

# **Completion of a Programmable DNA-Binding Small Molecule Library**

Thesis by

Carey Frank Hsu

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended October 30, 2008)

© 2009

Carey Frank Hsu

All Rights Reserved

*For my family*

## **Acknowledgements**

I would like to thank my research advisor, Professor Peter Dervan, from whom I have learned so much about research and life. Thank you for giving me the opportunity to work with you. In particular, I treasured the time that I spent as your teaching assistant, learning how to be a teacher. I would also like to thank the members of my thesis committee, Professor Linda Hsieh-Wilson, Professor Stephen Mayo, and Professor Brian Stoltz, for their support.

The Dervan group is a wonderful environment for research. Above all, the people make it a very special place. I would like to acknowledge my collaborators, John Phillips, James Puckett, Daniel Harki, Nicholas Nickols, and Michael Marques, who made this work possible. I would like to thank James Puckett, Ryan Stafford, and Justin Cohen, for their friendship.

I would like to acknowledge the members of the Dervan group: graduate students Timothy Best, Michael Brochu, David Chenoweth, Raymond Doss, Benjamin Edelson, Michelle Farkas, Eric Fechter, Claire Jacobs, Benjamin Li, David Montgomery, Katy Muzikar, Julie Poposki, Adam Poulin-Kerstien, James Sanchez, Sherry Tsai, and Fei Yang; and postdoctoral scholars Hans-Dieter Arndt, Christian Dose, Mareike Goeritz, Bogdan Olenyuk, and Anne Viger.

Finally, I would like to dedicate this work to my family.

## Abstract

Hairpin pyrrole-imidazole (Py-Im) polyamides are programmable oligomers that bind the DNA minor groove in a sequence-specific manner with affinities comparable to those of natural DNA-binding proteins. These cell-permeable small molecules have been shown to enter the nuclei of live cells and downregulate endogenous gene expression. We complete here a library of 27 hairpin Py-Im polyamides that bind 7-base-pair sequences of the general form 5'-WWGNNNW-3' (where W = A or T, N = W, G, or C). A table of binding affinities and sequence contexts for this completed 27-member library has been assembled for the benefit of the chemical biology community interested in molecular control of transcription. Quantitative fluorescence-based methods have been developed to determine the nuclear concentration of polyamide-fluorescein conjugates in cell culture. Confocal laser scanning microscopy and flow cytometry techniques are utilized to plot calibration curves, from which the nuclear concentration can be interpolated. Although confocal microscopy and flow cytometry generate disparate values, taken together these experiments suggest that the polyamide concentration inside the cell nucleus is lower than it is outside the cell. To further our understanding of C-terminal tail linkage effects on sequence specificity, the equilibrium association constants of hairpin polyamide conjugates were measured by quantitative DNase I footprint titration experiments. These results indicate that linkers and functional R groups on the tails of hairpin polyamide conjugates have recognition properties that should be considered in the design of these molecules to target DNA binding sites. Furthermore, these  $\beta$ -alanine- $C_3$ -linked polyamide conjugates are shown to decrease hypoxia-inducible transcription of vascular endothelial growth factor (VEGF) in cultured HeLa cells. In addition, polyamide conjugates designed to target the Oct4 octamer DNA element modulate the expression levels of Oct4-driven genes in P19 mouse embryonal carcinoma cells and R1 mouse embryonic stem (ES) cells.

## Table of Contents

	Page
Acknowledgements.....	iv
Abstract.....	v
Table of Contents.....	vi
List of Figures and Tables.....	vii
 Chapter 1     Introduction.....	 1
 Chapter 2     Completion of a Programmable DNA-Binding Small Molecule Library.....	  15
 Chapter 3     Quantitating the Concentration of Py-Im Polyamide-Fluorescein Conjugates in Live Cells.....	  38
 Chapter 4A    Effects of Tail Linkages on the Sequence-Specific Recognition of DNA by Hairpin Polyamide Conjugates.....	  59
 Chapter 4B    Inhibition of VEGF with $\beta$ -Alanine-C <sub>3</sub> -Linked Hairpin Polyamide Conjugates.....	  81
 Chapter 5     Disrupting the Oct4 Octamer DNA Element with Pyrrole-Imidazole Polyamide Conjugates.....	  98
 Appendix     Synthetic Experiments toward Fluorinated Heterocycles.....	 139

## List of Figures and Tables

	Page
 <b>Chapter 1</b>	
Figure 1.1    The structure of DNA.....	2
Figure 1.2    Structures of distamycin A bound to DNA as 1:1 and 2:1 complexes.....	4
Figure 1.3    Molecular recognition of the minor groove of DNA.....	5
Figure 1.4    Structure of ImHpPyPy- $\beta$ -Dp bound to DNA as a 2:1 complex.....	6
Figure 1.5    Binding model for hairpin polyamides.....	8
Figure 1.6    Synthesis of polyamides on solid support.....	9
Figure 1.7    Nuclear localization of polyamide-fluorophore conjugates.....	10
 <b>Chapter 2</b>	
Figure 2.1    Model for the complex formed between hairpin polyamide <b>24</b> and its match DNA sequence.....	19
Figure 2.2    Plasmid design (pCFH2, pCFH3, pCFH4, pCFH5, pPh2, and pMFST)...	21
Figure 2.3    DNase I footprint titrations ( <b>13</b> , <b>14</b> , <b>16</b> , <b>17</b> , <b>19</b> , <b>21</b> , <b>24</b> , and <b>27</b> ).....	22
Table 2.1    Equilibrium association constants.....	24
Table 2.2    Equilibrium association constants.....	25
Table 2.3    Equilibrium association constants.....	27
Figure 2.4    DNase I footprint titrations ( <b>6</b> , <b>12</b> , <b>20</b> , <b>22</b> , and <b>23</b> ).....	29

### Chapter 3

Figure 3.1	Structures of polyamide-fluorescein conjugates <b>1</b> and <b>2</b> .....	42
Figure 3.2	DNase I footprint titration experiments for polyamides <b>1</b> and <b>2</b> .....	43
Figure 3.3	Overlaid fluorescence emission spectra for polyamides <b>1</b> and <b>2</b> .....	44
Figure 3.4	Sample calibration curve for confocal microscopy.....	46
Figure 3.5	Confocal laser scanning microscopy images of calibration standards.....	46
Table 3.1	Calculated nuclear concentration from confocal microscopy.....	47
Figure 3.6	Sample calibration curve for flow cytometry.....	49
Table 3.2	Calculated cellular concentration from flow cytometry.....	50

### Chapter 4A

Figure 4.1	Structures of polyamide conjugates <b>1-10</b> .....	63
Figure 4.2	Designed binding sites on plasmids pCFH6 and pJWP6.....	69
Figure 4.3	Quantitative DNase I footprint titration experiments for polyamide conjugates <b>1</b> , <b>2</b> , and <b>5</b> on plasmid pCFH6.....	70
Table 4.1	Equilibrium association constants on plasmid pCFH6.....	71
Figure 4.4	Quantitative DNase I footprint titration experiments for polyamide conjugates <b>1</b> , <b>2</b> , and <b>5</b> on plasmid pJWP6.....	73
Table 4.2	Equilibrium association constants on plasmid pJWP6.....	74
Table 4.3	Nuclear localization of polyamide-fluorescein conjugates.....	75



## Chapter 4B

Figure 4.5	Structures of polyamide conjugates <b>1-13</b> .....	84
Figure 4.6	DNase I footprint titration experiments for polyamides <b>2</b> and <b>5</b> .....	90
Table 4.4	Equilibrium association constants for polyamides <b>2-8</b> .....	91
Figure 4.7	Nuclear localization of polyamides <b>2, 3</b> , and <b>7</b> in HeLa cells.....	92
Figure 4.8	Relative levels of VEGF mRNA in HeLa cells at 1 $\mu$ M polyamide.....	93
Figure 4.9	Relative levels of VEGF mRNA in HeLa cells at 0.2 and 1 $\mu$ M.....	94

## Chapter 5

Figure 5.1	Crystal structure of a POU/HMG/DNA ternary complex.....	101
Figure 5.2	Binding model for the hairpin polyamide bound to its target DNA sequence within the enhancer regions of Oct4 target genes.....	103
Figure 5.3	Structures of polyamide conjugates <b>1-7</b> .....	105
Figure 5.4	Quantitative DNase I footprint titrations for polyamide <b>1</b> .....	107
Figure 5.5	DNA melting temperature analysis for polyamides <b>1-5</b> .....	108
Figure 5.6	Nuclear localization of polyamide-fluorescein conjugate <b>1</b> .....	109
Figure 5.7	Time course for 48-hour incubation experiments with P19 cells.....	110
Figure 5.8	Quantitative RT-PCR experiments with polyamides <b>2</b> and <b>6</b> .....	111
Figure 5.9	Quantitative RT-PCR experiments with polyamides <b>2-5</b> .....	112
Figure 5.10	Quantitative RT-PCR experiments with polyamides <b>4</b> and <b>7</b> .....	113
Figure 5.11	Structures of polyamide conjugates <b>8-14</b> .....	115
Figure 5.12	Time course for 48-hour and 72-hour incubation experiments with R1 mouse embryonic stem cells.....	116

Figure 5.13	Quantitative RT-PCR experiments with polyamides <b>4</b> and <b>7</b> .....	117
Figure 5.14	Quantitative RT-PCR experiments with polyamide <b>4</b> .....	118
Figure 5.15	Images of R1 mouse embryonic stem cells after 24-hour treatment.....	120
Figure 5.16	Images of R1 mouse embryonic stem cells after 48-hour treatment.....	121
Figure 5.17	Images of R1 mouse embryonic stem cells after 72-hour treatment.....	122
Figure 5.18	The partial sequence of plasmid phOCT4-EGFP1.....	124

## Appendix

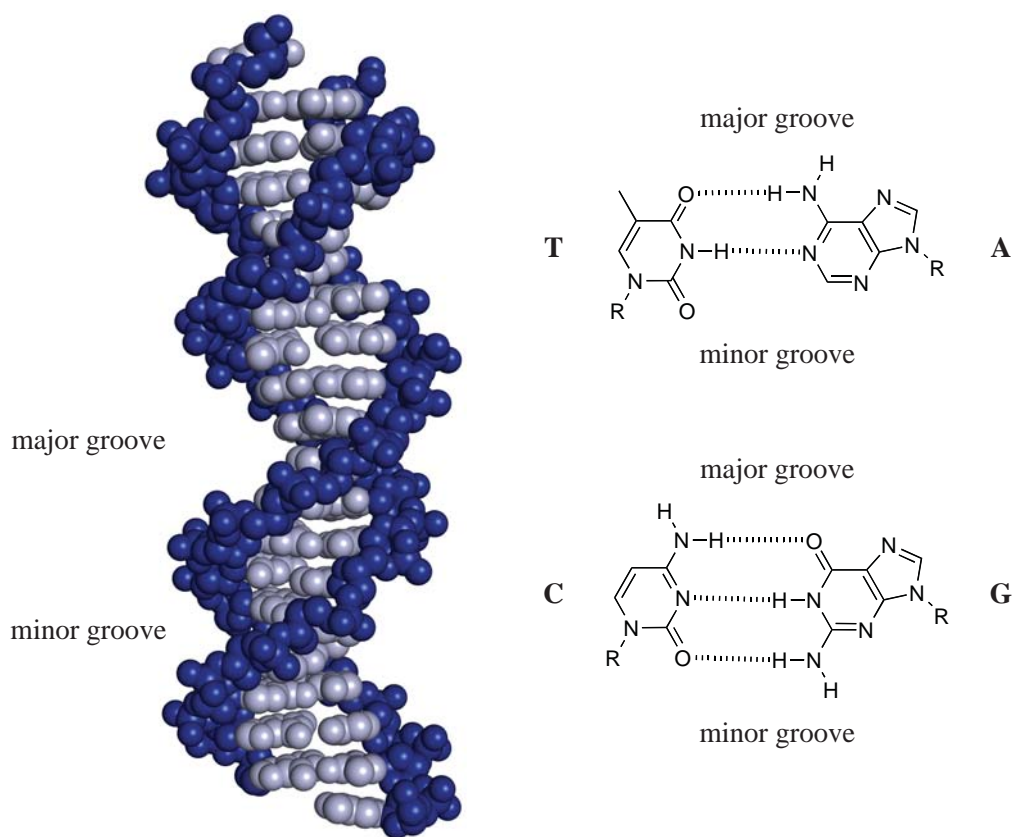
Figure A.1	Structure of the pyrrole-fluorobenzimidazole dimer <b>2</b> .....	141
Figure A.2	Structure of the key synthetic target fluorodiamine <b>4</b> .....	141
Figure A.3	Reactions attempted to achieve the desired tetrasubstitution of <b>4</b> .....	141

