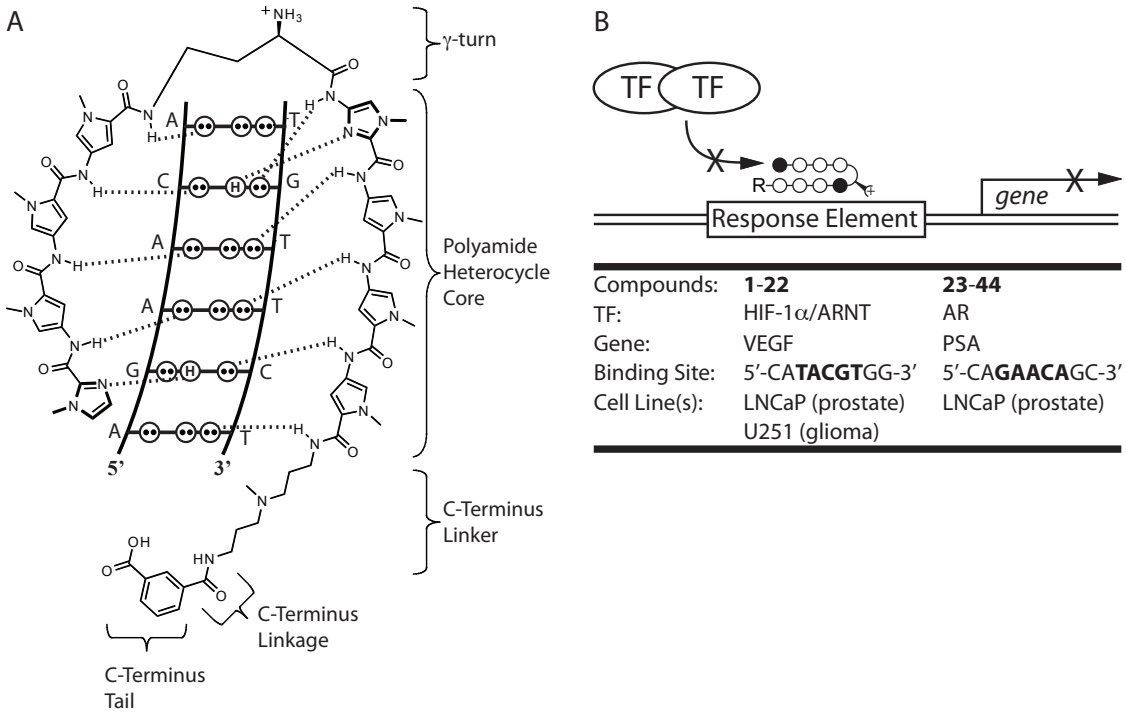


Figure 4.1. Project overview. A) Schematic illustration of polyamide-DNA recognition through formation of hydrogen bonds with the floor of the DNA minor groove. The γ-turn, polyamide heterocycle core, C-terminus linker, C-terminus linkage, and C-terminus tail of the polyamide are indicated by brackets and labeled for clarity. B) Schematic illustration of the mechanism by which polyamides affect gene expression. Sequence-specific binding of a polyamide core to a gene Response Element (RE) blocks Transcription Factor (TF) binding to the RE, thus blocking up-regulation of gene product expression under inducing conditions. Relative mRNA expression levels under inducing conditions were employed as a biological readout of polyamide cell uptake nuclear localization. Displayed in tabular form are the relevant transcription factors, gene products, DNA RE binding sequence, and cell lines for each set of match and mismatch polyamides.

The discovery of new moieties that facilitate polyamide cell uptake and nuclear localization for gene regulation studies is an area of keen interest as we aim to enhance further polyamide efficacy as we transition our studies towards animal disease models. The present study investigates the influence of linker, chemical linkage, and target cell line on polyamide nuclear localization utilizing a small library of C-terminus-modified polyamides (Figure 4.1). Quantitative real-time RT-PCR of relative mRNA expression levels in cells treated with polyamides was used as a measure of biological efficacy and the criterion for ranking

members of the polyamide library. Polyamides were targeted either to the vascular endothelial growth factor (VEGF) promoter in U251 (glioma) and LNCaP (prostate) cells, or to the prostate specific antigen (PSA) promoter in LNCaP cells (Figure 4.1B), as these two transcription-factor systems have been exploited for regulating such medically relevant genes by our group.[13-15] Although a variety of structural parameters were evaluated, it was discovered that a C-3 aliphatic linker tethered to an aromatic group through an oxime linkage resulted in enhanced polyamide potency. This motif has been employed recently by our group for the facile ^{18}F-labeling of polyamides for *in vivo* biodistribution by positron emission tomography.[16] This particular modification, as well as a variety of structural perturbations to the polyamide C-terminus, is discussed in detail.

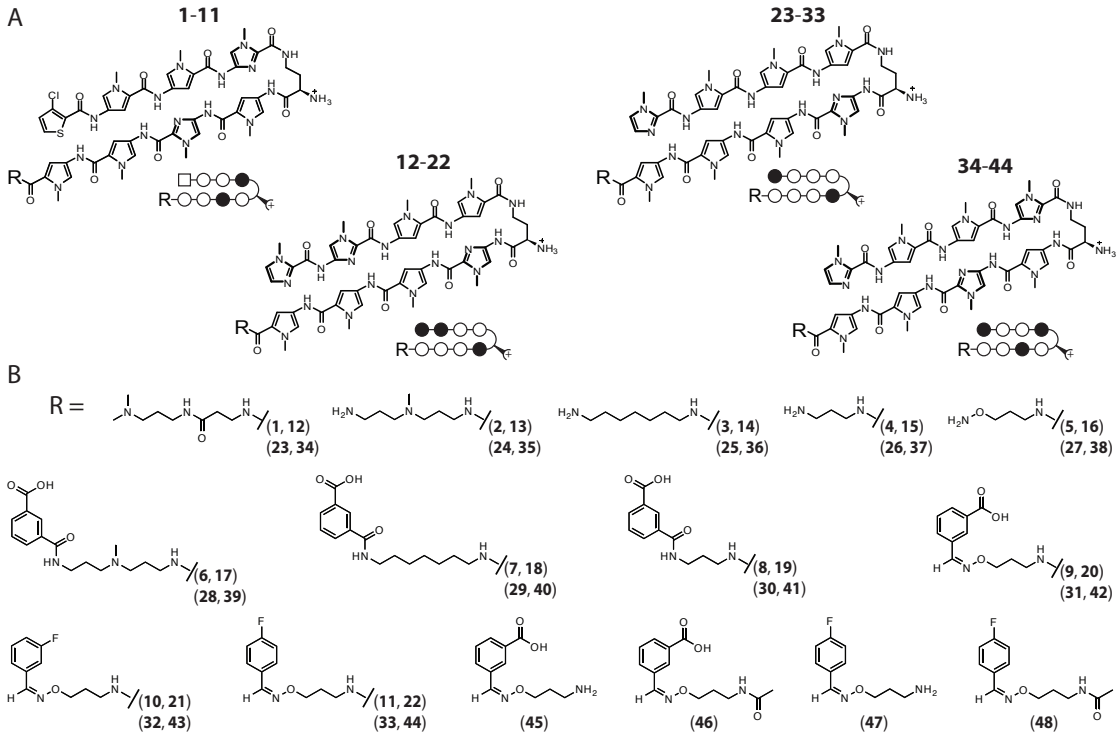


Figure 4.2. Schematic illustration of the polyamide cores and tail groups employed. A) At left are given the HRE match (**1**-**11**) and HRE mismatch (**12**-**22**) cores employed in U251 and LNCaP cells, and on the right the ARE match (**23**-**33**) and ARE mismatch (**34**-**44**) cores employed in LNCaP cells. Below the chemical structure is a schematic "ball-and-stick" shorthand representation of each core, in which pyrrole (Py) is symbolized by an open circle, imidazole (Im) by a filled circle, and chlorothiophene cap by an open square. B) The chemical structures of the tail groups "R" for compounds **1**-**44** are shown. Structures **45**-**48** are the "core-less" oxime-linked control compounds.

Results

Quantitative Real-Time RT-PCR data for Polyamide-Treated U251 and LNCaP cells

The biological activity against DFO-induced VEGF expression of HRE-targeted match (**6**-**11**) and mismatch (**12**-**22**) polyamides was studied in U251 and LNCaP cell culture, and of ARE-targeted match (**28**-**33**) and mismatch (**39**-**44**) polyamides against DHT-induced PSA expression in LNCaP cells. A schematic representation of polyamide-DNA binding and an overview of the project is shown in Figure 4.1, structures of the polyamide cores and tails are shown in Figure 4.2, and results for cell culture experiments in Figures 4.3-4.7. Polyamides that regulate VEGF expression under hypoxic (inducing) conditions target the hypoxia response element (HRE) 5'-A**TACGT**G-3' of the VEGF promoter (compounds **1**-**11**), and those that target PSA expression following dihydrotesterone (DHT) induction target the androgen response element (ARE) 5'-A**GAACA**G-3' of the PSA promoter (compounds **23**-**33**). For each gene, a set of mismatch control polyamides was included; mismatch polyamides are designed not to bind to the HRE or ARE, though some degree of biological activity is not unexpected, presumably due to off-target effects. The mismatch polyamides for the HRE series are **12**-**22** (target 5'-WGGWCW-3' sequences), and for the ARE series are **34**-**44** (target 5'-WGWCGW-3' sequences). These compounds are linked either through an amide linkage (**6**-**8**, **17**-**19**, **28**-**30** and **39**-**41**) or an oxime linkage (**9**-**11** and **20**-**22, 31**-**33** and **42**-**44**) to a C-terminal tail group. For each linker group, unconjugated free amine control compounds lacking a tail moiety (**2**-**5** and **13**-**16**, **24**-**27** and **35**-**38** (match and mismatch polyamides, respectively)) were included: due to its historical use and known modest level of biological activity, βDp tail (**1**, **12, 23** and **34**) polyamides were included for further comparison. Polyamide-untreated uninduced and polyamide-untreated induced conditions were used as controls, and VEGF or PSA mRNA expression levels in treated cells were normalized to those in the untreated, induced controls. Induced and uninduced DMSO-treated control conditions were included to rule out vehicle (DMSO) as the source of relative mRNA expression changes (DMSO concentration ≤ 0.1% for all biological

experiments)[17]. Unless otherwise noted, U251 cells were dosed with polyamides at 1, 0.2 and 0.02 μM, and LNCaP cells were dosed at 10, 2 and 0.2 μM; for both the VEGF and PSA systems, cells were dosed with control unconjugated control polyamides only at the highest concentration.

Biological Activity of HRE-targeted polyamides in U251 and LNCaP cells The data obtained for the amide-linked HRE series in U251 and LNCaP are shown in Figures 4.3A and 3B, respectively. The biological activity measured for **6** is comparable to previously obtained data for **6** at 1 μM in U251 cells.[8, 15] The general trend in this cell line indicates that the unconjugated control compounds lacking the IPA tail moiety all exhibited a modest degree of activity, but that activity is improved through IPA conjugation. IPA conjugation resulted in a greater increase in activity for compounds bearing the triamine linker than for compounds bearing the C7-linker or C3-linker. The βDp tail was the least active of all tails in this series. In LNCaP cells, the unconjugated controls were more active than in U251 cells, with the unconjugated parent compound **4** as active as IPA conjugates **6**-**8**. Conjugation of the IPA tail increased the activity of the triamine and C7-linked compounds, but the difference in activity between unconjugated parents and the IPA derivatives was more modest than in U251 cells. The βDp tail polyamides (**1** and **12**) were more active in LNCaP than U251.

The activity of the HRE oxime-linked series was measured in U251 and LNCaP, and the results are shown in Figure 4.6A and 4.6B, respectively. Compounds **9**-**11** were active in U251 cells under hypoxic conditions: VEGF mRNA expression was downregulated ~60% in U251 cells treated with 1 μM **9**-**11**, which is comparable to the activity of **6** at 1μM in U251 (Figure 4.3A). Mismatch polyamide compounds **20**-**22** were as active in U251 cells as their respective match polyamides **9**-**11.** In LNCaP cells, compounds **9**-**11** were also active: VEGF mRNA expression levels measured for cells treated with 10 μM **9**-**11** were downregulated ~80%. Mismatch oxime-linked polyamides were significantly active in LNCaP cells.