## **Chapter 1**

### **Introduction**

### 1.1. Deoxyribonucleic acid (DNA)

DNA is the biological macromolecule that encodes the genetic information necessary for life. The molecule consists of four different bases on a phosphodiester-linked deoxyribose sugar backbone. The intertwined strands of two DNA molecules form a double helix with hydrogen bonds between the Watson-Crick base pairs such that thymine (T) pairs with adenine (A), and cytosine (C) pairs with guanine (G). The structure of B-form DNA is a right-handed helix containing ten base pairs per turn, and the plane of each hydrogen-bonded base pair lies perpendicular to the helix axis.<sup>1,2</sup> A wide major groove and a narrow minor groove line the helix and are available for binding (Figure 1.1).<sup>2</sup> The human genome contains roughly three billion base pairs that encode 20,000



**Figure 1.1.** The structure of DNA. (left) The phosphodiester-linked deoxyribose backbone is shown in blue, and the Watson-Crick base pairs are shown in gray (PDB accession code: 1YSA<sup>5</sup>). (right) The chemical structures of the DNA bases thymine (T), adenine (A), cytosine (C), and guanine (G) are shown as hydrogen-bonded base pairs.<sup>40</sup>

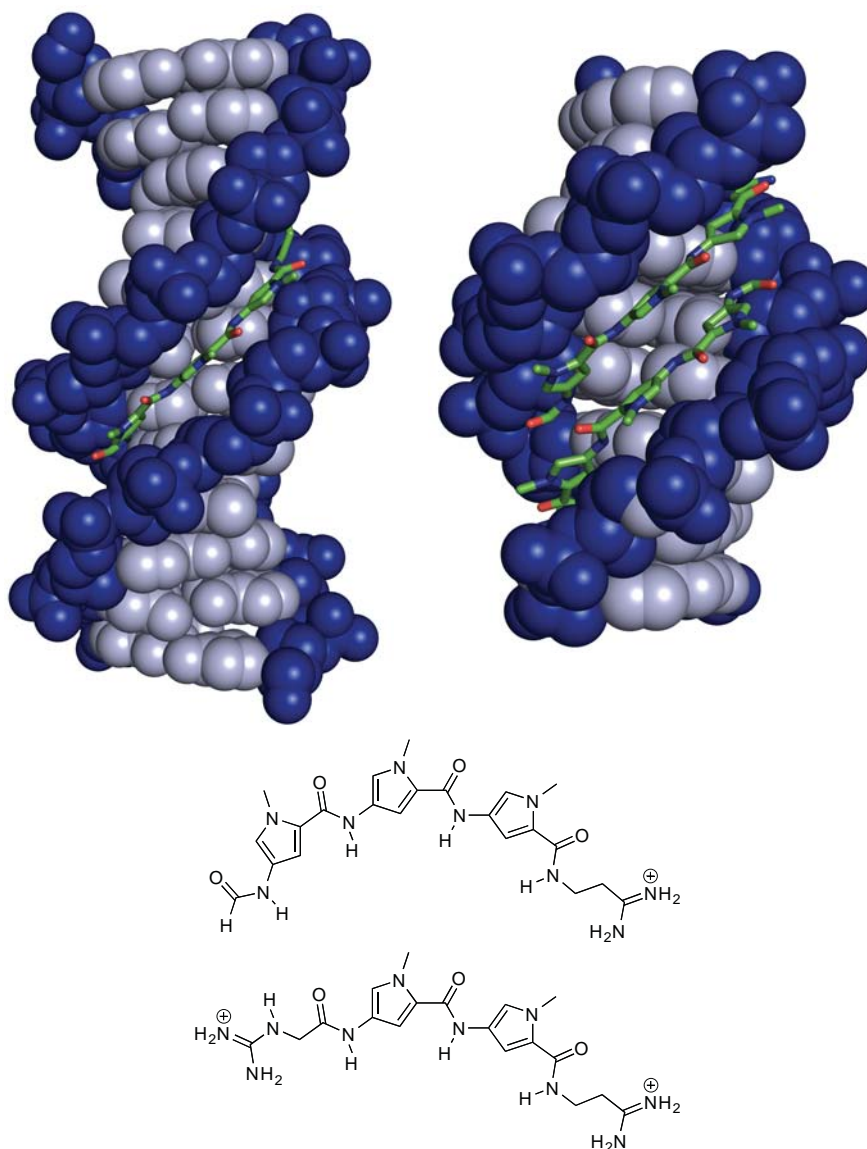
to 25,000 protein-coding genes.<sup>3</sup> In addition to DNA sequences that code for proteins, genes are associated with regulatory regions that control transcription. In the context of the living cell, DNA-binding proteins and the interactions between proteins modulate gene expression.<sup>4,5</sup>

## 1.2. DNA-binding small molecules

Nature has evolved a number of small molecules, including actinomycin, echinomycin, and calicheamicin, that bind to specific sequences of DNA. Distamycin A and netropsin are A,T-binding oligopeptidic antibiotic drugs of bacterial origin whose structures contain three and two *N*-methylpyrrole (Py) carboxamide units, respectively.<sup>6-9</sup> The X-ray crystal structure of the 1:1 complex between netropsin and 5'-CGCGAATTCGCG-3' shows that the crescent-shaped natural product binds in the minor groove of B-form DNA by displacing water molecules on the spine of hydration.<sup>10,11</sup> The NMR structure of the 2:1 complex between distamycin and 5'-CGCAAATTGGC-3' indicates that the two distamycin molecules bind the central 5'-AAATT-3' region in an antiparallel orientation with expansion of the minor groove relative to the 1:1 ligand-DNA complex.<sup>12</sup> The structures of the 1:1 and 2:1 distamycin-DNA complexes are depicted in Figure 1.2.<sup>13,14</sup>

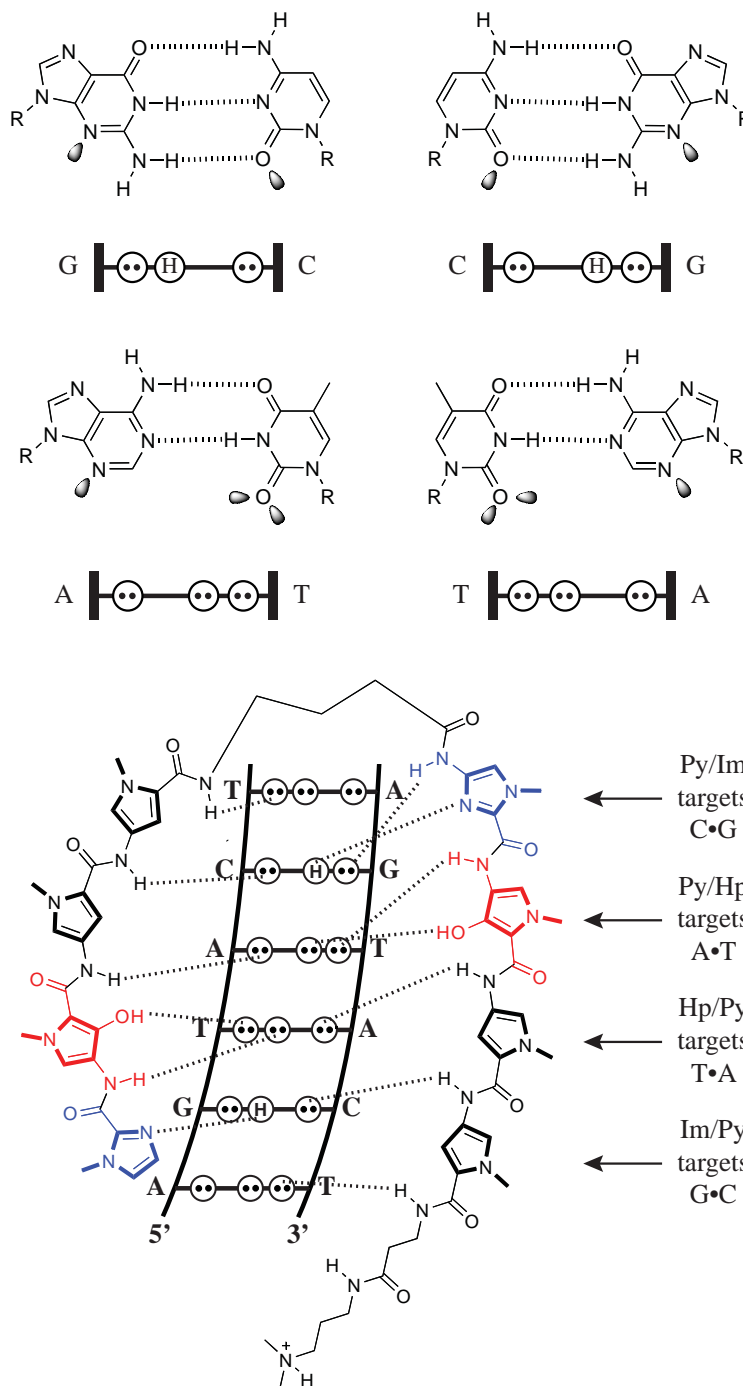
## 1.3. Recognition of the DNA minor groove by pyrrole-imidazole polyamides

A modular code for molecular recognition of the four Watson-Crick base pairs has been empirically determined using pyrrole-imidazole polyamides.<sup>15,16</sup> These synthetic ligands recognize the chemical features presented in the minor groove of DNA as a form of digital readout (Figure 1.3). The *N*-methylimidazole (Im) analog of distamycin binds the five base-pair sequence 5'-WGWCW-3' (where W = A or T) by formation of a 2:1 polyamide-DNA complex.<sup>17,18</sup> The two ImPyPy molecules are stacked together as an antiparallel dimer in the DNA minor groove with Im paired unsymmetrically across from Py to recognize the G•C base pair. Hydrogen bonding between the Im nitrogen and the