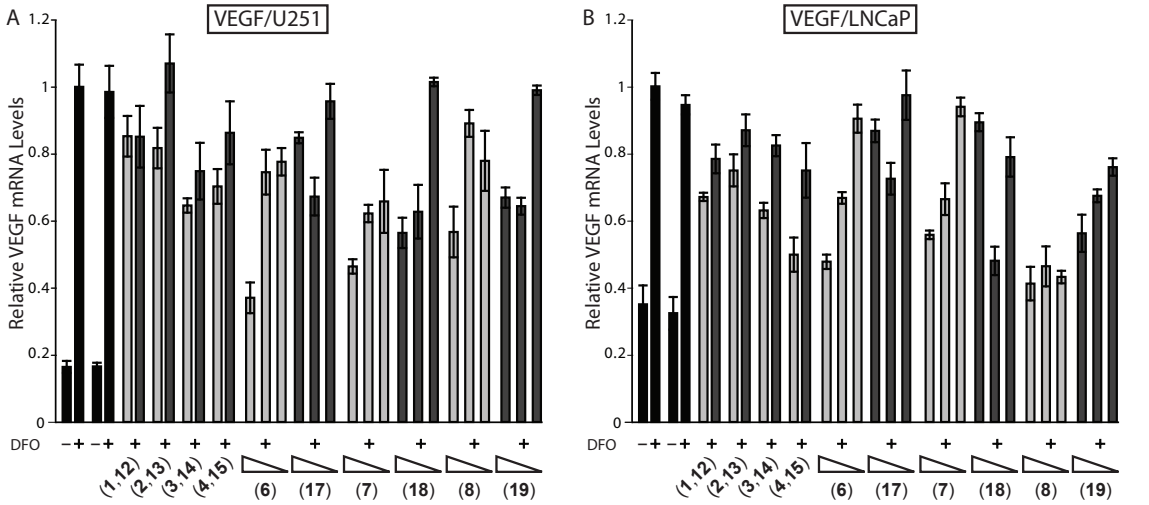


Figure 4.3. Quantitative real-time RT-PCR data showing the effect of HRE-targeted match compounds **1**-**4**, **6**-**8** and their mismatch congeners **12**-**15**, **17**-**19** on DFO-induced VEGF expression in cultured cells. Induced and uniduced control conditions are indicated by black bars, match core compounds by light grey bars, and mismatch core compounds by dark grey bars. Uninduced and induced control compounds without DMSO are on the left, and with 0.1% DMSO on the right. Errors shown are the fractional standard deviation. A) Expression of VEGF in cultured **U251** cells under DFO-induced hypoxic conditions of match core control compounds **1**-**4** and mismatch core control compounds **12**-**15** at 1 μM concentration. Cells were treated with match compounds **6**-**8** and respective mismatch compounds **17**-**19** as 1, 0.2 and 0.02 μM cell culture media solutions. B) Expression of VEGF in cultured **LNCaP** cells under DFO-induced hypoxic conditions of match core control compounds **1**-**4** and mismatch core control compounds **12**-**15** at 10 μM. Cells were treated with match compounds **6**-**8** and respective mismatch compounds **17**-**19** as 10, 2 and 0.2 μM cell culture media solutions.

Biological Activity of ARE-targeted polyamides in LNCaP cells The relative PSA mRNA levels in LNCaP cells induced with 1 nM DHT and dosed with polyamides **23**-**44** were measured (Figure 4.4 and Figure 4.6C.) In the amide series, the unconjugated free amine control compounds **23**-**26** were more active than mismatch controls **34**-**37**. Amide-linked ARE compounds **28**-**30** were appreciably active, and a modest linker effect was observed (C7 > C3 > triamine); mismatch compounds **39**-**41** were less active than their match mates. The difference in activity between each match **28**-**30** and respective mismatch **39**-**41** was greater than measured for VEGF regulation in U251 and LNCaP cells.

The activity of the oxime-linked series **31**-**33** and **42**-**44** in LNCaP cells was measured at 5, 1 and 0.1 μM and unconjugated controls **27** and **38** at 5 μM; the data are shown in Figure 4.6C. A significant increase in biological activity was measured for match polyamides **31**-**33**: all three compounds decreased PSA mRNA levels at 5 μM below untreated, uninduced levels. PSA mRNA levels were reduced to untreated, uninduced levels at 1 μM **31** and **33.**

Mismatch polyamides **42**-**44** were less active than their respective match polyamides **31**-**33**. The hydroxylamine match control compound **27** downregulated PSA mRNA levels by ~50%, while the hydroxylamine mismatch compound **38** showed very little activity. LNCaP cells treated with 5 μM **31**-**33** showed noticeable cytotoxic response upon incubation for 48 hours, an effect that was more pronounced for **32** and **33** as compared to **31**.

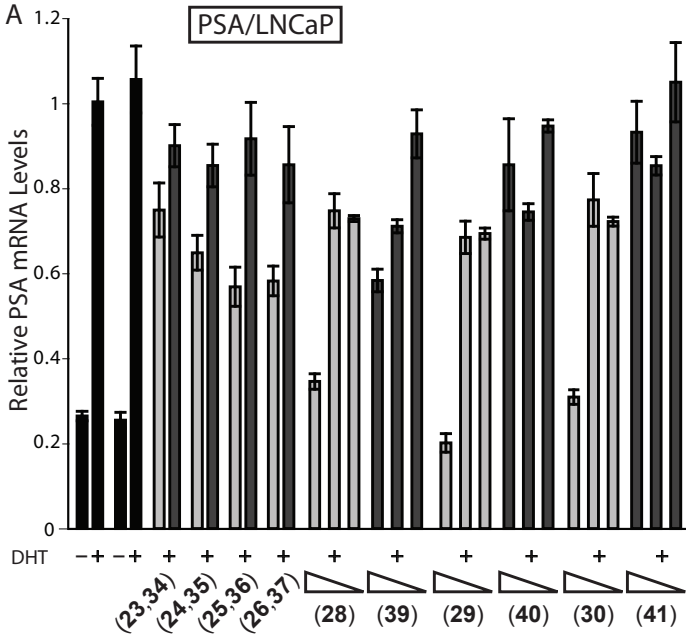


Figure 4.4. Quantitative real-time RT-PCR data measuring the effect on relative PSA mRNA expression levels under DHT induction in **LNCaP** cells dosed with compounds **23**-**26**, **28**-**30** and their mismatch congeners **34**-**37** and **39**-**41**. Induced and uniduced control conditions are indicated by black bars, match core compounds by light grey bars, and mismatch core compounds by dark grey bars. Uninduced and induced control compounds without DMSO are on the left, and with 0.1% DMSO on the right. Errors shown are the fractional standard deviation. Cells were treated with match core control compounds **23**-**26** and mismatch core control compounds **34**-**37** as cell media solutions 5 μM in polyamide. Cells were treated with match compounds **28**-**30** and respective mismatch compounds **39**-**41** as 5, 1, and 0.1 μM solutions in cell culture media.

DNA Melting Temperature (Tm) The effect of polyamide binding on 14-mer oligonucleotide melting temperature (Tm) was measured (Table SI 2) to determine whether differences in biological activity by quantitative real-time RT-PCR correlate to differences in binding affinities.[18] Tm measurements require minimal materials and present a rapid method for determining the rank-order of binding affinities, although a quantitative K_a value is not obtained. The 14-mer oligonucleotide sequences (Materials and Methods) are based on the HRE of VEGF and ARE of PSA, respectively. For both the HRE and ARE 14-mers, the melting temperature increased appreciably upon match compound binding (**1**-**11**

and **23**-**33**, respectively). Tms also generally increased upon mismatch binding (**12**-**22** and **34**-**44**, respectively), but to a more modest degree. Melting temperatures of the free amine parent compounds were higher than those of the IPA conjugates, with the exception of **36** versus **40**. The ΔTm values calculated for match and mismatch compounds ranged from a 0.1°C to 14.8°C increase for the HRE polyamide series, and from a -0.9°C drop to 17.3°C increase for the ARE series.

Table 4.1. Melting temperatures and ΔTm values measured for **1**-**44** on 14-mer oligonucleotides containing the HRE or ARE sequences (oligonucleotide sequences in Materials and Methods). All values reported are averages of at least three melting temperature experiments (standard deviations indicated in parentheses.)

	5′-G T G C **A T A C G T** G G G C-3′ 3′-C A C G **T A T G C A** C C C G-5′			5′-T T G C **A G A A C A** G C A A-3′ 3′-A A C G **T C T T G T** C G T T-5′	
Compound Number	Tm (°C)	ΔTm (°C)	**Compound Number**	Tm (°C)	ΔTm (°C)
DNA Only	64.7 +/- 0.2		**DNA Only**	60.8 +/- 0.2	
1	77.1 +/- 0.2	12.3	**23**	75.1 +/- 0.3	14.3
2	79.5 +/- 0.2	14.7	**24**	78.1 +/- 0.2	17.3
3	77.4 +/- 0.2	12.7	**25**	75.1 +/- 0.2	14.3
4	78.2 +/- 0.2	13.5	**26**	77.0 +/- 0.2	16.2
5	75.3 +/- 0.1	10.6	**27**	72.0 +/- 0.3	11.2
6	76.4 +/- 0.2	11.7	**28**	73.6 +/- 0.1	12.8
7	73.7+/- 0.0	9.0	**29**	71.2 +/- 0.2	10.4
8	74.2 +/- 0.2	9.5	**30**	71.9 +/- 0.2	11.1
9	76.1 +/- 0.2	11.4	**31**	72.3 +/- 0.2	11.5
10	74.1 +/- 0.2	9.4	**32**	73.3 +/- 0.1	12.5
11	75.2+/- 0.2	10.5	**33**	73.7 +/- 0.3	12.9
12	66.3 +/- 0.2	1.6	**34**	64.4 +/- 0.2	3.6
13	67.8 +/- 0.2	3.1	**35**	67.2 +/- 0.1	6.4
14	66.8 +/- 0.2	2.1	**36**	59.9 +/- 0.2	-0.9
15	68.9 +/- 0.0	4.1	**37**	66.1 +/- 0.2	5.3
16	64.8 +/- 0.2	0.1	**38**	62.8 +/- 0.2	2.0
17	65.6 +/- 0.2	0.9	**39**	64.0 +/- 0.2	3.2
18	65.4 +/- 0.2	0.7	**40**	63.4 +/- 0.2	2.6
19	65.2 +/- 0.2	0.5	**41**	62.3 +/- 0.2	1.5
20	65.6 +/- 0.3	0.8	**42**	63.4 +/- 0.0	2.6
21	65.7 +/- 0.2	0.9	**43**	64.8 +/- 0.1	4.0
22	65.5 +/- 0.2	0.8	**44**	64.2 +/- 0.2	3.4

ΔTm = Tm(DNA+polyamide) – Tm(DNA only).

Cell Viability Effects of ARE-targeted match compounds 28-33 and 39-44 in LNCaP cells LNCaP cells were dosed with each compound at concentrations spanning 30 to 0.01 μM, and metabolic activity measured with a standard WST-1 colorimetric assay. Metabolic activity decreased with increasing concentrations of **28**-**30**, but did not significantly surpass 50% viability; therefore IC_{50} values for growth inhibition could not be calculated. Cells treated with the mismatch mates (**39**-**41**) had little effect on cell growth. The three oxime-linked match compounds had measurable IC_{50} values for growth inhibition: **31** = 1.7 ± (0.3) μM, **32** = 0.6 ± (0.1) μM, and **33** =0.4 ± (0.0) μM. The mismatch oximes displayed a similar pattern; the growth inhibition of the fluorine-bearing oximes **43** and **44** was more pronounced than the 3-carboxy tail mismatch **42**, but the effect on cell growth of the mismatch polyamides was less pronounced than that of the match polyamides. The cell growth data for **28**-**33** and **39**-**44** in LNCaP are shown in Figure 4.6 and Table 4.2.

Cell Viability Effects of Control compounds 45-48 in U251 and LNCaP cells IC_{50} values for growth inhibition in U251 cells treated with oxime-linked polyamide core-less compounds **45**-**48** at a range of 10–0.002 μM and LNCaP cells at 50–0.02 μM were measured by a standard WST-1 colorimetric assay. Negligible cell viability effects were measured in either U251 or LNCaP cells after a 48-hour treatment course with **45**-**48** at the concentrations used. These data are shown in Figure 4.5.

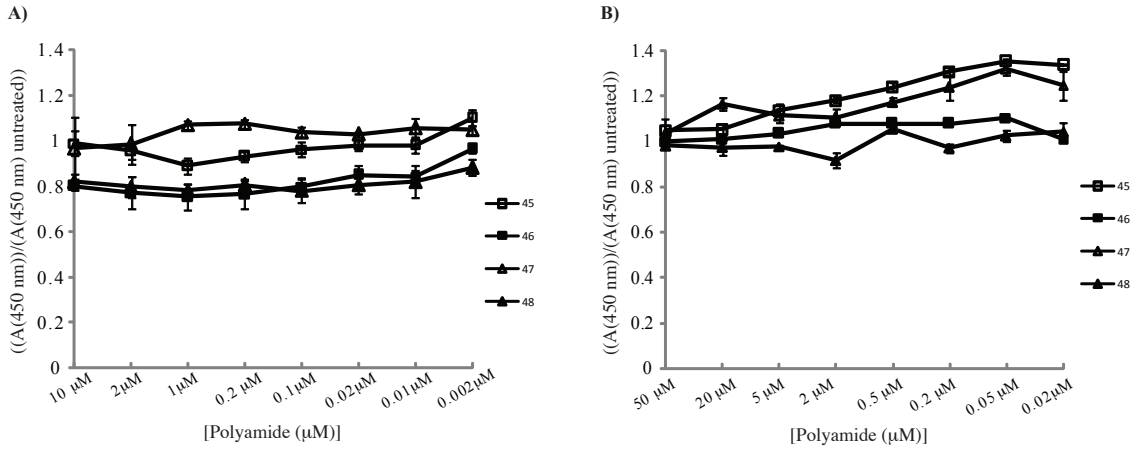


Figure 4.5. Cell viability data for **45**-**48** in U251 and LNCaP cells, performed according to the protocol provided in main text. A) Cell viability data for **45**-**48** in U251 cells at 10, 2, 1, 0.2, 0.1, 0.02, 0.01 and 0.002 μM. B) Cell viability data for **45**-**48** in LNCaP cells at 50, 20, 5, 2, 0.5, 0.2, 0.05 and 0.02 μM.

Gene Regulation IC_{50} Values for ARE-targeted oxime-linked compounds in LNCaP cells IC_{50} values for PSA mRNA expression levels were measured for **28**, **31** and **33** by quantitative real-time RT-PCR in LNCaP cells; **28** was used as a benchmark, and **31** and **33** were selected on the basis of their high biological activity and to provide variety in the tail substitution. Cells were dosed with **28** at a range of 30–0.1 μM and **31** and **33** at 10–0.03 μM. Isotherms were generated for each of three independently performed real-time quantitative RT-PCR experiments in LNCaP cells for **28**, **31** and **33**. An IC_{50} value was generated for each run, and the IC_{50} value given in Table 4.1 for each compound is the average value of three independent measurements (error given is the standard error of the mean (SEM) between the three independent runs). Isotherms (Figure 4.7B) were generated from the average value of the independent runs at each concentration (error bars are the SEM for each data point). Average IC_{50} values for gene regulation were as follows: **28** = 6.0 (± 2.6) μM, **31** = 0.28 (± 0.1) μM, and **33** = 0.6 (± 0.3) μM.

Discussion

A focused library of polyamides was synthesized with different modifications at the C-terminus, in which the linker, linkage and tail group were varied and tested in two separate cell lines for gene regulation activity of two different target genes. Prior to the current study, no direct comparison of the gene regulatory activity of a polyamide series between multiple cell lines had been performed. Measurement of VEGF mRNA expression levels in both U251 and LNCaP for the entire amide- and oxime-linked series was performed in an effort to evaluate whether modifications in tail or linker structures can have differing effects on biological activity in a cell-line dependent fashion. The polyamide dose required to effect gene regulation activity was found to be cell-line dependent, but the trend in relative activities of polyamides did not vary greatly between different cell lines. Two encouraging results of this study were the degree to which changes in the linker composition are well tolerated, and the significant increase in biological activity achieved upon replacement of an amide linkage with an oxime linkage between the linker and tail groups.

The biological activity of a series of polyamides targeted to the HRE of VEGF was measured under hypoxic (inducing) conditions in two different cell lines, U251 (glioma) and LNCaP (prostate cancer) lines. Polyamide regulation of DFO-induced VEGF expression in LNCaP cells has not previously been published. In the HRE-targeted amine series (**1**-**8** and **12**-**19**), a slight linker effect was observed, but varied between the two cell lines. Overall, changes in the linker group did not dramatically affect biological activity in either cell line. The most pronounced effect was conjugation of the IPA tail, which resulted in an increase for each linker group in both cell lines, with the exception of the C3 linker in LNCaP cells. IPA conjugation affected biological activity of the triamine-linked compounds most significantly in both U251 and LNCaP.

The amide-linked ARE series was assayed for activity against DHT-induced, AR-regulated PSA expression in LNCaP cells. Removal of the tertiary amine in the linker resulted in an increase in biological activity, as seen for **29** vs. **28**, while reduction in linker length from a C7 to a C3 linker led to a modest loss of activity (**29** vs. **30**). Although the mechanism by which mismatch polyamides affect gene expression is unknown, removal of the tertiary amine reduces mismatch activity (e.g. **39** is more active that **40** and **41**); it has previously been observed by DNase I titration footprint experiments that polyamides bearing multiple positive charges bind nonspecifically to DNA, and it may be here that the presence of the positive charge reduces the off-rate of DNA-bound polyamide and results in less specific binding compared to a charge-neutral compound. These results again indicate that there is tolerance of linker group modification, and that the ARE system in LNCaP cells benefits from removal of the amine group in the linker. The Tm data indicate that the absence of biological activity is not due to an inability of the polyamides to bind to DNA, as each of the IPA conjugates has a lower Tm than their respective free-amine parent compounds, but higher biological activity. The triamine-bearing polyamides **24** and **28**, in fact, have higher Tm values than the neutral, aliphatic linker-bearing polyamides **25** and **29** (C7 linker) and **26** and **30** (C3 linker) (free amine and IPA, respectively, for each pair). These data suggest

that the presence of the tertiary amine and the potential positive charge increase the binding affinity of these polyamides. The potential downside of the positive charge, however, is abrogation of binding specificity, presumably due to a slower polyamide off-rate, which would increase the chances of the polyamide binding to sites other than the designed target sequence. As this is not the same trend observed for the HRE series in U251 and LNCaP cells, it is difficult to be certain in absence of a quantitative study of polyamide nuclear concentration as a function of linker modifications.

Oxime-linked polyamides were synthesized for both the HRE (**9**-**11**, **20**-**22**) and ARE (**31**-**33**, **42**-**44**) systems. The oxime had been introduced as an efficient means of conjugating a polyamide core to ^{18}F-labeled 4-fluorobenzaldehyde for PET studies, but the ability of polyamides possessing this motif to regulate endogenous gene expression was not known. The aldehyde tail groups selected were 4-fluorobenzaldehyde, 3-fluorobenzaldehyde and 3-carboxybenzaldehyde; 3-fluorobenzaldehyde was selected to probe position effects, and 3-carboxybenzaldehyde provided an oxime-linked mimic of the amide-linked IPA tail.

In U251 cells (Figure 4.6A), the oxime linked match (**9**-**11**) polyamides had activity comparable to the triamine- or C7-linker amide-linked IPA conjugates (**6** and **7**, respectively). Introduction of the oxime linker had a more significant impact on the biological activity of the mismatch polyamides **20**-**22**, which were as active as their oxime-linked counterparts **9**-**11**. A similar result was seen for the C7- and C3-linker, amide-linked IPA conjugate pairs **7** and **18**, and **8** and **19**, in U251 cells. Thus, although there does appear to be tolerance of linker variations, the greatest difference in biological activity within an HRE match-mismatch pair in U251 cells is the triamine-linked IPA compounds **6** and **17**.

In LNCaP cells (Figure 4.6B), the HRE-oxime series **9**-**11** were more active than the amide-linked series **6**-**8**, and in fact reduced VEGF mRNA expression levels to below baseline levels. In this case, introduction of fluorine on the tail restored the match-mismatch activity difference to ~two-fold (i.e. **9** and **20**, or **10** and **21**, or **11** and **22**). The Tm data for the oxime-linked polyamides and their parent unconjugated polyamides do not indicate that

differences in binding affinity play any role in the differences in activity; the Tm values measured for the amide series, both match and mismatch cores, are comparable to those measured for the oxime series.

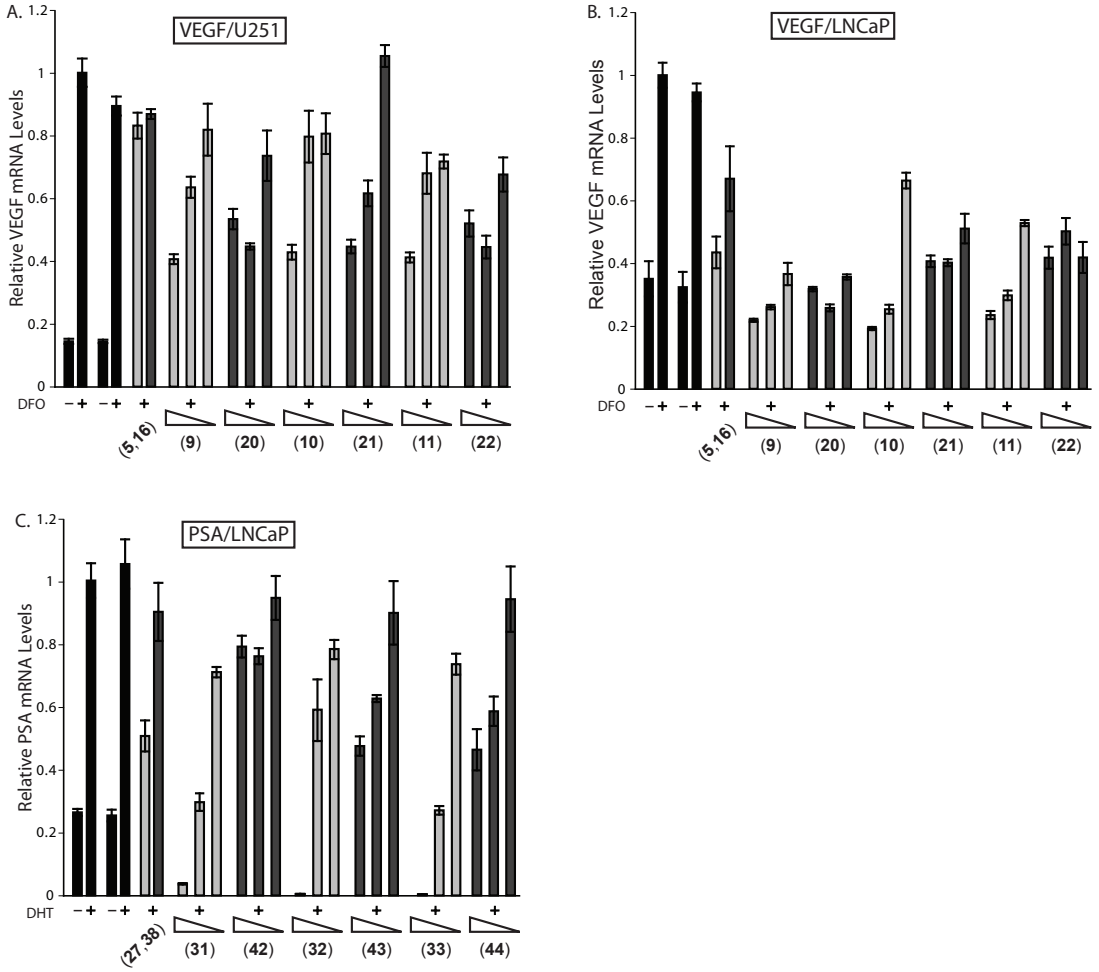


Figure 4.6. Gene-regulation effects measured by quantitative real-time RT-PCR for the oxime-linked polyamides **9**-**11**, **20**-**23**, **31**-**33** and **42**-**44**, and the hydroxylamine tail control compounds **5**, **16**, **27**, and **38**. Induced and uniduced control conditions are indicated by black bars, match core compounds by light grey bars, and mismatch core compounds by dark grey bars. Uninduced and induced control compounds without DMSO are on the left, and with 0.1% DMSO on the right. Errors shown are the fractional standard deviation. A) Effect on DFO-induced expression of VEGF by HRE-targeted match and mismatch polyamide cores in **U251** cells at 1, 0.2, and 0.02 μM; B) Effect on DFO-induced expression of VEGF by HRE-targeted match and mismatch polyamide cores in **LNCaP** cells at 10, 2 and 1 μM; C) Effect on DHT-induced expression of PSA mRNA by ARE-targeted match and mismatch polyamide cores in **LNCaP** cells at 5, 1 and 0.1 μM.

Polyamide regulation of VEGF expression in LNCaP had not been studied, but it was known that polyamide-effected regulation of PSA in LNCaP cells can be achieved with 10 and 5 μM **28**. A ten-fold higher concentration of HRE polyamides was required in LNCaP cells as compared to U251 cells to effect comparable VEGF gene regulation. Although the effects on VEGF expression in LNCaP and U251 cells of the HRE polyamide series

vary somewhat, it is encouraging that there are no drastic difference in biological activity measured for any of the compounds tested between the two cell lines. This suggests that a structural modification that results in increased gene regulation in one cell line may have a similar effect for the same core in other cell lines.