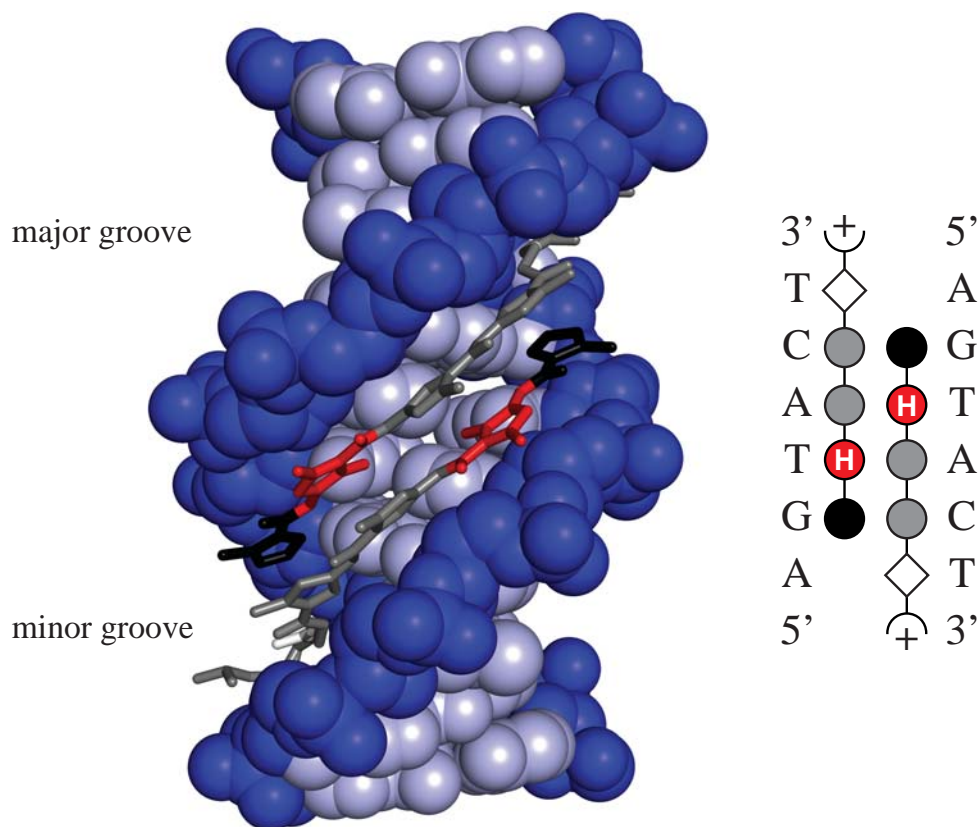
**Figure 1.2.** Structures of distamycin A bound to DNA as 1:1 and 2:1 complexes (PDB accession codes: 2DND<sup>13</sup> and 378D<sup>14</sup>). The DNA backbone is shown in blue, the base pairs are gray, and distamycin is shown in stick representation. The chemical structures of distamycin A (top) and netropsin (bottom) are shown below the crystal structures.<sup>40</sup>

exocyclic amine of guanine helps to confer specificity for G•C over C•G, A•T and T•A.<sup>19</sup> In this way, the Im/Py pair recognizes G•C and the Py/Im pair codes for C•G, while the degenerate Py/Py pair does not discriminate between A•T and T•A. To complete the task of recognizing all four Watson-Crick base pairs, *N*-methylhydroxypyrrole (Hp) was paired



**Figure 1.3.** Molecular recognition of the minor groove of DNA. (top) Minor groove hydrogen bonding patterns of Watson-Crick base pairs. Circles with dots represent lone pairs of N(3) of purines and O(2) of pyrimidines, and circles containing an H represent the 2-amino group of guanine. The R group represents the sugar-phosphate backbone of DNA. Electron pairs projecting into the minor groove are represented as shaded orbitals. (bottom) Binding model for the complex formed between ImHpPyPy- $\gamma$ -ImHpPyPy- $\beta$ -Dp and a 5'-AGTACT-3' sequence. Putative hydrogen bonds are shown as dashed lines.<sup>16</sup>

across from Py to recognize T•A over A•T.<sup>20-22</sup> The X-ray crystal structure of the 2:1 complex between ImHpPyPy and 5'-CCAGTACTGG-3' provides structural verification of the pairing rules (Figure 1.4).



**Figure 1.4.** Crystal structure of the 2:1 complex formed between ImHpPyPy-β-Dp and the 5'-CCAGTACTGG-3' sequence (PDB accession code: 407D<sup>21</sup>). Polyamides are shown at right in a ball-and-stick model, where a black circle represents Im, a red circle with an H represents Hp, a gray circle represents Py, a white diamond represents β-alanine, and a half circle with a plus represents Dp.



#### 1.4. Hairpin polyamide motif

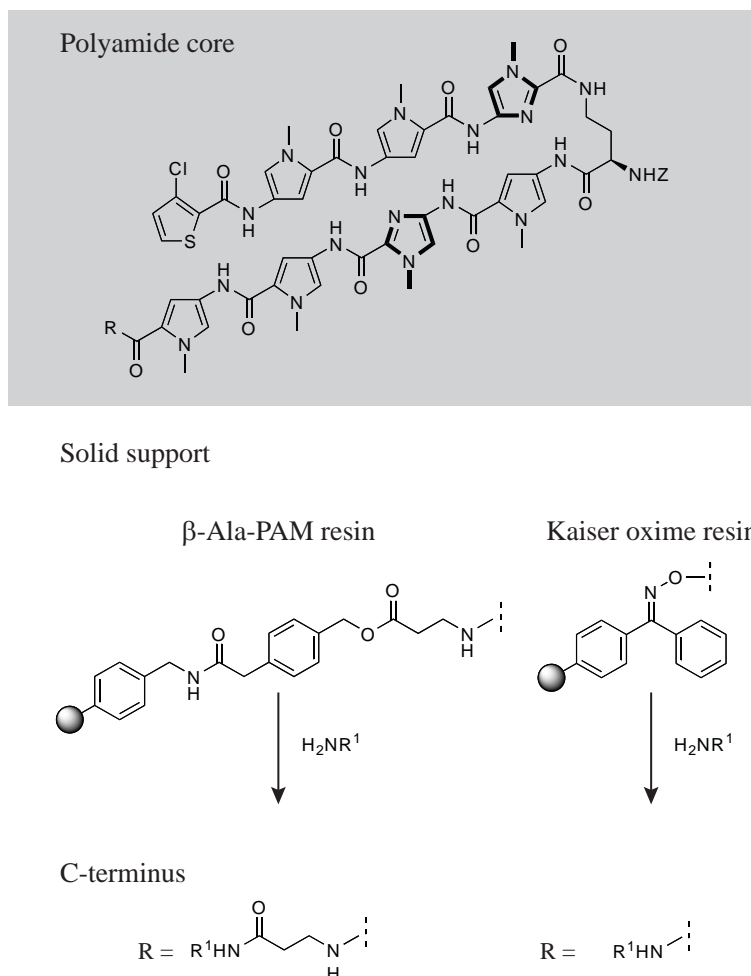
The pre-organization of two polyamide subunits by a covalent linkage improves the binding affinity and sequence specificity of these ligands.<sup>23-26</sup> While several different binding motifs have been explored, the hairpin motif provides excellent affinity and specificity for DNA and can be accessed through a facile synthetic route.<sup>23</sup> In the hairpin motif, a  $\gamma$ -aminobutyric acid ( $\gamma$ ) covalently links the carboxylic terminus of one polyamide with the amino terminus of the other. Eight-ring hairpin polyamides bind their six base-pair DNA match sequences with subnanomolar affinities.<sup>27,28</sup>

Structural features of the hairpin motif contribute to sequence-specific DNA recognition (Figure 1.5). The  $\gamma$  residue that forms the turn of the hairpin recognizes A•T and T•A base pairs. A chiral 2,4-diaminobutyric acid improves affinity while retaining specificity for A•T and T•A.<sup>29</sup> The increased binding affinity can be attributed to electrostatic interactions between the negatively charged DNA backbone and the positively charged free amine on the turn. Furthermore, the chiral turn allows stereochemical control, enforcing the N→C polyamide alignment with respect to the 5'→3' orientation of the adjacent DNA strand.<sup>29,30</sup>

The microstructure of DNA is sequence-dependent, and polyamides composed of Py and Im rings are overly curved compared to the DNA minor groove. Aliphatic  $\beta$ -alanine residues have been incorporated at internal positions to reset the register between contiguous aromatic ring pairings.<sup>31</sup> The  $\beta$ /Py and Py/ $\beta$  pairs, as well as the  $\beta$ / $\beta$  degenerate pair, recognize A•T and T•A base pairs.<sup>31,32</sup> The  $\beta$ -alanine residue on the tail, as well as the 3-(dimethylamino)propylamine (Dp) tail itself, both code for A•T and T•A.<sup>33</sup> Hairpin polyamide design typically includes an Im/Py pairing at the hairpin N-terminus in order to impart G•C specificity at this position. However, when 3-chlorothiophene (Ct) is paired across from Py at the N-terminus, the Ct/Py pair exhibits specificity for T•A.<sup>34</sup> In this way, the eight-ring hairpin polyamide CtPyPyIm-(R)<sup>RHN</sup> $\gamma$ -PyImPyPy- $\beta$ -Dp recognizes the seven base-pair sequence 5'-WWTWCGW-3' in the DNA minor groove.

Methods for solid-phase synthesis of hairpin polyamides have been developed using standard Boc coupling chemistry.<sup>35,36</sup> Aminolysis furnishes the polyamide, which can be further derivatized on the turn or tail residues. The selection of either  $\beta$ -Ala-PAM or Kaiser oxime resin yields different C-terminal groups following cleavage with a primary

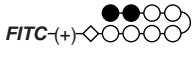
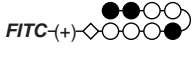
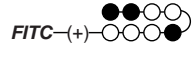
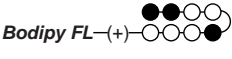
amine (Figure 1.6). The use of  $\beta$ -Ala-PAM resin leaves a residual A,T-specific  $\beta$ -alanine moiety at the C-terminus, while oxime resin produces a shorter polyamide. Access to a diverse range of polyamide binding motifs and linkers has proven valuable for experiments in cell culture.



**Figure 1.6.** Synthesis of polyamides on solid support.  $\beta$ -Ala-PAM resin and Kaiser oxime resin yield different C-terminal groups following cleavage with a primary amine. Z = H, Ac, or Boc, if further derivatization is required.<sup>16</sup>

### 1.5. Nuclear localization of hairpin polyamide conjugates in cell culture studies

Confocal microscopy has been used to determine the uptake profiles of polyamide-fluorophore conjugates in a range of cell lines.<sup>37-40</sup> A brief excerpt is presented in Figure 1.7. Polyamide **1** localizes to the nucleus, while polyamide **2** is excluded from the nucleus, which reveals that uptake is core-dependent. Removal of the  $\beta$ -alanine residue to produce

		DLD-1	HeLa	MCF-7	SK-BR-3	786-O	293	LN-CaP	PC3	MEL	NB4	Jurkat	CCRF-CEM	MEG-01
	1	+	++	++	++	+	+	++	++	+	++	++	++	++
	2	-	--	--	--	--	--	--	--	--	--	--	--	--
	3	+	++	++	++	++	++	++	++	++	++	++	++	++
	4	-	-	-	-	--	--	--	-	-	-	-	--	-

**Figure 1.7.** Uptake profile of polyamide-fluorophore conjugates **1-4** in 13 cell lines.<sup>38</sup> ++ indicates that nuclear staining exceeds that of the medium; +, nuclear staining less than or equal to that of the medium, but still prominent; -, very little nuclear staining, with the most fluorescence seen in the cytoplasm and/or medium; --, no nuclear staining.<sup>38,39</sup> In the ball-and-stick model, a black circle represents Im, a white circle represents Py, a white diamond represents  $\beta$ -alanine, “(+)” represents the 3,3’-diamino-*N*-methyldipropylamine linker, and a half-circle represents the  $\gamma$ -aminobutyric acid turn residue.

polyamide **3** rescues uptake, suggesting that the incorporation of  $\beta$ -alanine may be lethal for uptake. The poor nuclear localization of polyamide **4** indicates that fluorophore selection is crucial.<sup>41</sup> Based on these results, fluorescein has been the fluorescent dye of choice in recent experiments.

The gained insights into cellular uptake properties of polyamides have proven instrumental in the successful design of gene regulation studies. A polyamide-fluorescein conjugate has been shown to decrease hypoxia-inducible transcription of vascular endothelial growth factor (VEGF) in cultured HeLa cells.<sup>42,43</sup> The binding of this polyamide to the hypoxia response element (HRE) disrupts the binding of hypoxia-inducible factor to the HRE. Inhibition of VEGF expression has been used to assay the nuclear localization of HRE-targeted polyamides. These studies have shown that the fluorescein dye can be minimized to a isophthalic acid moiety, yielding higher-affinity conjugates that exhibit comparable biological activity.<sup>44</sup> A hairpin polyamide-isophthalic acid conjugate has been shown to downregulate prostate-specific antigen (PSA) expression in prostate cancer cells by binding the androgen response element in the PSA promoter.<sup>45</sup> This prevents the binding of androgen receptor (AR) and inhibits androgen-induced PSA expression.