The data for the oxime-linked ARE series (**31**-**33**, **42**-**44**) in LNCaP cells (Figure 4.6C) indicate that the oxime linkage greatly increases the gene regulation activity of the ARE polyamides. PSA mRNA expression was negligible in cells treated with **31**-**33** at 5 μM, but there was evidence of cell death upon visual inspection (vide infra) and levels of total mRNA were very low. The average gene expression titration curve generated by quantitative real-time RT-PCR for **28**, **31** and **33** is shown in Figure 4.7B. These data illustrate the impact of the oxime linkage on the biological activity, as the activity curves of the oxime compounds **31** and **33** are shifted down by a factor of ≥10-fold from the amide-linked triamine IPA parent polyamide **28** (IC_{50} values in Table 4.1).

Table 4.2. Table of IC_{50} values (μM) for Growth Inhibition and Gene Regulation in LNCaP cells for **28**-**33**; data is shown in Figure 4.6. IC_{50} for gene regulation was measured only for **28**, **31** and **33**. For **28**, **31** and **33**, the (IC_{50} for Growth Inhibition)/(IC_{50} for Gene Regulation) ratio was calculated.

Compound Number	Growth Inhibition IC50 (μM)	Gene Regulation IC50 (μM)	Growth Inhibition IC50 (μM) / Gene Regulation IC50 (μM)
28	> 30	6.0 (+/- 2.6)	≥5
29	> 30	------	
30	> 30	------	
31	1.7 (+/- 0.3)	0.3 (+/- 0.1)	6.1
32	0.6 (+/- 0.1)	------	
33	0.4 (+/- 0.0)	0.6 (+/- 0.3)	0.7

To quantitate the effect of **28**-**33** and **39**-**44** on cell growth in LNCaP cells, cells were treated with compounds and viability was measured (Figure 4.7A **(28**-**33**) and 4.7C (**39**-**44**)). Cells treated with the amide-linked compounds at these concentrations did not show ≥50% decrease in cell growth, so an IC_{50} value for growth inhibition could not be calculated. IC_{50} values for growth inhibition were calculated for **31**-**33** (Table 4.1.) Oxime mismatch polyamide **42** minimally affected cell viability, while **43** and **44** had estimated IC_{50} values for cell viability are ≥ 30 μM**.** The IC_{50} values for gene regulation for **31** and **33** are 20- and

10-fold lower than **28**, respectively. In the case of the 4-Fluoro oxime-linked compound **33**, the IC_{50} for growth inhibition and for gene regulation are within error of each other. The IC_{50} for growth inhibition of **28** and **31** is ~5-fold higher than the IC_{50} values for gene regulation, but both these values are ~20-fold lower for **31** that **28**, indicating that **31** is ~20-fold more potent than **28** and more specific than **33**.

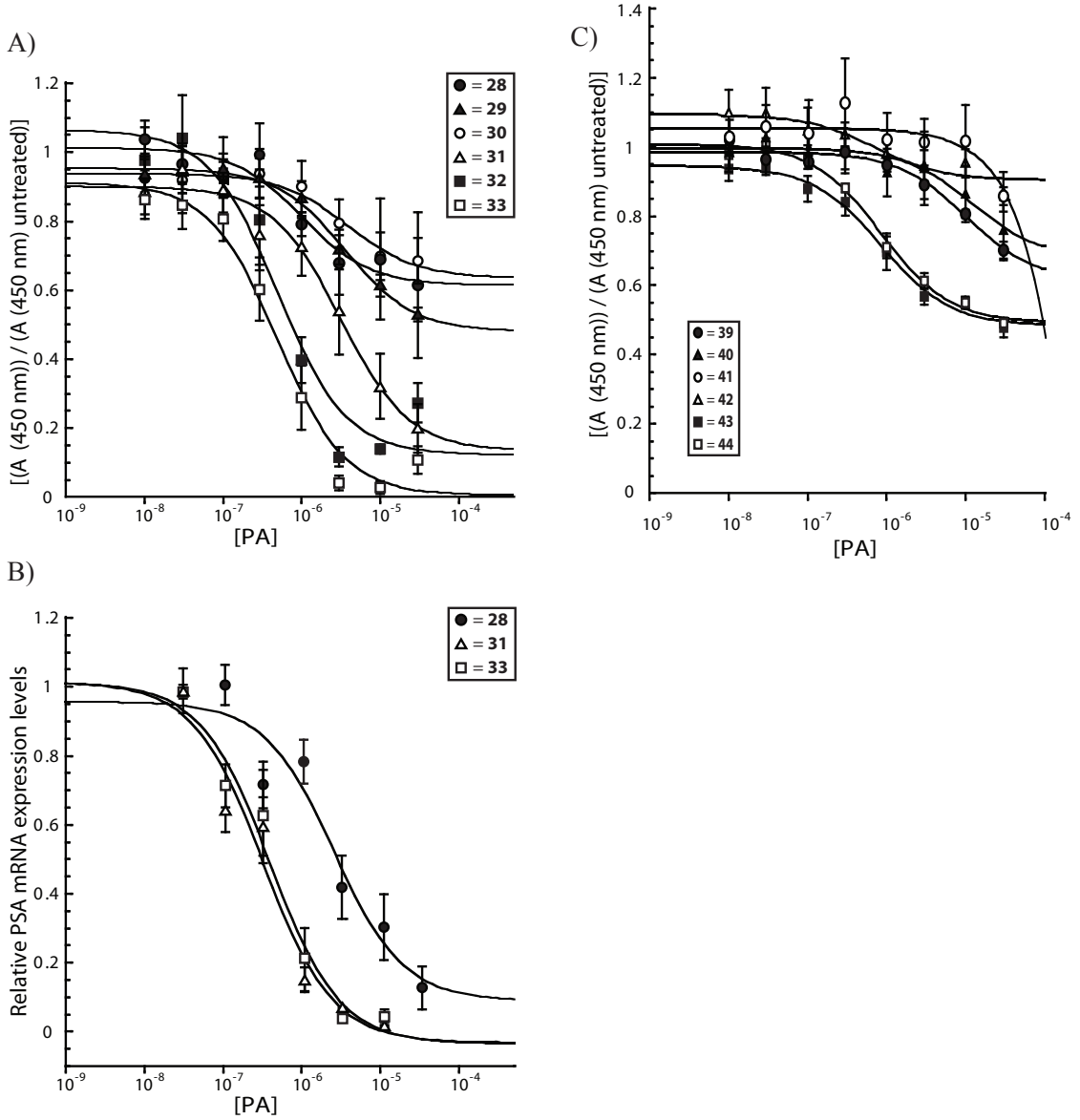


Figure 4.7. Cell viability data and IC_{50} values for gene regulation for in **LNCaP** cell culture. A) Cell viability was measured by a WST-1 colorimetric assay for **LNCaP** cells dosed for 48 hours with polyamides **28-33** at the following concentrations: 30, 10, 3, 1, 0.3, 0.1, 0.03 and 0.01 μM. For **31-33**, IC_{50} values for cell viability were extracted from isotherms generated in KaleidaGraph from the data. B) Quantitative real-time RT-PCR data for **28**, **31** and **33** in **LNCaP** cells to establish IC_{50} values for PSA gene regulation under inducing conditions. Concentrations employed of **28** were 30, 10, 3, 1, 0.3 and 0.1 μM, and of **31** and **33** were 10, 3, 1, 0.3, 0.1 and 0.03 μM. IC_{50} values were extracted from isotherms generated in Kaleidagraph from the RT-PCR titrations. C) IC_{50} values (μM) for cell viability estimated or calculated for **28-33** and IC_{50} values (μM) for PSA gene regulation for **28**, **31** and **33** shown in tabular format. Values given are average values of three independent experiments; the standard error of the mean is given in parentheses.

These results indicate that the oxime linkage can effect a significant increase in the cell uptake and gene regulation activity of polyamides. For unconjugated hydroxylamine polyamides **27** and **38**, a ~50% and 10% downregulation was measured, respectively, of induced PSA mRNA expression in LNCaP cells at 5 μM, and **27** and **38** were significantly less active than the oxime conjugates. Although previous studies have rigorously verified the stability of oxime-linked conjugates *in vivo*, a control experiment employing polyamide-lacking compounds **45**-**48** were synthesized and their cytotoxicity measured.[19] It was rationalized that equivalent cytotoxicity to oxime polyamide conjugates would serve as an indirect measure of bond lability, whereas a lack to toxicity would reinforce the robustness of this chemical bond. The growth inhibitory effects of these compounds were studied in U251 and LNCaP cells (Figure 4.6). Treatment with **45**-**48** in both cells lines failed to cause dramatic cell death at any concentrations used (10–0.002 μM in U251, 50–0.02 μM in LNCaP), indicating that the oxime linkage did not decompose and invoke cell death. These results suggest that the oxime-linkage is promising for future work aimed at increasing polyamide potency.

Although not the focus of the current study, the mechanism(s) of polyamide cell entry, nuclear localization and cell exit are active areas of investigation.[11] Our current working hypothesis for polyamide biological activity is that passive or non-specific mechanisms of polyamide cell uptake and competing efflux pathways influence polyamide accumulation in the cell nucleus, and structural modifications that result in improved polyamide biological activity may reflect a decrease in efflux. Polyamides are believed in part to undergo active efflux by p-glycoprotein pumps, which have more pronounced activity in immortalized cancer cells such as those employed in the current study.[20,21] The chemical agent verapamil, which inhibits p-glycoprotein pump activity, has been found to result in nuclear localization for some fluorophore-polyamide conjugates.[10] The mechanistic basis for the improved activity observed upon introduction of structural features such as the IPA tail is unknown, but we speculate that it is partially due to reduced polyamide efflux.

The tolerance of modifications to the linker group suggest that further modifications to the linker group may lead to improved polyamide biological activity. Comparison of the data gathered for the HRE match and mismatch polyamides in U251 and LNCaP cells indicate that, while the concentrations required to achieve acceptable levels of gene regulation activity will vary between cell lines, the effect of structural modifications will most likely be comparable between cell lines, i.e. the trends in activity will translate between cell lines. In addition, the use of an oxime linkage in place of an amide linkage between the tail and linker groups was found to increase the gene regulation activity of polyamides. This increase was most noticeable for the ARE oxime-linked series **31**-**33** in LNCaP cells. The dramatic increase in activity measured for **31** compared to the previously published **28** makes this compound and related derivatives useful for future ARE studies. In addition, these results suggest that the oxime linkage may be a useful structural feature to increase the biological activity of polyamides in other biological studies. Although this study did not uncover a set of predictive structural requirements for polyamide cell uptake, the discovery of the utility of the oxime linkage for gene regulation studies is a welcome addition to the known set of structural modifications to increase polyamide cell uptake.

Materials and Methods

Synthesis of Polyamides. Polyamides were synthesized by solid-phase methods on Kaiser oxime resin (NovaBiochem), or Boc-β-alanine-PAM resin (Peptides International), were cleaved from resin with 3,3'-diamino-*N*-methyl-dipropylamine, N, N'-dimethylpropane-1, 3-diamine (Dp), heptane-1, 7-diamine, propane-1, 3-diamine or *tert*-butyl 3-aminopropoxycarbamate, and purified by reverse-phase HPLC.[16, 22-24] Synthesis of IPA conjugates and oxime conjugates was as previously described.[8, 16] Reaction progress was monitored by analytical HPLC at 310 nm. Turn deprotection was done through addition of 1 mL 1:1 TFA:DCM at ambient temperature, and purification by reverse-phase preparatory HPLC performed immediately following successful deprotection. Reverse-phase HPLC solvent systems were 0.1% TFA (aqueous) and acetonitrile. Polyamide purity and identity

were assessed by analytical HPLC and MALDI-ToF MS, and polyamides quantitated by UV-vis at 310 nm (ε = 69,360 M^{-1} cm^{-1} for eight-ring hairpin polyamides). All polyamides were ≥95% by analytical HPLC (310 nm). Supporting Information Table SI 1 provides expected and observed ($M+H^+$) (MALDI-ToF), and purity by analytical HPLC for **1**-**44**. ^1H and ^{13}C NMR spectra were obtained on a 500 MHz spectrometer.

Synthesis of 45. 1 equivalent (84.3 mg, 0.44 mmol) of *tert*-butyl 3-(aminooxy) propylcarbamate was dissolved in 300 μL DMF, and 1.2 equivalents of 3-formyl benzoic acid added as a 50 mM solution in DMF.[25] Reaction progress was monitored by analytical HPLC at 254 nm for starting material consumption and production of a mixture of the E and Z -3-(10,10-dimethyl-8-oxo-3,9-dioxa-2,7-diazaundec-1-enyl)benzoic acid intermediate. Upon consumption of starting material, the Boc protecting group was removed by addition of 4 mL of 1:1 TFA:DCM. After 15 minutes, the reaction was concentrated by rotovap, and the product purified by reverse phase preparatory HPLC. The E and Z products were separable by preparatory HPLC, but interconvert at room temperature on the time scale of the experiments, so were recombined to yield 43.3 mg (44% yield) of 3-((3-aminopropoxyimino)methyl)benzoic acid (**45**). ^1H NMR (DMSO-d_6): δ 13.23 (s, 1H), 8.36 (s, 1H), 8.20 (app t, J = 1.8 Hz, 1H), 7.96 (app dt, 7.8, 1.8, 1 Hz, 1H), 7.88 (br s, 3H), 7.84 (app dt, 8.0, 1.5 Hz, 1H), 7.55 (app t, 7.8 Hz, 1H), 4.20 (t, 6.3 Hz, 2H), 2.91 (t, 7.3 Hz, 2H), 1.96 (m, 2H). ^{13}C NMR (DMSO-d_6): δ 166.8, 148.6, 132.3, 131.4, 131.0, 130.6, 129.2, 127.4, 70.6, 36.2, 27.0. Exact mass ($M+H^+$): calc'd, 223.1083; found, 223.1078.

Synthesis of 46. 1 equivalent (83.8 mg, 0.44 mmol) of *tert*-butyl 3-(aminooxy) propylcarbamate was used to synthesize **45**, as above. Following Boc deprotection and prior to purification, 2 equivalents $(Ac)_2O$ (83.2 μL, 0.88 mmol) and 2.5 equivalents DIEA (191.6 μL, 1.1 mmol) were added. Reaction progress was monitored by analytical HPLC at 254 nm, and the product was purified by preparatory reverse phase HPLC. As above, the E and Z enantiomers were separable, but were recombined to yield 108.3 mg (93% yield) of the 3-((3-acetamidopropoxyimino)methyl)benzoic acid product **46**. ^1H NMR

(DMSO-d_6): δ 13.14 (s, 1H), 8.33 (s, 1H), 8.18 (t, 1.8 Hz, 1H), 7.95 (app dt, 7.5, 1.5 Hz, 1H), 7.87 (br t, 4.8 Hz, 1H), 7.83 (app dt, 8.3, 1.5 Hz, 1H), 7.54 (app t, 7.8 Hz, 1H), 4.14 (t, 6.5 Hz, 2H), 3.12 (m, 2H), 1.79 (s, 3H), 1.78 (m, 2H). ^{13}C NMR (DMSO-d_6): δ 169.1, 166.8, 148.0, 132.5, 131.3, 130.9, 130.4, 129.2, 127.4, 71.5, 35.4, 28.9, 22.6. Exact mass ($M+H^+$): calc'd, 265.1188; found, 265.1199.

Synthesis of 47. The synthesis and purification of **47** proceeded as described for **45**, but with 4-fluorobenzaldehyde. Yield: 66.8 mg (74%). ^{1}H NMR (DMSO-d_6): δ 8.27 (s, 1H), 7.86 (br s, 3H), 7.66 (m, 2H), 7.26 (m, 2H), 4.17 (t, 6.3 Hz, 2H), 2.90 (m, 2H), 1.94 (m, 2H). ^{13}C NMR (DMSO-d_6): δ 164.0, 162.0, 129.0 (J = 28 Hz), 128.5 (J = 11 Hz), 115.9 (J = 86 Hz), 70.4, 36.2, 27.0. Exact Mass ($M+H^+$): calc'd, 197.1090; found, 197.1091.

Synthesis of 48. The synthesis and purification of **48** proceeded as described for **46**, but with 4-fluorobenzaldehyde. Yield: 103.3 mg (89%). ^{1}H NMR (DMSO-d_6): δ 8.24 (s, 1H), 7.87 (br s, 1H), 7.65 (m, 2H), 7.24 (m, 2H), 4.10 (t, 6.5 Hz, 2H), 3.11 (m, 2H), 1.78 (s, 3H), 1.76 (m, 2H). ^{13}C NMR (DMSO-d_6): δ 169.1, 164.0, 162.0, 129.0 (J = 33 Hz), 128.6 (J = 13 Hz), 115.8 (J = 87.5 Hz), 71.3, 35.5, 28.9, 22.6. Exact Mass ($M+H^+$): calc'd, 239.1196; found, 239.1189.

Determination of DNA Melting Temperature (Tm) values. Melting Temperatures were measured according to a previously published methodology on a Varian Cary 100 spectrophotometer equipped with a thermocontrolled cuvette holder, and quartz cuvettes (1 cm pathlength)[18]. The buffer was an aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0. Oligonucleotide sequences were: 5'-GTGCATACGTGGGC-3' (HRE 14mer top); 5'-GCCCACGTATGCAC-3' (HRE 14mer bottom); 5'-TTGCAGAACAGCAA-3' (ARE 14mer top); and 5'-TTGCTGTTCTGCAA-3' (ARE 14–mer bottom).

Cell Culture Experiments. For cell culture experiments, a polyamide stock solution in DMSO and DNAse/RNAse-free water was prepared by dissolving 5-30 nmol polyamide in DMSO (1-3 μL), followed by addition of DNAse/RNAse-free water to a target concentration

of 50 M. Suspensions were centrifugated to remove undissolved material, and solution concentration determined by UV-vis measurement of a 100× diluted solution at 310 nm (ε = 69,360 M^{-1} cm^{-1} for eight-ring hairpin polyamides.) The DMSO concentration of cell media solutions was kept below 0.1%, and a DMSO control was included in quantitative real-time RT-PCR (see below).[17] Isotherms were generated using the following modified Hill equation: $Y = m_1 + (m_2 - m_1) / (1 + (m_0 / m_3))$; $m_1 = 100$, $m_2 = 5000$, $m_3 = 5e^{-7}$.

Measurement of Cell Viability in U251 and LNCaP Cells. U251 and LNCaP cells were maintained as previously described.[8, 14] For the growth inhibition assay, cells were plated in 96-well plates in 0.2 mL at 10-15 x 10^2 cells per well (15-20 x 10^3 cells/mL). After 24 hours (U251) or 48 hours (LNCaP), 150 μL of media was removed and replaced with 50 μL of 2x (polyamide concentration) cell media solutions. After 48 hours, the media was replaced with 100 μL fresh media, and cells were allowed to recover for 24 hrs before addition of 10 μL WST-1 reagent (Roche) to each well. Cells were incubated for 30 minutes, and absorbance at 450 nm measured. Untreated controls and cell-free, media-only controls were included on each plate, and each well corrected for average background absorption of media-only controls before normalizing average absorption for each concentration to untreated controls.

Measurement of Induced Gene Expression in U251 and LNCaP Cells. Measurement of hypoxia-induced VEGF expression in U251 cells and DHT-induced PSA expression in LNCaP cells was as previously described.[8, 14] Measurement of hypoxia-induced VEGF expression in LNCaP cells followed the protocol for DHT-induced PSA expression in LNCaP cells, with the following changes: gene induction was by DFO, and FBS (not charcoal stripped) was used. Primer sequences were as follows: VEGF, L 5'- AGGGCAGAATCATCACGAAG -3', R 5'-GGGTACTCCTGGAAGATGTCC -3'; PSA, L 5'-TCTGCGGCGGTGTTCTG-3', R 5'-GCCGACCCAGCAAGATCA -3'; ß-glucuronidase, L 5'-CTCATTTGGAATTTTGCCGATT-3', R 5'-CCGAGTGAAGATCCCCTTTTTA-3'.

Table 4.3. MALDI-ToF and Purity data for compounds **1**-**44**.

Compound Number	MALDI-ToF: Expected ($M+H^+$); Observed ($M+H^+$)	Purity*	Compound Number	MALDI-ToF Expected (M+H): Observed (M+H):	Purity*
1	1274.49; 1274.51	99.3	**23**	1237.59; 1237.48	95.0
2	1246.50; 1246.81	99.3	**24**	1209.59; 1209.89	95.6
3	1231.49; 1231.77	99.7	**25**	1194.58; 1194.53	100.0
4	1175.43; 1175.61	98.7	**26**	1138.52; 1138.50	99.7
5	1191.42; 1191.38	97.2	**27**	1154.51; 1154.62	96.8
6	1394.52; 1394.46	98.0	**28**	1357.61; 1357.59	98.8
7	1379.50; 1379.72	99.3	**29**	1342.60; 1342.55	99.6
8	1323.44; 1323.52	99.9	**30**	1286.64; 1286.59	99.0
9	1323.44; 1323.44	95.6	**31**	1286.54; 1286.56	96.1
10	1297.44; 1297.48	95.9	**32**	1260.54; 1260.63	95.3
11	1319.42; 1319.52[a]	97.2	**33**	1260.54; 1260.51	96.4
12	1238.58; 1238.51	96.5	**34**	1238.58; 1238.56	99.0
13	1210.59; 1210.76	98.1	**35**	1210.59; 1210.59	98.0
14	1195.58; 1195.57	98.7	**36**	1195.58; 1195.56	99.3
15	1139.51; 1139.49	98.8	**37**	1139.51; 1139.77	95.1
16	1155.51; 1155.83	98.4	**38**	1155.51; 1155.40	97.4
17	1358.60; 1358.66	97.8	**39**	1358.60; 1358.70	95.4
18	1343.59; 1343.76	99.9	**40**	1343.59; 1343.52	100.0
19	1287.53; 1287.56	99.4	**41**	1287.53; 1287.60	98.4
20	1287.53; 1287.73	97.6	**42**	1287.53; 1287.53	95.9
21	1261.53; 1261.87	97.3	**43**	1261.53; 1261.56	98.9
22	1261.53; 1261.51	98.8	**44**	1261.53; 1261.51	96.2

MALDI-ToF data gives calculated exact mass for ($M+H^+$) and observed mass for ($M+H^+$) for polyamides **1**-**44**. * Purity measured by analytical HPLC at 310 nm; purity for oxime-linked compounds **9**-**11**, **20**-**22**, **31**-**33** and **42**-**44** included both E and Z isomers. [a] Mass found [$M+Na^+$], compound and data provided by Dr. Daniel Harki.

Focused Library of Polyamide-Tail Conjugates

The following section contains unpublished data.

The ability to design and implement potential future applications of polyamides in cell culture or whole organisms depends in part on knowledge of the extent of possible modifications to polyamides available, i.e. whether the FITC and its IPA derivative are discrete, singular solutions to the C-terminal tail identity, isolated but not unique solutions, or members of a much more extensive structure space around the C-terminal tail available for biologically active polyamides. The latter case would indicate that biologically active conjugates of polyamide DNA-recognition elements and other function groups are accessible, e.g. conjugates with by DNA intercalators, DNA cross-linking groups, recognition elements for endogenous or exogenous chemical species, or substrate partners for DNA/polyamide-templated reactions.

To this end, in addition to the linker and linkage modifications detailed above, a library of polyamides was synthesized in which the C-terminal tail group was varied and assayed for changes in biological activity. The HRE- and ARE-targeted match and mismatch polyamide cores were selected and polyamide-tail conjugate biological activity was assessed via induced VEGF (U251 and LNCaP cells) and PSA (LNCaP cells) gene expression evaluated by quantitative real-time RT-PCR measurement of gene mRNA levels. The set of tail moieties selected for study are shown in Figure 4.8, and were designed as a structure-activity relationship series in an attempt to ascertain the structural features of the IPA tail responsible for its enhanced biological activity.

Polyamides **49-60** are HRE match polyamides, **61-72** are HRE mismatch polyamides, **73-84** are ARE match polyamides, and **85-96** are ARE mismatch polyamides. With each set of polyamide cores, a series of 13 tail variations was introduced, including the previously described benzoic acid, 3-hydroxybenzoic acid and IPA conjugates (Chapter 2). As discussed previously, these tails had been initially selected to determine whether the

cell uptake and nuclear localization of the original FITC conjugate was due to the presence of an aromatic ring, a phenol, or a benzoic acid group. Upon identification of the IPA tail as the best combination of ease of synthesis and biological activity, it became of interest to determine, if possible, what structural feature of IPA lay at the heart of its positive activity profile, or how tolerant the C-terminal tail group is to modifications.