## 1.6. Scope of this work

In Chapter 2, we complete a library of 27 hairpin Py-Im polyamides that bind 7-base-pair sequences of the general form 5'-WWGNNNW-3' (where W = A or T, N = W, G, or C). A table of binding affinities and sequence contexts for this completed 27-member library has been assembled for the benefit of the chemical biology community interested in molecular control of transcription. In Chapter 3, quantitative fluorescence-based methods have been developed to determine the nuclear concentration of polyamide-fluorescein conjugates in cell culture. Confocal laser scanning microscopy and flow cytometry techniques are utilized to plot calibration curves, from which the nuclear concentration can be interpolated. Although confocal microscopy and flow cytometry generate disparate values, taken together these experiments suggest that the polyamide concentration inside the cell nucleus is lower than it is outside the cell. In Chapter 4, to further our understanding of C-terminal tail linkage effects on sequence specificity, the equilibrium association constants of hairpin polyamide conjugates were measured by quantitative DNase I footprint titration experiments. These results indicate that linkers and functional R groups on the tails of hairpin polyamide conjugates have recognition properties that should be considered in the design of these molecules to target DNA binding sites. Furthermore, these  $\beta$ -alanine- $C_3$ -linked polyamide conjugates are shown to decrease hypoxia-inducible transcription of vascular endothelial growth factor (VEGF) in cultured HeLa cells. In Chapter 5, polyamide conjugates designed to target the Oct4 octamer DNA element modulate the expression levels of Oct4-driven genes in P19 mouse embryonal carcinoma cells and R1 mouse embryonic stem (ES) cells.

## References

1. Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E., *Nature* **1980**, 287, 755-758.
2. Dickerson, R. E.; Drew, H. R.; Conner, B. N.; Wing, R. M.; Fratini, A. V.; Kopka, M. L., *Science* **1982**, 216, 475-485.
3. Collins, F. S.; Lander, E. S.; Rogers, J.; Waterston, R. H., *Nature* **2004**, 431, 931-945.
4. Pabo, C. O.; Sauer, R. T., *Annu. Rev. Biochem.* **1984**, 53, 293-321.
5. Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C., *Cell* **1992**, 71, 1223-1237.
6. Arcamone, F.; Nicoletti, V.; Penco, S.; Orezzi, P.; Pirelli, A., *Nature* **1964**, 203, 1064-&.
7. Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X., *J. Am. Chem. Soc.* **1951**, 73, 341-343.
8. Vantamelen, E. E.; White, D. M.; Kogon, I. C.; Powell, A. D. G., *J. Am. Chem. Soc.* **1956**, 78, 2157-2159.
9. Waller, C. W.; Wolf, C. F.; Stein, W. J.; Hutchings, B. L., *J. Am. Chem. Soc.* **1957**, 79, 1265-1266.
10. Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E., *Proc. Natl. Acad. Sci. U. S. A.* **1985**, 82, 1376-1380.
11. Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E., *J. Mol. Biol.* **1985**, 183, 553-563.
12. Pelton, J. G.; Wemmer, D. E., *Proc. Natl. Acad. Sci. U. S. A.* **1989**, 86, 5723-5727.
13. Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A., *Proc. Natl. Acad. Sci. U. S. A.* **1987**, 84, 8385-8389.
14. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M., *Acta Crystallogr. Sect. D-Biol. Crystallogr.* **1999**, 55, 602-609.
15. Dervan, P. B., *Bioorg. Med. Chem.* **2001**, 9, 2215-2235.

16. Dervan, P. B.; Edelson, B. S., *Curr. Opin. Struct. Biol.* **2003**, 13, 284-299.
17. Wade, W. S.; Mrksich, M.; Dervan, P. B., *J. Am. Chem. Soc.* **1992**, 114, 8783-8794.
18. Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B., *Proc. Natl. Acad. Sci. U. S. A.* **1992**, 89, 7586-7590.
19. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C., *Nat. Struct. Biol.* **1998**, 5, 104-109.
20. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *Nature* **1998**, 391, 468-471.
21. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *Science* **1998**, 282, 111-115.
22. Kielkopf, C. L.; Bremer, R. E.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *J. Mol. Biol.* **2000**, 295, 557-567.
23. Mrksich, M.; Parks, M. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1994**, 116, 7983-7988.
24. Greenberg, W. A.; Baird, E. E.; Dervan, P. B., *Chem.-Eur. J.* **1998**, 4, 796-805.
25. Melander, C.; Herman, D. M.; Dervan, P. B., *Chem.-Eur. J.* **2000**, 6, 4487-4497.
26. Heckel, A.; Dervan, P. B., *Chem.-Eur. J.* **2003**, 9, 3353-3366.
27. Trauger, J. W.; Baird, E. E.; Dervan, P. B., *Nature* **1996**, 382, 559-561.
28. Trauger, J. W.; Dervan, P. B., *Methods in Enzymology* **2001**, 340, 450-466.
29. Herman, D. M.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1998**, 120, 1382-1391.
30. White, S.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1997**, 119, 8756-8765.
31. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1998**, 120, 6219-6226.
32. Wang, C. C. C.; Ellervik, U.; Dervan, P. B., *Bioorg. Med. Chem.* **2001**, 9, 653-657.

33. Swalley, S. E.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1999**, 121, 1113-1120.
34. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., *Bioorg. Med. Chem.* **2003**, 11, 4333-4340.
35. Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1996**, 118, 6141-6146.
36. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., *Bioorg. Med. Chem.* **2002**, 10, 2767-2774.
37. Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B., *Bioorg. Med. Chem.* **2002**, 10, 3313-3318.
38. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., *Proc. Natl. Acad. Sci. U. S. A.* **2003**, 100, 12063-12068.
39. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., *Nucleic Acids Res.* **2004**, 32, 2802-2818.
40. Stafford, R. L. Ph.D. Thesis, California Institute of Technology, Pasadena, CA, 2008.
41. Crowley, K. S.; Phillion, D. P.; Woodard, S. S.; Schweitzer, B. A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J. K., *Bioorg. Med. Chem. Lett.* **2003**, 13, 1565-1570.
42. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B., *Proc. Natl. Acad. Sci. U. S. A.* **2004**, 101, 16768-16773.
43. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., *ACS Chem. Biol.* **2007**, 2, 561-571.
44. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., *Nucleic Acids Res.* **2007**, 35, 363-370.
45. Nickols, N. G.; Dervan, P. B., *Proc. Natl. Acad. Sci. U. S. A.* **2007**, 104, 10418-10423.



## Chapter 2

### **Completion of a Programmable DNA-Binding Small Molecule Library**

The text of this chapter was taken in part from a manuscript co-authored with John W. Phillips, John W. Trauger, Michelle E. Farkas, Jason M. Belitsky, Alexander Heckel, Bogdan Z. Olenyuk, James W. Puckett, Clay C. C. Wang, and Peter B. Dervan (Caltech).

(Hsu, C. F., Phillips, J. W., Trauger, J. W., Farkas, M. E., Belitsky, J. M., Heckel, A., Olenyuk, B. Z., Puckett, J. W., Wang, C. C. C., Dervan, P. B. (2007) *Tetrahedron* **63**, 6146-6151.)

**Abstract**

Hairpin pyrrole-imidazole (Py-Im) polyamides are programmable oligomers that bind the DNA minor groove in a sequence-specific manner with affinities comparable to those of natural DNA-binding proteins. These cell-permeable small molecules have been shown to enter the nuclei of live cells and downregulate endogenous gene expression. We complete here a library of 27 hairpin Py-Im polyamides that bind 7-base-pair sequences of the general form 5'-WWGNNNW-3' (where W = A or T, N = W, G, or C). Their equilibrium association constants ( $K_a$ ) range from  $K_a = 1 \times 10^8 \text{ M}^{-1}$  to  $4 \times 10^{10} \text{ M}^{-1}$  with good sequence specificity. A table of binding affinities and sequence contexts for this completed 27-member library has been assembled for the benefit of the chemical biology community interested in molecular control of transcription.

## 2.1. Introduction

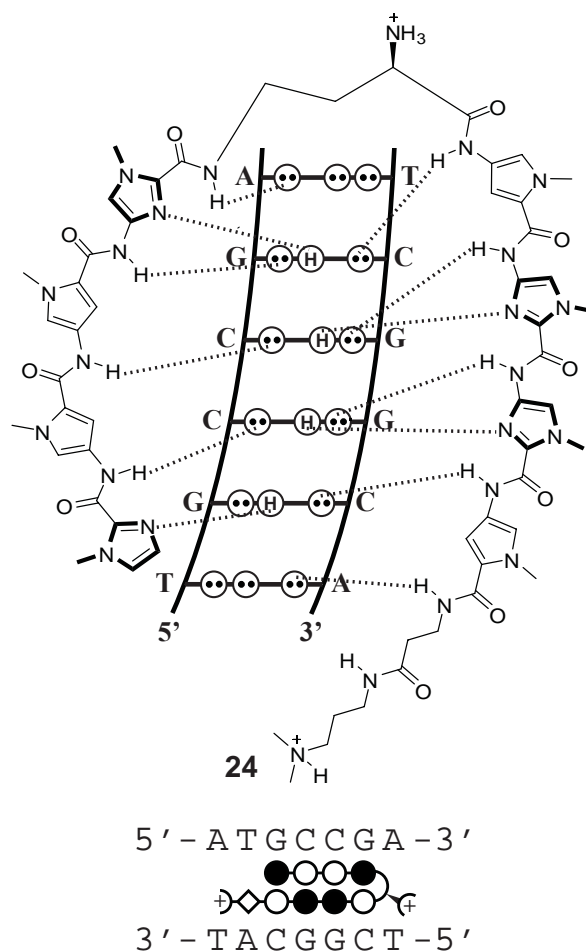
The biological applications of sequence-specific DNA-binding small molecules are a subject of intense research but are still far from being routine.<sup>1-11</sup> Py-Im polyamides have been shown to influence a number of protein-DNA interactions, demonstrating both repression and activation of gene expression. Confocal microscopy experiments have confirmed that polyamide-fluorophore conjugates traffic unaided to the nuclei of living cells.<sup>12-14</sup> Since many diseases are attributed to aberrant gene expression, the regulation of transcriptional pathways with small molecules could have an important effect on human medicine.<sup>15-17</sup> For researchers interested in selectively targeting protein-DNA interfaces in promoters of specific genes with small molecules, access to well-characterized libraries of polyamides which bind a repertoire of different sequences with high affinity and specificity could enable development in this area.

Polyamides constructed from *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methylhydroxypyrrole (Hp) amino acids comprise a class of synthetic ligands that bind within the minor groove of DNA in a sequence-specific manner.<sup>18,19</sup> Inspired by the natural products netropsin and distamycin A, these programmed molecules recognize the Watson-Crick base pairs (bp) according to a series of pairing rules, where aromatic heterocycles paired in an antiparallel fashion are able to discriminate one Watson-Crick base pair from the other three combinations. The Py/Py pair recognizes A,T over C,G.<sup>20</sup> The Im/Py pair distinguishes G•C from C•G.<sup>21,22</sup> The exocyclic amine of guanine presents steric hindrance to the C3-H of Py, while the N3 of Im accommodates the amine and accepts one of its hydrogen bonds. The Hp/Py pair discriminates T•A over A•T due to the steric fit of the hydroxy group protruding into the minor groove, thus completing the pairing rules.<sup>23,24</sup> NMR and X-ray crystallography studies reveal that the crescent-shaped polyamide side-by-side dimer binds B-form DNA, a remarkable example of shape-selective recognition of the deep minor groove of DNA.<sup>20,22,25,26</sup>

Within the framework of the pairing rules, covalent linkages between two antiparallel polyamide strands result in several possible structures, including the hairpin, cycle, H-pin, and U-pin binding motifs.<sup>27-30</sup> These linked structures show improved affinity and specificity when compared with the unlinked dimers. The eight-ring hairpin polyamide provides a good compromise between synthetic ease (linear vs. branched oligomers) and molecular recognition properties. In this binding motif, a  $\gamma$ -aminobutyric acid residue connects the carboxylic terminus of one strand to the amino terminus of the other.<sup>27</sup> The turn residue also serves as a DNA recognition element, as it has been shown to bind A•T and T•A base pairs with greater than 25-fold specificity over C•G and G•C, presumably for steric reasons.<sup>31</sup> Use of a chiral diaminobutyric acid turn residue increases the overall binding affinity of the molecule by 10-fold without a loss of sequence specificity.<sup>32</sup>

When hairpin polyamides are synthesized using solid phase methods on Boc- $\beta$ -Ala-PAM resin and cleaved with 3-dimethylamino-1-propylamine, the product contains a  $\beta$ -alanine residue and a dimethylaminopropylamide tail at the C-terminus. Both of these elements are specific for W (where W = A or T) over G•C and C•G, again for steric reasons. In the hairpin motif, the  $\beta$ -alanine residue exhibits greater than 210-fold specificity for A•T and T•A base pairs over G•C and C•G.<sup>31</sup> The dimethylaminopropylamide tail shows a 20-fold preference for A•T and T•A over G•C and C•G.<sup>31</sup> Therefore, an eight-ring hairpin polyamide can bind seven base pairs with specificity for W over the turn,  $\beta$ -alanine residue, and tail (Figure 2.1).

We have excluded Hp in this study in favor of more stable Py residue. The electron-rich Hp ring degrades in the presence of strong acid and free radical impurities. Based on the paradigm of unsymmetrical ring pairings, more robust rings such as the hydroxybenzimidazole/Py pair can replace Hp for T•A recognition.<sup>33</sup> In our work, hairpin polyamide design typically includes an Im/Py pairing at the hairpin terminus in order to impart G•C specificity at this position. As a result, eight-ring pyrrole-imidazole hairpin polyamides can specifically bind 7-bp sequences of the general form 5'-WWGNNNW-



**Figure 2.1.** (Top) Model for the complex formed between hairpin polyamide **24** and its match DNA sequence. Circles with two dots represent the lone pairs of N(3) of purines and O(2) of pyrimidines. Circles containing an H represent the N(2) hydrogen of G. Hydrogen bonds are illustrated by dotted lines. (Bottom) Ball-and-stick binding model for the hairpin motif with the polyamide bound to its target DNA sequence. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits.

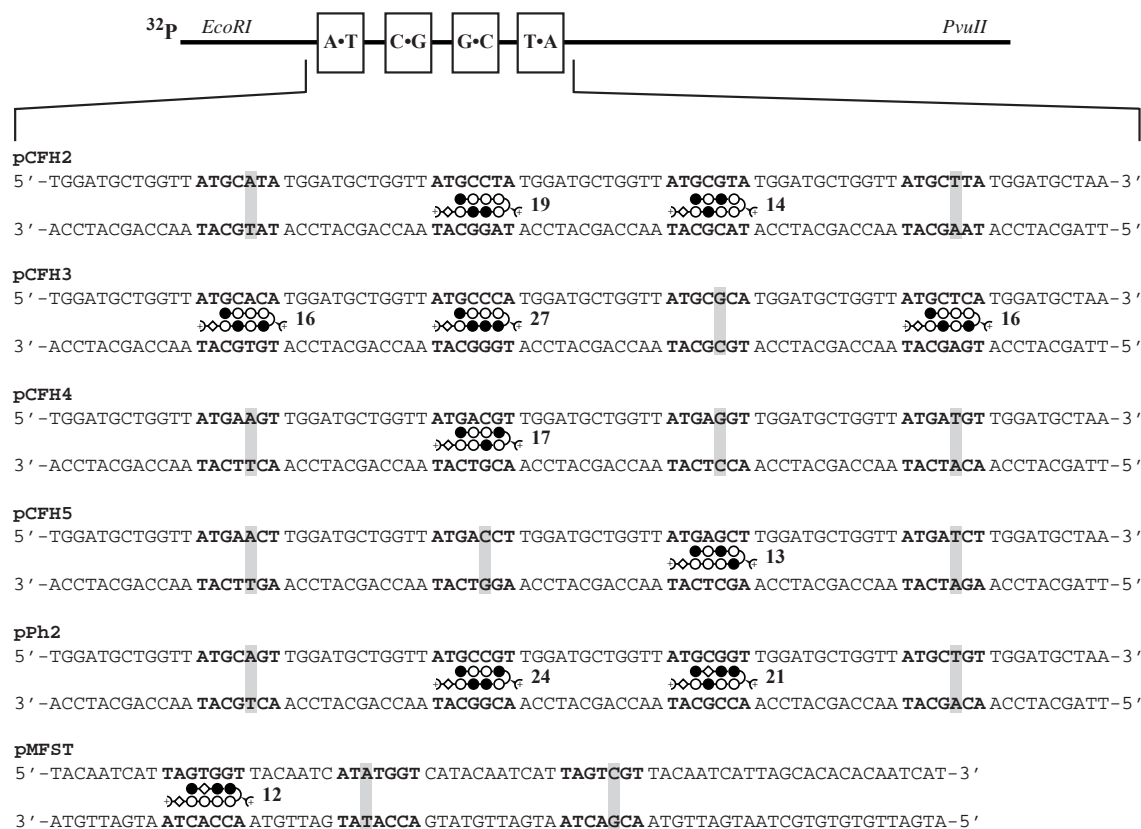
3' (where N = W, G, or C). There are 27 possible permutations that fall within these guidelines. Our group has published the energetics of 11 hairpin polyamides binding 11 distinct cognate sequences. Over half of the 27 sequences remain unreported. Searching our theses and notebooks confirmed 7 additional previously unpublished characterizations. Our studies during the past decade left only 9 out of 27 sequences unexamined; these are: 5'-WWGCGWW-3', 5'-WWGCCWW-3', 5'-WWGCWCW-3', 5'-WWGCCCW-3', 5'-

WWGWCGW-3', 5'-WWGWGCW-3', 5'-WWGCGGW-3', 5'-WWGCCGW-3', and 5'-WWGWGGW-3'. To complete the library, we have synthesized nine hairpin polyamides designed to target these remaining sequences and assayed their binding affinity and sequence specificity by DNase I footprint titration experiments. In this study, we complete the table of polyamide-DNA binding affinities, indicating each of the 27 general DNA sequences and an eight-ring hairpin polyamide that sequence specifically binds that 7-bp site. We hope that this centralized source of previously unpublished data proves helpful for other research groups currently modulating protein-DNA interfaces with DNA-binding small molecules.

## 2.2. Results

**2.2.1. Polyamide synthesis.** Nine polyamides **12**, **13**, **14**, **16**, **17**, **19**, **21**, **24**, and **27** were synthesized on Boc- $\beta$ -Ala-PAM resin according to published manual solid-phase synthesis protocols.<sup>34</sup> After cleavage with 3-dimethylamino-1-propylamine and reverse-phase HPLC purification, polyamides were characterized by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization/time-of-flight mass spectrometry.

**2.2.2. DNA binding energetics.** Quantitative DNase I footprint titration experiments (10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub>, pH 7.0, 22°C) were performed on the 295 bp, 5'-<sup>32</sup>P-end-labeled PCR product of plasmids pCFH2, pCFH3, pCFH4, pCFH5, pPh2, and pMFST.<sup>35</sup> Each of the nine polyamides was assayed on a plasmid containing its 7-bp match site according to the pairing rules, as well as three formal mismatch binding sites (Table 2.1 and Figure 2.2). The energetics of polyamide binding in the minor groove of DNA can be calculated from the Hill equation isotherms following DNase I cleavage and gel separation of the fragments. The equilibrium association constants ( $K_a$ ) determined in this way provide a quantitative measure of polyamide affinity



**Figure 2.2.** Plasmid design for pCFH2, pCFH3, pCFH4, pCFH5, pPh2, and pMFST, indicating the designed match and mismatch sites for hairpin polyamides **12**, **13**, **14**, **16**, **17**, **19**, **21**, **24**, and **27**. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits.

at a given DNA binding site. Comparing these constants across the four potential binding sites allows a relative measure of specificity for each base pair at the targeted position.

The nine polyamides **12**, **13**, **14**, **16**, **17**, **19**, **21**, **24** and **27** bind their respective 7-bp match sites with equilibrium association constants that range from  $K_a = 7 \times 10^8 \text{ M}^{-1}$  to  $4 \times 10^{10} \text{ M}^{-1}$  (Table 2.1 and Figure 2.3). The sequence specificities ( $K_{\text{match}}/K_{\text{single bp mismatch}}$ ) for these compounds vary from 4-fold to greater than 100-fold, further validating the pairing rules.

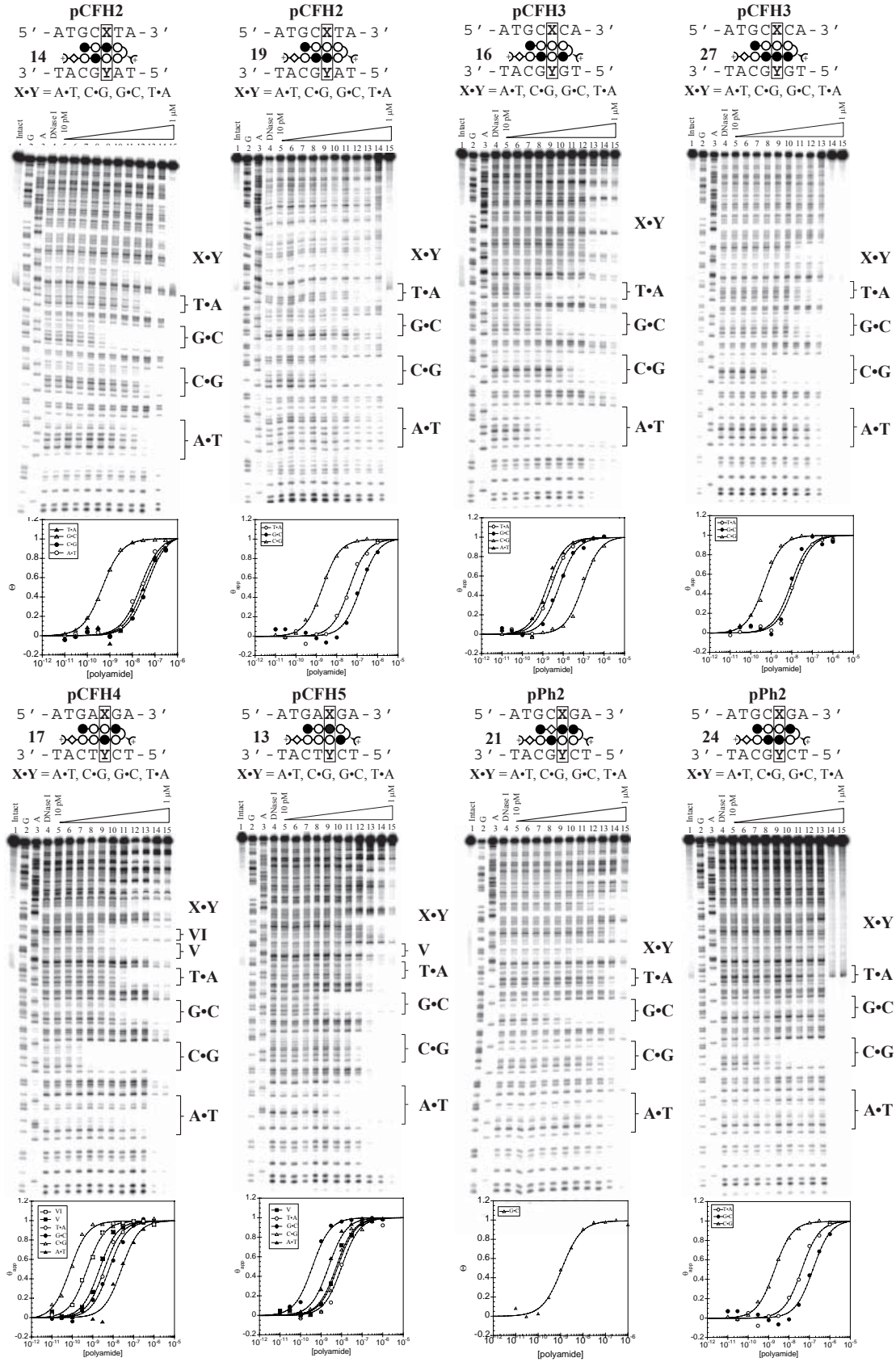


Figure 2.3



**Figure 2.3.** Quantitative DNase I footprint titration experiments for polyamides **13**, **14**, **16**, **17**, **19**, **21**, **24**, and **27** on the 295 bp, 5'-<sup>32</sup>P-end-labeled PCR product of plasmids pCFH2, pCFH3, pCFH4, pCFH5, and pPh2: lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase standard; lanes 5-15, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, 100 nM, 300 nM, 1  $\mu$ M, respectively. Binding isotherms for the four designed sites are shown below each footprinting gel;  $\theta_{\text{norm}}$  values were calculated according to published methods. A binding model for the hairpin motif is shown above each gel with the polyamide bound to its target DNA sequence. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits.