To examine the effect of moving the carboxylic acid group from the *meta* to the *para* position on the ring relative to the amide linkage, the terephthalic acid tail was selected. -polyamide conjugates were synthesized. Trimesic acid was introduced as a tail group to determine whether the possible role of the carboxylic acid group in the biological activity of FITC- and IPA-conjugates would be enhanced through addition of another carboxylic acid group. The methyl-ester derivative was designed to probe whether the ability of the carboxylic acid to be deprotonated and carry a negative charge was responsible for biological activity by acting as an endogenous buffer group and facilitating polyamide escape from endosomes and/or acidic lysosomes. The amide and nitro aromatic derivatives are similar in shape to the carboxylic acid group, but differ in charge distribution and surface potential. Both the glutaric acid and cyclohexane-1,3-dicarboxylic acid tail conjugates carry a carboxylic acid group, but lack the aromatic phenyl group of the IPA tail. Methylamine conjugates (e.g. **57**) can be protonated to carry an overall positive charge, but lack the carbonyl group of the amide-bearing conjugates (e.g. **53**). The 3-hydroxy benzoic acid tail had demonstrated some activity in earlier studies and was selected for use here for benchmark comparison to the 3-Fluoro and 4-Fluoro benzoic acid derivatives (Chapter 2).

The cell culture results of polyamides bearing these tails in U251 and LNCaP cells are shown in Figures 4.9-4.11. U251 cells were dosed with HRE match and mismatch core compounds at 1 and 0.5 μM, and the βDp and unconjugated triamine tail controls were included. Differences in the biological activity of these new tail derivatives was less pronounced at 0.5 mM than at 1 μM, but less of a disparity in biological activity was seen

between the match and mismatch compounds at 1 mM than at 0.5 μM. This is similar to what was observed for the oxime-linked compounds above (Figure 4.6). Discussion of HRE cores in U251 will focus on data at the 1 μM dose.

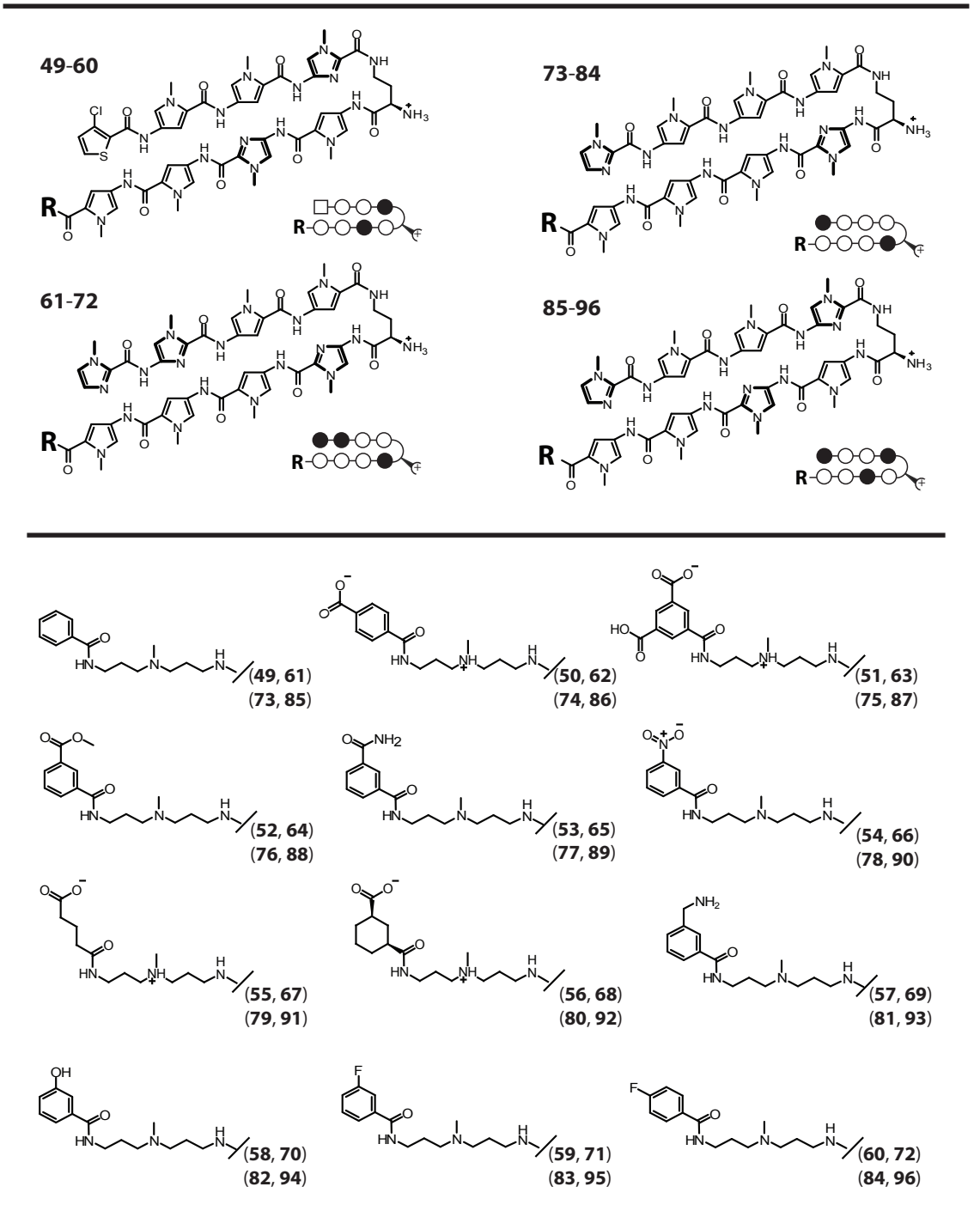


Figure 4.8. Structural representation of polyamide-tail conjugates **49-96**. Shown at top are structural representations of the polyamide heterocycle cores for the VEGF (**49-72**) and PSA (**73-96**) systems. At bottom are shown the structures of the focused library of tails selected for this study.

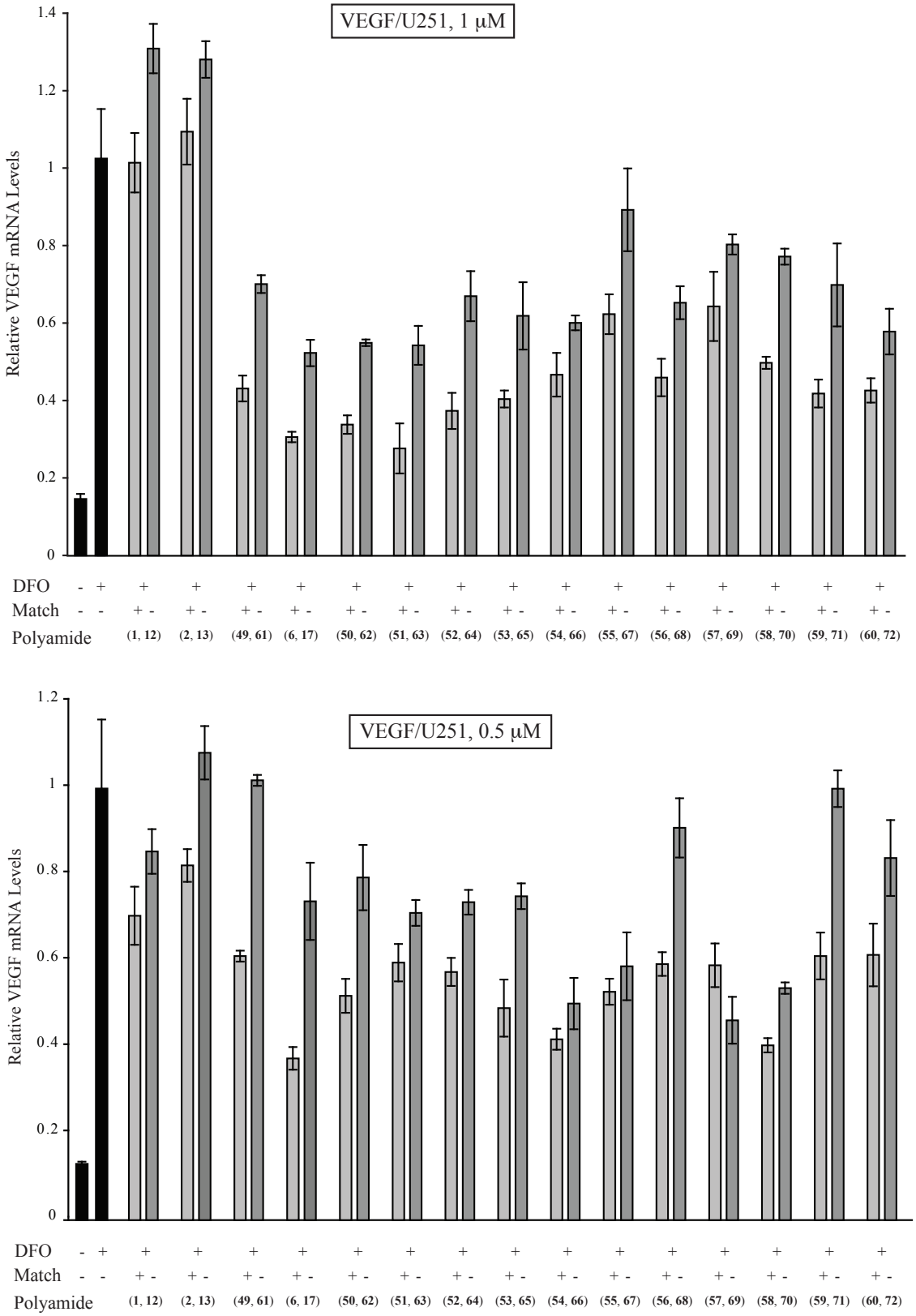


Figure 4.9. Quantitative real-time RT-PCR data in cultured U251 cells induced with DFO and treated with **49**-**72** and controls **1**, **2**, **6**, **12**, **13**, and **17**. U251 cells treated with polyamides at A) 1 μM and B) 0.5 μM.

In U251 cell cultures dosed with polyamides **49**-**72** at 1 μM, the tail identity had an effect on polyamide biological activity, but not in a fashion that facilitated extraction of a set of predictive uptake guidelines. It had been suggested that the presence of the carboxylic acid group in the fluorescein dye acted as an internal buffering group and facilitated polyamide escape from lysosomal acidification. Indeed, moving from the benzoic acid (**49**, **61**) to the IPA (**6**, **17**), terephthalic acid (**50**, **62**) and trimesic acid (**51**, **63**) tails, a slight increase in activity is observed, suggesting that the carboxylic acid is a positive determinant for biological activity. Under the hypothesis suggested above, which implies that activity is predominantly a function of uptake that is facilitated by the acid moiety, it was expected that switching from the IPA tail to the methyl ester would abrogate biological activity. The biological activity of the methyl ester tail derivatives (**52**, **64**) was, however, at best slightly lower than the measured activity of the IPA, terephthalic acid and trimesic acids tails, and was clearly within error of these acid-bearing tail derivatives.

The 3-nitro benzoic acid tail was selected as a means of probing whether the potential negative charge of the carboxylic acid strongly influences polyamides biological activity. Examination of the data in Figure 4.9A reveals a modest decrease in biological activity for the 3-nitro tail derivative compared to the acid-bearing tails, and comparable to the activity of the benzoic acid conjugate (**49**, **61**).

The lack of activity measured in this system for the triamine and βDp tails strongly suggested that the use of an aliphatic linker terminating in an amine group is a negative determinant for polyamide biological activity, and thus that incorporation into the C-terminal tail group of an aromatic phenyl ring rather than the carboxylic acid group is critical to biological activity. In order to evaluate the importance of the aromatic phenyl ring on biological activity, two derivatives lacking the aromatic ring were synthesized and tested: the glutaric acid tail (**55**, **67**), and the 3-carboxycyclohexanoic acid tail (**56**, **68**). The activity of the glutaric acid tail was approximately half that of the IPA tail, and less than that measured for the 3-carboxycyclohexanoic acid tail, which was closer in activity

to the 3-nitro, 3-amide and benzoic acid tails. These results suggest that the presence of a ring is important, and while it need not be an aromatic ring, aromaticity serves further to increase activity in cell culture. The 3-methylamine tail (**57**, **69**) was designed to probe the effect of concurrent removal of the carbonyl group and introduction of a tail capable of carrying a positive charge. This resulted in the most inactive group of the series. Next was a set of three tails comprised of aromatic rings bearing electronegative substituents: 3-hydroxybenzoic acid (**58**, **70**), 3-fluorobenzoic acid (**59**, **71**), and 4-fluorobenzoic acid (**60**, **72**). All three of these derivatives were as active as the benzoic acid tail conjugate.

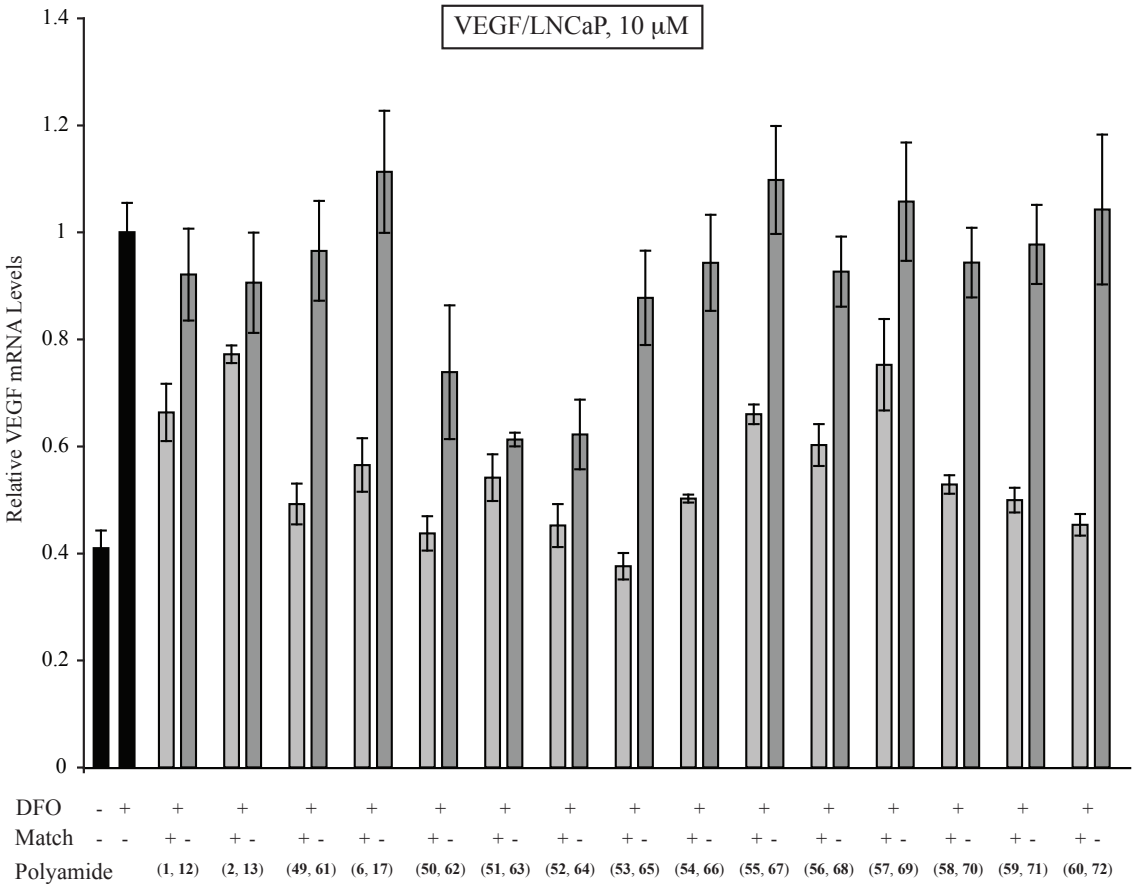


Figure 4.10. Quantitative real-time RT-PCR data in cultured LNCaP cells induced with DFO and treated with **49**-**72** and controls **1**, **2**, **6**, **12**, **13**, and **17**. LNCaP cells treated with polyamides at 10 μM.

To determine whether the trend in activities of these tails was specific to cultured U251 cells, the biological activity of this same series of HRE match and mismatch core tail derivatives was measured in LNCaP cells. The results in LNCaP cells were reasonably

comparable to those in U251 cells. The fold induction of VEGF mRNA was again less in LNCaP cells than in U251. At the concentration tested, there was also a much less significant difference in activity between the control polyamides (**1**, **2**, **12**, **13**) and the IPA tail standard (**6**, **17**) than was measured in the U251 cells at 1 μM. It was encouraging, however, that the data in LNCaP cells did not differ wildly from that in U251 cells, i.e. compounds found to be active in U251 cells were not inactive in LNCaP cells. Further, examination of the LNCaP results (Figure 4.10) shows that there is again appreciable tolerance for modifications to the aromatic ring through introduction of various substituents, while destruction of the aromatic ring or introduction of the methylamine decreased biological activity.

The cell culture results obtained in LNCaP cells for the ARE match and mismatch core series at 10 and 5 μM under inducing (DHT) conditions are shown in Figure 4.11. In contrast to VEGF mRNA induction in LNCaP cells, a 10-fold difference in PSA expression was measured between the uninduced and induced conditions, and a greater degree of variation in biological activity was measured for ARE-core conjugates in LNCaP cells. As for the HRE cores, destruction of the aromatic ring, either in the form of the glutaric acid (**79**, **91**) or 3-carboxylcyclohexanoic acid (**80**, **92**) derivatives resulted in a significant decline in biological activity. The methylamine derivative (**81**, **93**) again showed low levels of biological activity. As with the HRE series in U251 cells, a significant level of tolerance for changes to the ring substituents was observed. With the exception of the glutaric acid, 3-carboxycyclohexanoic acid, methyl amine and trimesic acid tail derivatives, appreciable levels of activity were observed for all other ARE polyamide-tail conjugates.

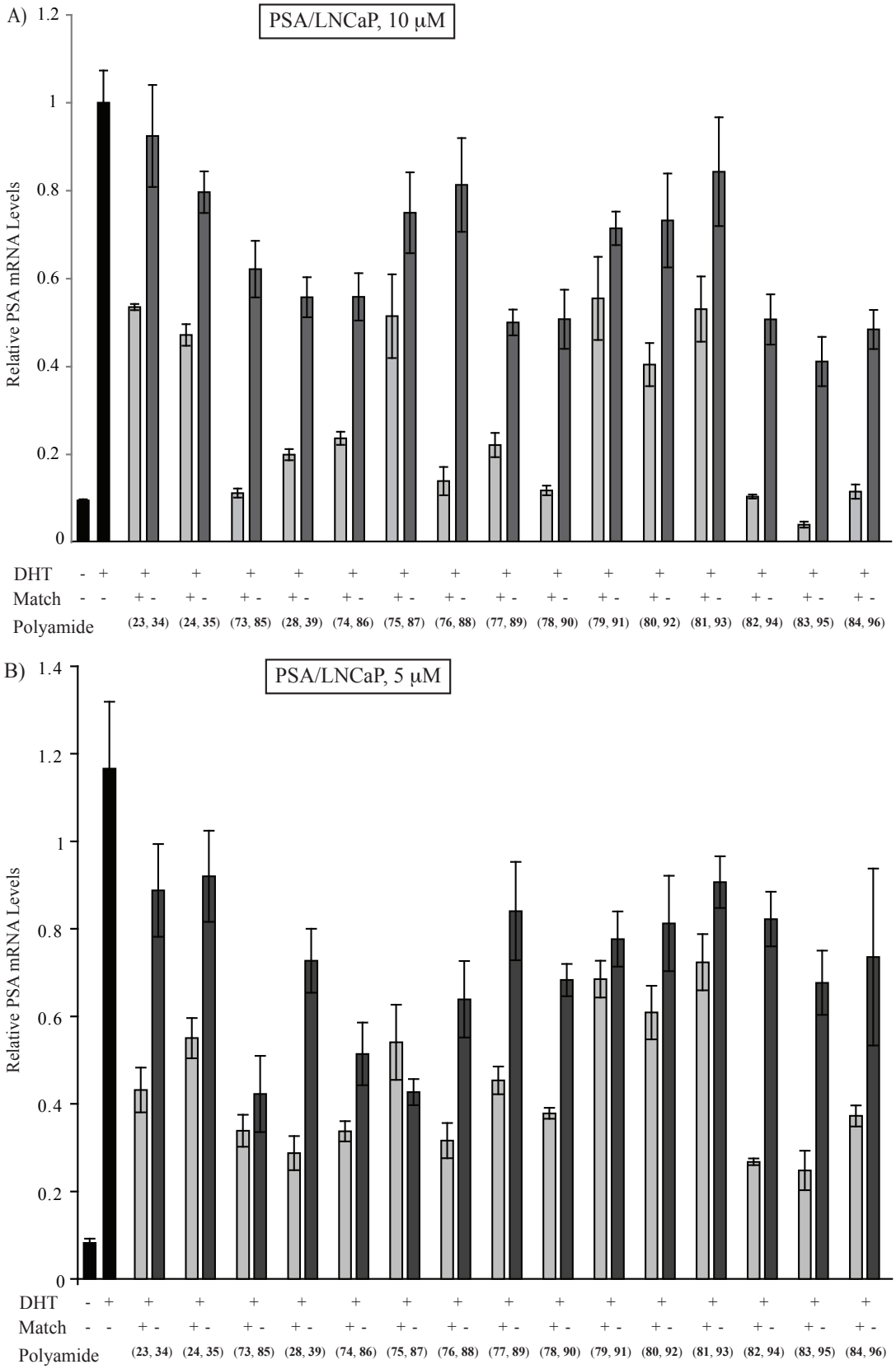


Figure 4.11. Quantitative real-time RT-PCR data in cultured LNCaP cells induced with DHT and treated with **73-96** and controls **23**, **24**, **28**, **34**, **35**, and **39** at A) 10 μM and B) 5 μM.

Although the focused library presented here did not fulfill the primary goal of uncovering structural determinants in the C-terminal tail group of biological activity of polyamides in cell culture, it did provide valuable information. The high level of tolerance for variations in the ring substituents suggests that the IPA tail is not a unique solution to polyamide uptake, but rather that there is an appreciable structure space available for cell-permeable polyamides. In combination with the results presented above for linker and linkage modifications, it appears that the C-terminal portion of polyamides is fungible without endangering compound activity in cell culture. The relative lack of activity measured for polyamides terminating in a βDp tail may reflect the poor activity of eight-ring βDp polyamides rather than polyamides as a compound class. It should be noted at this juncture, however, that other research groups have found success with βDp-terminated two-β-two polyamides.

The results for the ARE-core polyamides in LNCaP were more encouraging than the HRE-core polyamides in either LNCaP or U251 cells. This may reflect lower levels of biological activity of the HRE-match core itself, as the difference in activity of HRE-match and -mismatch congeners is often less than for ARE match/mismatch pairs. In absence of a larger data set of the activity of various cores in numerous cell lines, however, it is difficult to determine whether the HRE-match core is intrinsically less active than pure Im/Py cores. The general concordance between cell lines in the biological activity of this focused library is encouraging, as it suggests that positive hits in one cell line will likely translate into multiple other cell lines. An investigative tool that would aid greatly in efforts to parse whether the modifications being introduced in this study affect biological activity through improved cell uptake, decreased efflux, or another mechanism, is the ability to assay the intranuclear concentration of structurally diverse polyamides. Some efforts have been made in this direction by other researchers in the Dervan group, but have not yet been refined to provide a rapid, reliable intranuclear polyamide concentration assay.

Table 4.4: Tm Data for HRE Cores. Tm DNA only = 64.7 +/- 0.2

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

Compound Number	Tm (°C)	ΔTm (°C)	Compound Number	Tm (°C)	ΔTm (°C)
49	76.3 +/- 0.3	15.5	**61**	66.5 +/- 0.3	5.7
50	75.5 +/- 0.3	14.7	**62**	65.0 +/- 0.2	4.2
51	73.6 +/- 0.1	12.8	**63**	64.0 +/- 0.2	3.2
52	76.1 +/- 0.2	15.3	**64**	66.3 +/- 0.2	5.5
53	76.5 +/- 0.3	15.7	**65**	65.9 +/- 0.1	5.1
54	76.3+/- 0.2	15.5	**66**	66.1 +/- 0.1	5.3
55	74.7 +/- 0.2	13.9	**67**	64.3 +/- 0.4	3.5
56	74.7 +/- 0.4	13.9	**68**	64.3 +/- 0.2	3.5
57	78.8 +/- 0.2	18.0	**69**	68.6 +/- 0.4	7.8
58	77.0+/- 0.2	16.1	**70**	66.6 +/- 0.4	5.8
59	76.6 +/- 0.3	15.8	**71**	66.2 +/- 0.2	5.4
60	76.6 +/- 0.3	15.8	**72**	65.9 +/- 0.4	5.1

Oligonucleotide sequences for HRE-core compounds given in Table 4.1; conditions and buffers as for **1**-**44**.