**Table 2.1.** Equilibrium association constants  $K_a$  ( $M^{-1}$ )<sup>a</sup>

pCFH2 (5'→3')	ATGCATA	ATGCCTA	ATGCGTA	ATGCTTA
14	$7.5 (\pm 5.0) \times 10^7$	$3.0 (\pm 1.0) \times 10^7$	<b><math>2.1 (\pm 0.2) \times 10^9</math></b>	$4.6 (\pm 2.6) \times 10^7$
19	$< 1 \times 10^7$	<b><math>7.0 (\pm 2.2) \times 10^8</math></b>	$1.4 (\pm 0.5) \times 10^7$	$3.6 (\pm 1.3) \times 10^7$
pCFH3 (5'→3')	ATGCACA	ATGCCCA	ATGCGCA	ATGCTCA
16	<b><math>7.0 (\pm 1.3) \times 10^9</math></b>	$2.5 (\pm 1.3) \times 10^8$	$2.5 (\pm 1.0) \times 10^9$	<b><math>9.3 (\pm 2.5) \times 10^9</math></b>
27	$< 1 \times 10^7$	<b><math>1.8 (\pm 0.3) \times 10^9</math></b>	$9.8 (\pm 1.3) \times 10^7$	$7.1 (\pm 0.6) \times 10^7$
pCFH4 (5'→3')	ATGAAGA	ATGACGA	ATGAGGA	ATGATGA
17	$5.1 (\pm 3.3) \times 10^7$	<b><math>1.3 (\pm 0.1) \times 10^{10}</math></b>	$1.8 (\pm 0.7) \times 10^8$	$2.9 (\pm 1.0) \times 10^8$
pCFH5 (5'→3')	ATGAACT	ATGACCT	ATGAGCT	ATGATCT
13	$5.3 (\pm 1.4) \times 10^8$	$9.5 (\pm 4.4) \times 10^7$	<b><math>2.2 (\pm 0.4) \times 10^9</math></b>	$8.2 (\pm 4.1) \times 10^7$
pPh2 (5'→3')	ATGCAGA	ATGCCGA	ATGCGGA	ATGCTGA
21	$< 3.3 \times 10^7$	$< 3.3 \times 10^7$	<b><math>9.3 (\pm 0.5) \times 10^8</math></b>	$< 3.3 \times 10^7$
24	$< 1 \times 10^7$	<b><math>2.4 (\pm 0.5) \times 10^9</math></b>	$1.1 (\pm 0.3) \times 10^7$	$2.2 (\pm 1.4) \times 10^7$
pMFST (5'→3')	TAGTGGT	ATATGGT	TAGTCGT	
12	<b><math>4.3 (\pm 0.3) \times 10^{10}</math></b>	$9.9 (\pm 0.6) \times 10^8$	$6.2 (\pm 0.2) \times 10^9$	

<sup>a</sup>  $K_a$  ( $M^{-1}$ ) values reported are the mean values from at least three DNase I footprint titration experiments, with the standard deviation given in parentheses. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; β-alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits.

**2.2.3. Table of DNA Binding Affinities for a Hairpin Polyamide Library.** Having determined the binding affinities for these nine hairpin polyamides, a table of equilibrium association constants can be populated for the general sequence 5'-WWGNNNW-3' (where N = W, G, or C) (Table 2.2). Eight-ring hairpin polyamides can be used to sequence specifically target all of the 27 general sequences with an equilibrium association constant  $K_a \geq 1 \times 10^8 M^{-1}$ . Each of these 27 compounds shown represents a current best solution for its DNA sequence.

**Table 2.2.** Equilibrium association constants  $K_a$  ( $M^{-1}$ )<sup>a</sup>

	General (5'→3')	Polyamide	$K_a$ ( $M^{-1}$ )	Sequence context	Ref
1	WWGWWWW		$3 \times 10^9$	5' - TAGTATT - 3'	23
2	WWGGWWWW		$5 \times 10^8$	5' - CTGGTTA - 3'	42
3	WWGWWGW		$4 \times 10^9$	5' - TAGTGAA - 3'	43
4	WWGWWGW		$9 \times 10^9$	5' - TAGTAGT - 3'	44
5	WWGWWCW		$3 \times 10^{10}$	5' - TAGTACT - 3'	23
6	WWGWCWW		$2 \times 10^9$	5' - GAGTCTA - 3'	c
7	WWGCWWW		$5 \times 10^9$	5' - ATGCAAA - 3'	45
8	WWGGGWW		$3 \times 10^8$	5' - AAGGGAA - 3'	46
9	WWGGWGW		$1 \times 10^{10}$	5' - TAGGTGT - 3'	51
10	WWGGWCW		$1 \times 10^{10}$	5' - ATGGTCA - 3'	24
11	WWGGCWW		$4 \times 10^8$	5' - AAGGCAT - 3'	47
12	WWGWGGW		$4 \times 10^{10}$	5' - TAGTGGT - 3'	b
13	WWGWGCW		$2 \times 10^9$	5' - ATGAGCT - 3'	b
14	WWGCGWW		$2 \times 10^9$	5' - ATGCGTA - 3'	b
15	WWGCWGW		$2 \times 10^9$	5' - TAGCAGT - 3'	48
16	WWGCWCW		$9 \times 10^9$	5' - ATGCTCA - 3'	b
17	WWGWCGW		$1 \times 10^{10}$	5' - ATGACGT - 3'	b
18	WWGWCCW		$2 \times 10^9$	5' - TAGACCA - 3'	49
19	WWGCCWW		$7 \times 10^8$	5' - ATGCCTA - 3'	b
20	WWGGGGW		$2 \times 10^8$	5' - GAGGGGT - 3'	c
21	WWGCGGW		$9 \times 10^8$	5' - ATGCGGT - 3'	b
22	WWGGCGW		$2 \times 10^8$	5' - CAGGCGT - 3'	c
23	WWGGGCW		$1 \times 10^8$	5' - CTGGGCA - 3'	c
24	WWGCCGW		$2 \times 10^9$	5' - ATGCCGT - 3'	b
25	WWGGCCW		$9 \times 10^9$	5' - ATGGCCA - 3'	50
26	WWGCGCW		$3 \times 10^9$	5' - ATGCGCA - 3'	37
27	WWGCCCW		$1 \times 10^9$	5' - ATGCCCA - 3'	b


















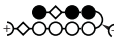

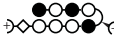


























<sup>a</sup>  $K_a$  ( $M^{-1}$ ) values reported are the mean values from at least three DNase I footprint titration experiments. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; the achiral  $\gamma$ -aminobutyric acid turn residue is shown as a semicircle connecting the two subunits; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits; <sup>b</sup> Table 2.1; <sup>c</sup> Figure 2.4

### 2.3. Discussion

The microstructure of DNA is sequence dependent, and each hairpin polyamide acts as a microcaliper for the shape (width, depth) of the minor groove of DNA. Hence, it is not surprising that the energetics of hairpin binding for match sequences varies over two orders of magnitude ( $10^8$ - $10^{10}$  M<sup>-1</sup>) and is DNA sequence dependent. We would assume for transcriptional factor inhibition experiments one would prefer the highest affinity binding molecules. In the hairpin binding motif,  $\beta$ -alanine has been found to be a good structural replacement for Py, as the  $\beta$ /Im pair is specific for C•G and  $\beta$ /Py and  $\beta$ / $\beta$  both code for A,T.<sup>36,37</sup> Furthermore, when  $\beta$  replaces Py adjacent to Im, binding is generally improved, as this flexible residue is believed to allow the amino acid pairings to reset their register with the DNA base pairs.<sup>37,38</sup> A key example of the use of  $\beta$  to replace Py can be seen in polyamide **26**. ImPyImPy- $\gamma$ -ImPyImPy- $\beta$ -Dp binds its target sequence 5'-ATGCGCA-3' with a binding affinity of  $3 \times 10^7$  M<sup>-1</sup>. Replacement of the internal Py residues with  $\beta$  residues yields the polyamide Im- $\beta$ -ImPy- $\gamma$ -Im- $\beta$ -ImPy- $\beta$ -Dp, which has a binding affinity of  $3 \times 10^9$  M<sup>-1</sup>, a 100-fold improvement.<sup>37</sup>

The chiral substitution of the  $\gamma$ -aminobutyric acid turn residue has been shown to improve binding affinity and sequence specificity.<sup>32</sup> For polyamides **14**, **16**, **19**, and **24**, the chiral diaminobutyric acid turn residue increases equilibrium association constants for these polyamides by 7- to 50-fold over their achiral variants (Table 2.3). In the example of polyamide **21**, the achiral pyrrole-imidazole compound (ImPyImIm- $\gamma$ -PyPyImPy- $\beta$ -Dp) binds its target sequence with an affinity of less than  $1 \times 10^7$  M<sup>-1</sup>, while the chiral version with  $\beta$ -alanine (Im- $\beta$ -ImIm-(*R*)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyPyImPy- $\beta$ -Dp) exhibits greater than 90-fold enhancement, binding with an affinity of  $9 \times 10^8$  M<sup>-1</sup>. The combination of  $\beta$ -alanine substitution when appropriate and incorporation of the chiral turn can yield significantly improved DNA binding.

**Table 2.3.** Equilibrium association constants  $K_a$  ( $M^{-1}$ )<sup>a</sup>

General (5'→3')	Achiral	K <sub>a</sub> (M <sup>-1</sup> )	Sequence context	Ref	Chiral/β	K <sub>a</sub> (M <sup>-1</sup> )	Sequence context	Ref	
1	WWGWWW		3×10 <sup>9</sup>	5' - TAGTATT - 3'	23	-	-	-	
2	WWGGWWW		5×10 <sup>8</sup>	5' - CTGGTTA - 3'	42	-	-	-	
3	WWGWGW		7×10 <sup>8</sup>	5' - TAGTGAA - 3'	43		4×10 <sup>9</sup>	5' - TAGTGAA - 3'	43
4	WWGWGW		-	-	-		9×10 <sup>9</sup>	5' - TAGTAGT - 3'	44
5	WWGWWCW		3×10 <sup>10</sup>	5' - TAGTACT - 3'	23	-	-	-	
6	WWGWCWW		-	-	-		2×10 <sup>9</sup>	5' - GAGTCTA - 3'	c
7	WWGCWWW		-	-	-		5×10 <sup>9</sup>	5' - ATGCCAA - 3'	45
8	WWGGGWW		3×10 <sup>8</sup>	5' - AAGGGAA - 3'	46	-	-	-	
9	WWGGWGW		-	-	-		1×10 <sup>10</sup>	5' - TAGGTGT - 3'	51
10	WWGGWCW		1×10 <sup>10</sup>	5' - ATGGTCA - 3'	24	-	-	-	
11	WWGGCWW		4×10 <sup>8</sup>	5' - AAGGCAT - 3'	47	-	-	-	
12	WWGWGGW		-	-	-		4×10 <sup>10</sup>	5' - TAGTGGT - 3'	b
13	WWGWGCW		-	-	-		2×10 <sup>9</sup>	5' - ATGAGCT - 3'	b
14	WWGCGWW		9×10 <sup>7</sup>	5' - ATGCGTA - 3'	d		2×10 <sup>9</sup>	5' - ATGCGTA - 3'	b
15	WWGCWGW		-	-	-		2×10 <sup>9</sup>	5' - TAGCAGT - 3'	48
16	WWGCWCW		6×10 <sup>8</sup>	5' - ATGCACA - 3'	d		9×10 <sup>9</sup>	5' - ATGCTCA - 3'	b
17	WWGWCGW		-	-	-		1×10 <sup>10</sup>	5' - ATGACGT - 3'	b
18	WWGWCCW		2×10 <sup>9</sup>	5' - TAGACCA - 3'	49	-	-	-	
19	WWGCCWW		1×10 <sup>8</sup>	5' - ATGCCCTA - 3'	d		7×10 <sup>8</sup>	5' - ATGCCCTA - 3'	b
20	WWGGGGW		2×10 <sup>7</sup>	5' - ATGGGGA - 3'	50		2×10 <sup>8</sup>	5' - GAGGGGT - 3'	c
21	WWGCGGW		< 1×10 <sup>7</sup>	5' - ATGCGGT - 3'	d		9×10 <sup>8</sup>	5' - ATGCGGT - 3'	b
22	WWGGCGW		-	-	-		2×10 <sup>8</sup>	5' - CAGGCGT - 3'	c
23	WWGGGCW		-	-	-		1×10 <sup>8</sup>	5' - CTGGGCA - 3'	c
24	WWGCCGW		4×10 <sup>7</sup>	5' - ATGCCGT - 3'	d		2×10 <sup>9</sup>	5' - ATGCCGT - 3'	b
25	WWGGCCW		9×10 <sup>9</sup>	5' - ATGGCCA - 3'	50	-	-	-	
26	WWGCGCW		3×10 <sup>7</sup>	5' - ATGCGCA - 3'	50		3×10 <sup>9</sup>	5' - ATGCGCA - 3'	37
27	WWGCCCW		< 2×10 <sup>8</sup>	5' - TAGCCCA - 3'	45		1×10 <sup>9</sup>	5' - ATGCCCA - 3'	b