Table 4.5: Tm Data for ARE Cores: Tm DNA only = 60.8 +/- 0.2

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

Compound Number	Tm (°C)	ΔTm (°C)	Compound Number	Tm (°C)	ΔTm (°C)
73	76.3 +/- 0.3	15.5	**85**	66.5 +/- 0.3	5.7
74	75.5 +/- 0.3	14.7	**86**	65.0 +/- 0.2	4.2
75	73.6 +/- 0.1	12.8	**87**	64.0 +/- 0.2	3.2
76	76.1 +/- 0.2	15.3	**88**	66.3 +/- 0.2	5.5
77	76.5 +/- 0.3	15.7	**89**	65.9 +/- 0.1	5.1
78	76.3+/- 0.2	15.5	**90**	66.1 +/- 0.1	5.3
79	74.7 +/- 0.2	13.9	**91**	64.3 +/- 0.4	3.5
80	74.7 +/- 0.4	13.9	**92**	64.3 +/- 0.2	3.5
81	78.8 +/- 0.2	18.0	**93**	68.6 +/- 0.4	7.8
82	77.0+/- 0.2	16.1	**94**	66.6 +/- 0.4	5.8
83	76.6 +/- 0.3	15.8	**95**	66.2 +/- 0.2	5.4
84	76.6 +/- 0.3	15.8	**96**	65.9 +/- 0.4	5.1

Oligonucleotide sequences for ARE-core compounds given in Table 4.1; conditions and buffers as for **1**-**44**.

References

1. Brennan, P.; Donev, R.; Hewamana, S., Targeting transcription factors for therapeutic benefit. *Mol Biosyst* **2008,** 4, (9), 909-919.

2. Berg, T., Inhibition of transcription factors with small organic molecules. *Curr Opin Chem Biol* **2008,** 12, (4), 464-471.

3. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol* **2003,** 13, (3), 284-299.

4. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998,** 391, (6666), 468-471.

5. Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C., Structural basis for G.C recognition in the DNA minor groove. *Nat Struct Biol* **1998,** 5, (2), 104-109.

6. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., Shape selective recognition of T.A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues. *Bioorg Med Chem* **2003,** 11, (20), 4333-4340.

7. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C.; Dervan, P. B., Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron* **2007,** 63, (27), 6146-6151.

8. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res* **2007,** 35, (2), 363-370.

9. Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B., Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. *Bioorg Med Chem* **2002,** 10, (10), 3313-3318.

10. 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., Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorg Med*

Chem Lett **2003,** 13, (9), 1565-1570.

11. 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, (21), 12063-12068.

12. 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. *Nucleic Acids Res* **2004,** 32, (9), 2802-2818.

13. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G., Jr.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A* **2004,** 101, (48), 16768-16773.

14. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci U S A* **2007,** 104, (25), 10418-10423.

15. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem Biol* **2007,** 2, (8), 561-571.

16. Harki, D. A.; Satyamurthy, N.; Stout, D. B.; Phelps, M. E.; Dervan, P. B., In vivo imaging of pyrrole-imidazole polyamides with positron emission tomography. *Proc Natl Acad Sci U S A* **2008,** 105, (35), 13039-13044.

17. Tjernberg, A.; Markova, N.; Griffiths, W. J.; Hallen, D., DMSO-related effects in protein characterization. *J Biomol Screen* **2006,** 11, (2), 131-137.

18. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J Am Chem Soc* **2008,** 130, (21), 6859-6866.

19. Kalia, J.; Raines, R. T., Hydrolytic stability of hydrazones and oximes. *Angew Chem Int Ed Engl* **2008,** 47, (39), 7523-7526.

20. Juliano, R. L.; Ling, V., A surface glycoprotein modulating drug permeability in

Chinese hamster ovary cell mutants. *Biochim Biophys Acta* **1976,** 455, (1), 152-162.

21. Ueda, K.; Cardarelli, C.; Gottesman, M. M.; Pastan, I., Expression of a full-length cDNA for the human "MDR1" gene confers resistance to colchicine, doxorubicin, and vinblastine. *Proc Natl Acad Sci U S A* **1987,** 84, (9), 3004-3008.

22. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg Med Chem* **2002,** 10, (8), 2767-2774.

23. Baird, E. E.; Dervan, P.B., Solid Phase Synthesis of polyamides containing imidazole and pyrrole amino acids. *J Am Chem Soc* **1996,** 118, 6141-6146.

24. Salisbury, C. M.; Maly, D. J.; Ellman, J. A., Peptide Microarrays for the Determination of Protease Substrate Specificity *J Am Chem Soc* **2002,** 124, 14868-14870.

25. Wallace, E.; Hurley, B.; Yang, H.; Lyssikatos, J.; Blake, J.; Marlow, A. Heterocyclic inhibitors of MEK and methods of use thereof., US Patent Application 20050054701 A1, 2005.

Appendix A

Experimental Time Courses

Experimental Time Courses for Quantitative Real-Time RT-PCR

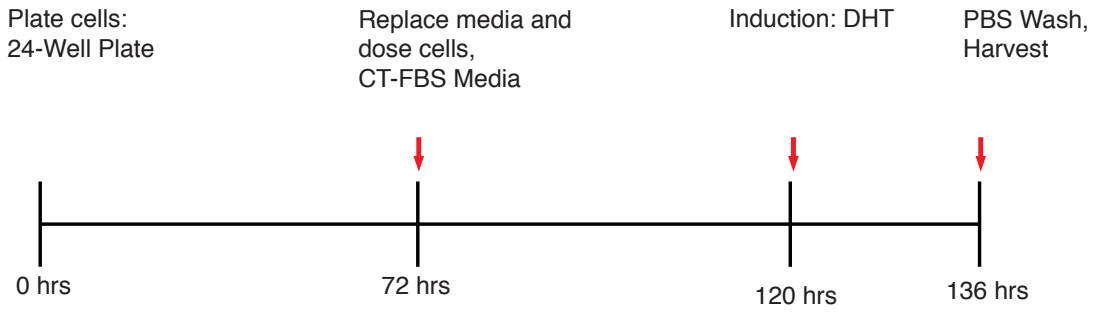
Provided here are schematic illustrations of the experimental time courses followed for quantitative real-time RT-PCR and cytotoxicity experiments in Chapters 2-4.

Quantitative Real-Time RT-PCR experiments

VEGF mRNA expression:

Plating density (Cell Line): 30-40K/mL, 0.5 mL (HeLa), 30-60K/mL, 0.5 mL/well (U251), 30-60K/mL, 0.5 mL/well (LNCaP)

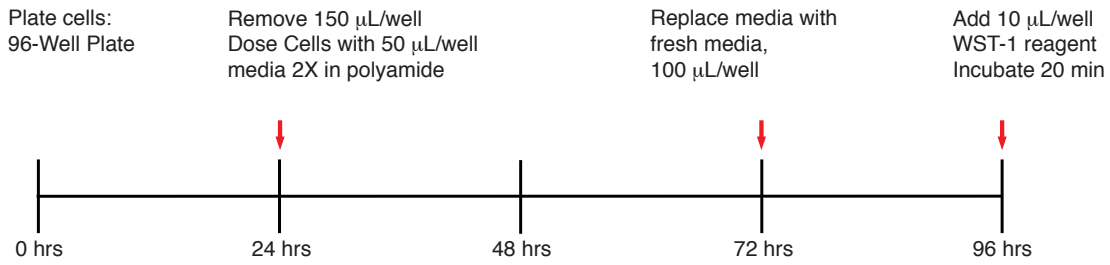
PSA mRNA expression:



Plating density for LNCaP cells: 30-60K/mL, 0.5 mL/well

Cytotoxicity and IC_{50} Measurements

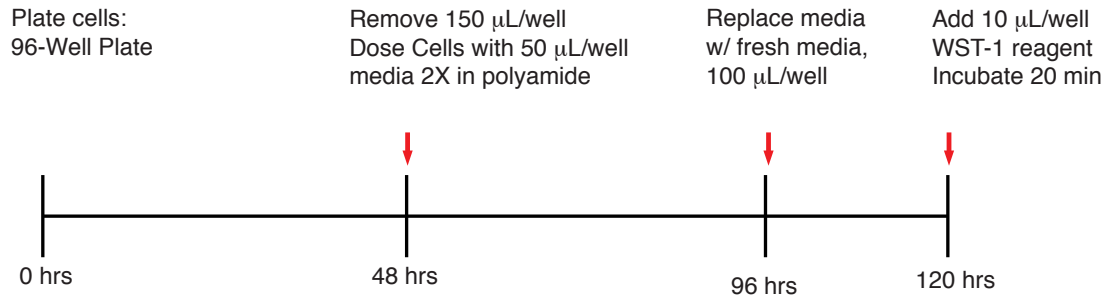
U251 cells:



Plating density in 96-well plates (Cell Density): 10-20K/mL, 0.2 mL/well (U251 cells)

No induction.

LNCaP Cells



Plating density in 96-well plates (Cell Density): 10-20K/mL, 0.2 mL/well (LNCaP cells), regular FBS (not CT-FBS), no induction. Replacement of media must be done slowly and carefully with LNCaP cells.

Appendix C

Plasmid Insert Sequences for pGL2-VEGF-*Luc* and pCSJ-FLT1 Plasmids

Plasmid Insert Sequences for Quantitative DNase I Titration Footprint Analysis

Plasmid Information for pGL2-VEGF-*Luc*, which was used for DNase I footprint titration K_a measurements of HRE polyamides in Chapters 2 and 3. This plasmid was a kind gift from the Kaelin group, who graciously provided the plasmid and plasmid insert sequence.

pGL2-VEGF-*Luc*: Contains WT VEGF HRE Sequence (bold), 197 bp sequence

5′-**CTCAGTTCCCTGGCAACATCT**GGGGTTGGGGGGGCAGCAGGAACAAGGGCCTCTGT

3′-GAGTCAAGGGACCGTTGTAGACCCCAACCCCCCCGTCGTCCTTGTTCCCGGAGACA

CTGCCCAGCTGCCTCCCCCTTTGGGTTTTGCCAGACTCCACAGTGCA**TACGTG**GGCTCC

GACGGGTCGACGGAGGGGGAAACCCAAAACGGTCTGAGGTGTCACGT**ATGCAC**CCGAGG

AACAGGTCCTCTTCCCTCCCAGTCACTGACTAACCCCGGAACCACACAGCTTCCCGTTC

TTGTCCAGGAGAAGGGAGGGTCAGTGACTGATTGGGGCCTTGGTGTGTCGAAGGGCAAG

TCAGCTCCACAAACTTGGTGCCA-3′

AGTCGAGGTGTTTGAACCACGGT-5′

Primer sequences used for ^{32}P-plasmid labeling: P2 primer should be labeled with ^{32}P

VEGFP2: 5’- TGG CAC CAA GTT TGT GGA GCT - 3’
VEGFP1: 5’- CTC AGT TCC CTG GCA ACA TCT - 3’

Plasmid insert sequence for pCSJ-FLT1, which was used for DNAse I footprint titration K_a measurements in Chapter 3. Primer sequences were standard for PUC19 inserts.

Top Strand: 5′-GAT CCC TTT GAG GTA ATG GAG ACA TAA TTG AGG AAC
Bottom Strand: 3′— GG AAA CTC CAT TAC CTC TGT ATT AAC TCC TTG

AAC GTG GAA TTA GTG TCA TAG CAA ATG ATC TAG GGC CT - 3′
TTG CAC CTT AAT CAC AGT ATC GTT TAC TAG ATC CCG GAT CGA -5′

Appendix B

Primer and Oligonucleotide Sequences for Quantitative Real-Time RT-PCR and Melting–Temperature Measurements

Primer and Oligonucleotide Sequences for Chapters 2-4

Primers for Quantitative Real-Time RT-PCR mRNA measurement

Chapters 2-4

VEGF:
VEGF-99-L: 5’ - AGG GCA GAA TCA TCA CGA AG - 3’
VEGF-99-R: 5’ - GGG TAC TCC TGG AAG ATG TCC - 3’

PSA:
PSA(KLK3) L: 5’ - TCT GCG GCG GTG TTC TG - 3’
PSA(KLK3) R: 5’ - GCC GAC CCA GCA AGA TCA - 3’

GUSB:
GUSB-human-L: 5’ - CTC ATT TGG AAT TTT GCC GAT T - 3’
GUSB-human-R: 5’ - CCG AGT GAA GAT CCC CTT TTT A - 3’

Primers for Chromatin Immunoprecipitation

Chapters 2 and 3

VEGF	L 5’- CCTTTGGGTTTTGCCAGA-3’
	R 5’- CCAAGTTTGTGGAGCTGA-3’

Oligonucleotide Sequences for Tm Measurements

Chapter 4

HRE:
HRE Top: 5’ - GTG CAT ACG TGG GC - 3’
HRE Bottom: 5’ - GCC CAC GTA TGC AC - 3’

ARE:
ARE Top: 5’ - TTG CAG AAC AGC AA - 3’
ARE Bottom: 5’ - TTG CTG TTC TGC AA - 3’