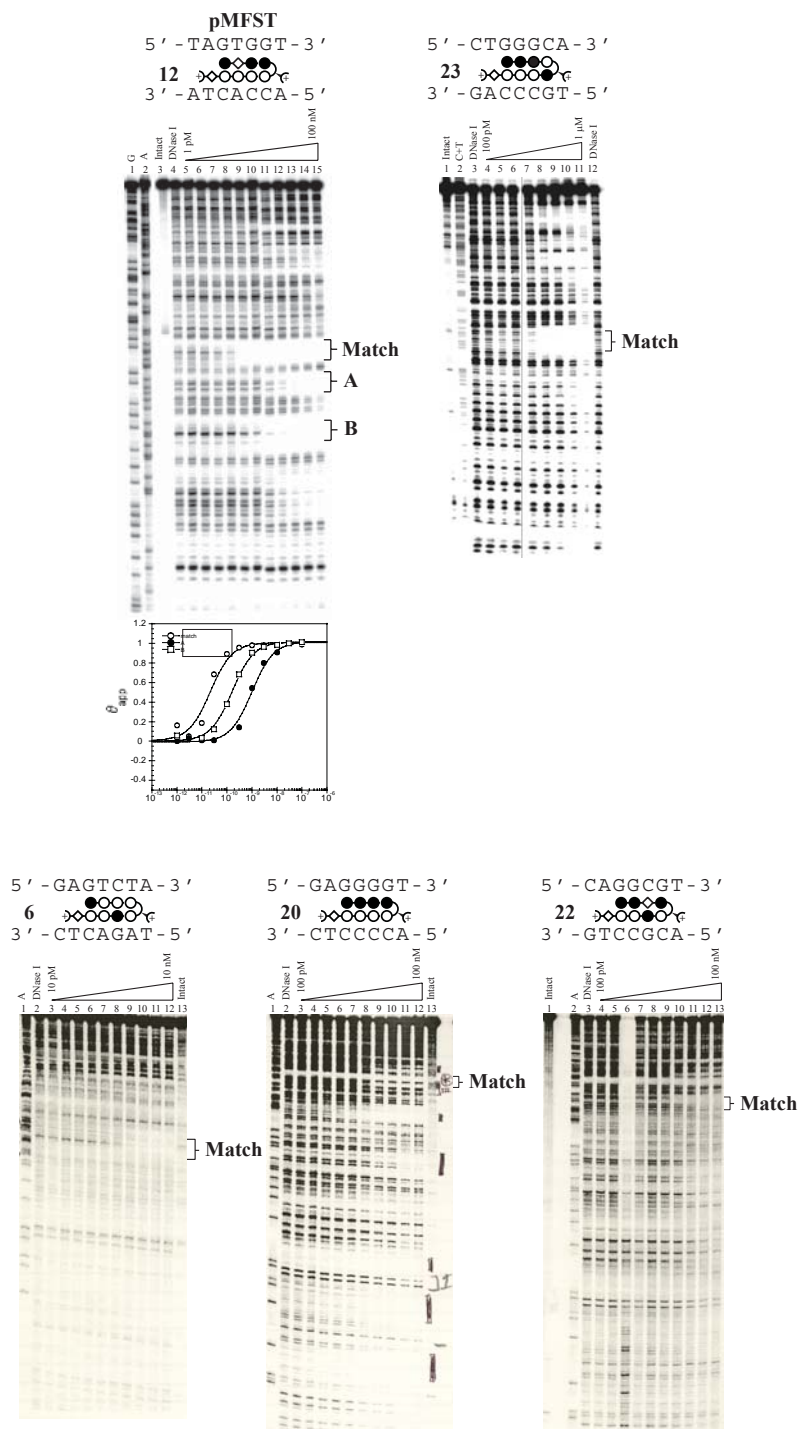
<sup>a</sup>  $K_a$  ( $M^{-1}$ ) values reported are the mean values from at least three DNase I footprint titration experiments. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; the achiral  $\gamma$ -aminobutyric acid turn residue is shown as a semicircle connecting the two subunits; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits; <sup>b</sup> This paper, Table 2.1; <sup>c</sup> This paper, Figure 2.4; <sup>d</sup> Previously unpublished data

## 2.4. Conclusion

We report here nine new eight-ring hairpin polyamides and assay their DNA-binding properties. These characterizations mark the completion of a 27-member polyamide library that covers all 7-bp DNA sequences 5'-WWGNNNW-3' (N = W, G, or C), where each compound binds its respective sequence with  $K_a \geq 1 \times 10^8 \text{ M}^{-1}$ . Typical fold change for match versus single base-pair mismatch sites is in the range of 4 to 100. Looking forward, one could imagine using a DNA microarray-based approach to interrogate the entire sequence-recognition profile of each member of the 27 hairpin polyamide library on every permutation of DNA sites seven base pairs in size.<sup>52</sup> We hope that this compilation of polyamide-DNA binding affinities will serve as a resource for ongoing small molecule gene regulation projects.

## 2.5. Experimental

**2.5.1 Materials.** (*tert*-Butoxycarbonyl)- $\beta$ -alanine-PAMresin(Boc- $\beta$ -Ala-PAMresin, divinylbenzene 1%, 200-400 mesh, 0.81 meq/g loading), *N* $^{\alpha}$ -9-fluorenylmethoxycarbonyl-*N'*-*tert*-butoxycarbonyl-D-2,4-diaminobutyric acid (Fmoc-D-Dab(Boc)-OH), and *O*-(benzotriazol-1-yl)-*N,N,N',N'*-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Peptides International. 1-Methyl-2-pyrrolidinone (NMP), *N,N*-diisopropylethylamine (DIEA), and 3-dimethylamino-1-propylamine (Dp) were purchased from Aldrich. *N,N*-Dimethylformamide (DMF) was purchased from EMD Biosciences, dichloromethane (DCM) was purchased from Fisher Scientific, and trifluoroacetic acid (TFA) was purchased from Halocarbon. Oligonucleotide inserts were synthesized by Integrated DNA Technologies. Bam HI, Hind III, polynucleotide kinase, DNase I, rapid DNA ligation kit, PCR core kit, and glycogen were purchased from Roche. JM109 competent cells and Wizard Plus Midipreps DNA Purification System were purchased from Promega. pUC19 plasmid DNA,  $\beta$ -mercaptoethanol, bromophenol blue, and xylene cyanole were purchased from Sigma. Calf thymus DNA and ProbeQuant G-50 Micro



**Figure 2.4.** Quantitative DNase I footprint titration experiments for polyamides **6**, **12**, **20**, **22**, and **23**. A binding model for the hairpin motif is shown above each gel with the polyamide bound to its target DNA sequence. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits.

Columns were purchased from Amersham Biosciences. Adenosine 5'-triphosphate [ $\gamma$ - $^{32}\text{P}$ ], DL-dithiothreitol, and Tris were purchased from MP Biomedicals. 0.5 M EDTA, pH 8.0, phenol:chloroform:isoamyl alcohol (25:24:1 v/v), and formamide were purchased from Invitrogen. Potassium chloride, magnesium chloride, calcium chloride, and sodium chloride were purchased from Mallinckrodt. RNase-free water was purchased from USB. 8% Gene-PAGE PLUS, 7 M urea, denaturing acrylamide blend was purchased from Amresco. Tris borate EDTA was purchased from National Diagnostics. All reagents were used without further purification.

HPLC analysis was performed on a Beckman Gold system using a Phenomenex Gemini  $4.6 \times 250$  mm,  $5 \mu\text{m}$   $100 \text{ \AA}$   $\text{C}_{18}$  reverse-phase column in 0.1% (w/v) TFA with acetonitrile as the eluent. Preparatory HPLC was carried out on a Beckman Gold system using either a Waters Delta-Pak  $25 \times 100$  mm,  $15 \mu\text{m}$   $300 \text{ \AA}$   $\text{C}_{18}$  PrepPak Cartridge reverse-phase column or a Varian Dynamax  $21.4 \times 250$  mm Microsorb  $8 \mu\text{m}$   $300 \text{ \AA}$   $\text{C}_8$  reverse-phase column in 0.1% (w/v) TFA with acetonitrile as the eluent. UV spectra were measured on an Agilent Technologies 8453 UV-Vis ChemStation spectrophotometer. Matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF-MS) was carried out on an Applied Biosystems Voyager DE-PRO. Storage phosphor autoradiography was performed on a Molecular Dynamics Typhoon 8600 phosphorimager.  $18\text{M}\Omega$  water was obtained from an AquaMAX Ultra water purification system, and all buffers were  $0.2 \mu\text{m}$  filtered.

**2.5.2. Polyamide synthesis.** Polyamides were synthesized using pre-loaded Boc- $\beta$ -Ala-PAM resin (50 mg, 0.81 meq/g) according to published manual solid-phase synthesis protocols.<sup>34</sup> The resin was cleaved with neat 3-dimethylamino-1-propylamine (1 mL) at  $37^\circ\text{C}$  with agitation for 16 h. Products were purified by preparatory reverse-phase HPLC and characterized by analytical HPLC, UV-visible spectroscopy, and MALDI-TOF mass spectrometry.



**2.5.2.1. Im- $\beta$ -ImIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyPyPyPy- $\beta$ -Dp (12).** Polyamide **12** was isolated upon lyophilization of the appropriate fractions as a white powder (0.1 mg, 0.1% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 309 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1187.86 (1187.57 calcd for [M + H]<sup>+</sup> C<sub>54</sub>H<sub>71</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.2. ImPyImPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy- $\beta$ -Dp (13).** Polyamide **13** was isolated upon lyophilization of the appropriate fractions as a white powder (0.1 mg, 0.1% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 314 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1238.59 (1238.58 calcd for [M + H]<sup>+</sup> C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.3. ImPyImPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyPyImPy- $\beta$ -Dp (14).** Polyamide **14** was isolated upon lyophilization of the appropriate fractions as a white powder (0.1 mg, 0.1% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 315 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1238.58 (1238.58 calcd for [M + H]<sup>+</sup> C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.4. ImPyPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyImPy- $\beta$ -Dp (16).** Polyamide **16** was isolated upon lyophilization of the appropriate fractions as a white powder (0.4 mg, 1.5% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 315 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1238.49 (1238.58 calcd for [M + H]<sup>+</sup> C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.5. ImPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -Dp (17).** Polyamide **17** was isolated upon lyophilization of the appropriate fractions as a white powder (1.8 mg, 7.1% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 316 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1238.63 (1238.58 calcd for [M + H]<sup>+</sup> C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.6. ImPyPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImImPy- $\beta$ -Dp (19).** Polyamide **19** was isolated upon lyophilization of the appropriate fractions as a white powder (0.9 mg, 3.8% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 312 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1238.63 (1238.58 calcd for [M + H]<sup>+</sup> C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.7. Im- $\beta$ -ImIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyPyImPy- $\beta$ -Dp (21).** Polyamide **21** was isolated upon lyophilization of the appropriate fractions as a white powder (0.3 mg, 0.2% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 309 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1188.67 (1188.57 calcd for [M + H]<sup>+</sup> C<sub>53</sub>H<sub>70</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.8. ImPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImImPy- $\beta$ -Dp (24).** Polyamide **24** was isolated upon lyophilization of the appropriate fractions as a white powder (1.9 mg, 7.5% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 310 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1239.67 (1239.58 calcd for [M + H]<sup>+</sup> C<sub>56</sub>H<sub>71</sub>N<sub>24</sub>O<sub>10</sub><sup>+</sup>).

**2.5.2.9. ImPyPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImImImPy- $\beta$ -Dp (27).** Polyamide **27** was isolated upon lyophilization of the appropriate fractions as a white powder (0.1 mg, 0.1% recovery). UV (H<sub>2</sub>O)  $\lambda_{\text{max}}$  = 310 nm; MALDI-TOF-MS (monoisotopic)  $m/z$  1239.60 (1239.58 calcd for [M + H]<sup>+</sup> C<sub>56</sub>H<sub>71</sub>N<sub>24</sub>O<sub>10</sub><sup>+</sup>).

**2.5.3. DNase I footprinting experiments.** Plasmids pCFH2, pCFH3, pCFH4, pCFH5, pPh2, and pMFST were constructed according to standard protocols for DNA manipulation.<sup>39</sup> PCR products (295 bp, 5'-<sup>32</sup>P-end-labeled) were isolated and DNase I footprint titrations were performed according to standard protocols.<sup>35</sup> Chemical sequencing reactions were performed according to published methods.<sup>40,41</sup>

**Acknowledgements.** We are grateful to the National Institutes of Health for research support.

## References

- (1) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B., *Nature* **1997**, *387*, 202-205.
- (2) Dickinson, L. A.; Gulizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B., *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12890-12895.
- (3) Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3930-3935.
- (4) Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U. K., *Mol. Cell* **2000**, *6*, 1013-1024.
- (5) Maeshima, K.; Janssen, S.; Laemmli, U. K., *EMBO J.* **2001**, *20*, 3218-3228.
- (6) Gygi, M. P.; Ferguson, M. D.; Mefford, H. C.; Lund, K. P.; O'Day, C.; Zhou, P.; Friedman, C.; van den Engh, G.; Stolowitz, M. L.; Trask, B. J., *Nucleic Acids Res.* **2002**, *30*, 2790-2799.
- (7) Crowley, K. S.; Phillion, D. P.; Woodard, S. S.; Schweitzer, B. A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J. K., *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1565-1570.
- (8) Kodadek, T.; Reddy, M. M.; Olivos, H. J.; Bachhawat-Sikder, K.; Alluri, P. G., *Acc. Chem. Res.* **2004**, *37*, 711-718.
- (9) Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B., *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 16768-16773.
- (10) Bando, T.; Narita, A.; Sasaki, S.; Sugiyama, H., *J. Am. Chem. Soc.* **2005**, *127*, 13890-13895.

- (11) Burnett, R.; Melander, C.; Puckett, J. W.; Son, L. S.; Wells, R. D.; Dervan, P. B.; Gottesfeld, J. M., *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 11497-11502.
- (12) Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B., *Bioorg. Med. Chem.* **2002**, *10*, 3313-3318.
- (13) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12063-12068.
- (14) Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., *Nucleic Acids Res.* **2004**, *32*, 2802-2818.
- (15) Pandolfi, P. P., *Oncogene* **2001**, *20*, 3116-3127.
- (16) Darnell, J. E., *Nat. Rev. Cancer* **2002**, *2*, 740-749.
- (17) Arkin, M. R.; Wells, J. A., *Nat. Rev. Drug Discov.* **2004**, *3*, 301-317.
- (18) Dervan, P. B., *Bioorg. Med. Chem.* **2001**, *9*, 2215-2235.
- (19) Dervan, P. B.; Edelson, B. S., *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299.
- (20) Pelton, J. G.; Wemmer, D. E., *Proc. Natl. Acad. Sci. USA* **1989**, *86*, 5723-5727.
- (21) Wade, W. S.; Mrksich, M.; Dervan, P. B., *J. Am. Chem. Soc.* **1992**, *114*, 8783-8794.
- (22) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B., *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 7586-7590.
- (23) Trauger, J. W.; Baird, E. E.; Dervan, P. B., *Nature* **1996**, *382*, 559-561.
- (24) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *Nature* **1998**, *391*, 468-471.

- (25) Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *Nat. Struct. Biol.* **1998**, *5*, 104-109.
- (26) Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *Science* **1998**, *282*, 111-115.
- (27) Mrksich, M.; Parks, M. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1994**, *116*, 7983-7988.
- (28) Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1999**, *121*, 1121-1129.
- (29) Greenberg, W. A.; Baird, E. E.; Dervan, P. B., *Chem. Eur. J.* **1998**, *4*, 796-805.
- (30) Heckel, A.; Dervan, P. B., *Chem. Eur. J.* **2003**, *9*, 3353-3366.
- (31) Swalley, S. E.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1999**, *121*, 1113-1120.
- (32) Herman, D. M.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1998**, *120*, 1382-1391.
- (33) Renneberg, D.; Dervan, P. B., *J. Am. Chem. Soc.* **2003**, *125*, 5707-5716.
- (34) Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146.
- (35) Trauger, J. W.; Dervan, P. B., *Methods Enzymol.* **2001**, *340*, 450-466.
- (36) Swalley, S. E.; Baird, E. E.; Dervan, P. B., *Chem. Eur. J.* **1997**, *3*, 1600-1607.
- (37) Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1998**, *120*, 6219-6226.
- (38) Wang, C. C. C.; Ellervik, U.; Dervan, P. B., *Bioorg. Med. Chem.* **2001**, *9*, 653-657.

- (39) Sambrook, J.; Fritsch, E. F.; Maniatis, T., *Molecular Cloning: A Laboratory Manual*. 2nd ed.; Cold Spring Harbor Laboratory: Plainview, NY, 1989.
- (40) Maxam, A. M.; Gilbert, W., *Methods Enzymol.* **1980**, *65*, 499-560.
- (41) Iverson, B. L.; Dervan, P. B., *Nucleic Acids Res.* **1987**, *15*, 7823-7830.
- (42) Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B.; Gottesfeld, J. M., *J. Mol. Biol.* **1997**, *274*, 439-445.
- (43) Dickinson, L. A.; Trauger, J. W.; Baird, E. E.; Ghazal, P.; Dervan, P. B.; Gottesfeld, J. M., *Biochemistry* **1999**, *38*, 10801-10807.
- (44) Wang, C. C. C. Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2002.
- (45) Trauger, J. W. Ph.D. thesis, California Institute of Technology, Pasadena, CA, 1998.
- (46) Swalley, S. E.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1996**, *118*, 8198-8206.
- (47) Gottesfeld, J. M.; Melander, C.; Suto, R. K.; Raviol, H.; Luger, K.; Dervan, P. B., *J. Mol. Biol.* **2001**, *309*, 615-629.
- (48) Belitsky, J. M. Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2002.
- (49) Coull, J. J.; He, G. C.; Melander, C.; Rucker, V. C.; Dervan, P. B.; Margolis, D. M., *J. Virol.* **2002**, *76*, 12349-12354.
- (50) Swalley, S. E.; Baird, E. E.; Dervan, P. B., *J. Am. Chem. Soc.* **1997**, *119*, 6953-6961.

- (51) Tsai, S. M.; Farkas, M. E.; Chou, C. J.; Gottesfeld, J. M.; Dervan, P. B., *Nucleic Acids Res.* **2007**, *35*, 307-316.
- (52) Warren, C. L.; Kartochvil, N. C. S.; Hauschild, K. E.; Foister, S.; Brezinski, M. L.; Dervan, P. B.; Phillips, G. N., Jr.; Ansari, A. Z., *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 867-872.

## Chapter 3

### Quantitating the Concentration of Py-Im Polyamide-Fluorescein Conjugates in Live Cells

The text of this chapter was taken in part from a manuscript co-authored with Peter B. Dervan (Caltech).

(Hsu, C. F., Dervan, P. B. *Bioorg. Med. Chem. Lett.* (2008) article in press, doi:10.1016/j.bmcl.2008.05.063)



**Abstract**

Quantitative fluorescence-based methods have been developed to determine the nuclear concentration of polyamide-fluorescein conjugates in cell culture. Confocal laser scanning microscopy and flow cytometry techniques are utilized to plot calibration curves, from which the nuclear concentration can be interpolated. Upon treatment with polyamide, the concentration in the nucleus of live HeLa cells is calculated to be between 0.1-0.5  $\mu\text{M}$ , which is significantly lower than the 2  $\mu\text{M}$  dosage concentration. In contrast, the observed nuclear concentration in U251 cells is closer to the dosage concentration, indicating a cell line-specific increase in uptake for this class of compounds. Although confocal microscopy and flow cytometry generate disparate values, taken together these experiments suggest that the polyamide concentration inside the cell nucleus is lower than it is outside the cell.

### 3.1. Introduction

Hairpin pyrrole-imidazole polyamides are synthetic ligands that target predetermined DNA sequences with affinities comparable to those of naturally occurring DNA-binding proteins.<sup>1,2</sup> These cell-permeable small molecules have been shown to localize to the nucleus of living cells<sup>3-7</sup> and regulate endogenous gene expression.<sup>8-12</sup> Polyamides selectively bind in the minor groove of DNA according to a set of “pairing rules,” where each heterocyclic ring pair targets a specific Watson-Crick base pair.<sup>1,2</sup> The antiparallel pairing of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) aromatic rings targets the C•G base pair, while Im/Py discriminates G•C.<sup>13-16</sup> The Py/Py pair recognizes A•T and T•A, while chlorothiophene can be paired with Py to distinguish T•A at the N-terminus.<sup>17</sup>

Polyamide-small molecule conjugates have been utilized in a number of applications.<sup>18-21</sup> In particular, with laser microscopy it is possible to visualize the uptake of polyamide-fluorophore conjugates as they traffic unaided to the nucleus of live cells.<sup>3-6</sup> This advance has helped to usher in an era of gene regulation.<sup>8-11</sup> Traditionally, we have used a qualitative scoring system to rate the extent of nuclear localization in cell culture.<sup>5,6</sup> Upon incubation with polyamide and direct imaging of cells, the fluorescence intensity in the nucleus is compared to that in the medium. Positive scores are assigned when the nuclear staining exceeds that of the medium. The degree of nuclear localization varies across polyamide design, fluorophore selection, and cell type.<sup>3-6</sup> Quantitating the polyamide concentration within the cell nucleus would be an improvement over the current “yes/no” rating, thus creating a digital readout of cellular uptake.

Our quantitative approach utilizes two different fluorescent imaging technologies. First, with confocal laser scanning microscopy, we can compare images of cells to images of calibration standards containing known polyamide concentrations, then interpolate to determine the nuclear concentration in the cell. Next, we can corroborate this value with additional evidence from flow cytometry. Here, the fluorescence of live cells is compared to a calibration curve built using beads functionalized with known amounts of fluorophore,

yielding the interpolated number of fluorophores per cell. Taken together, these values provide a numerical range that is not apparent from previous qualitative experiments. In this study, we investigate the nuclear concentration of two polyamide-fluorescein conjugates in two cell lines using confocal laser scanning microscopy and flow cytometry.

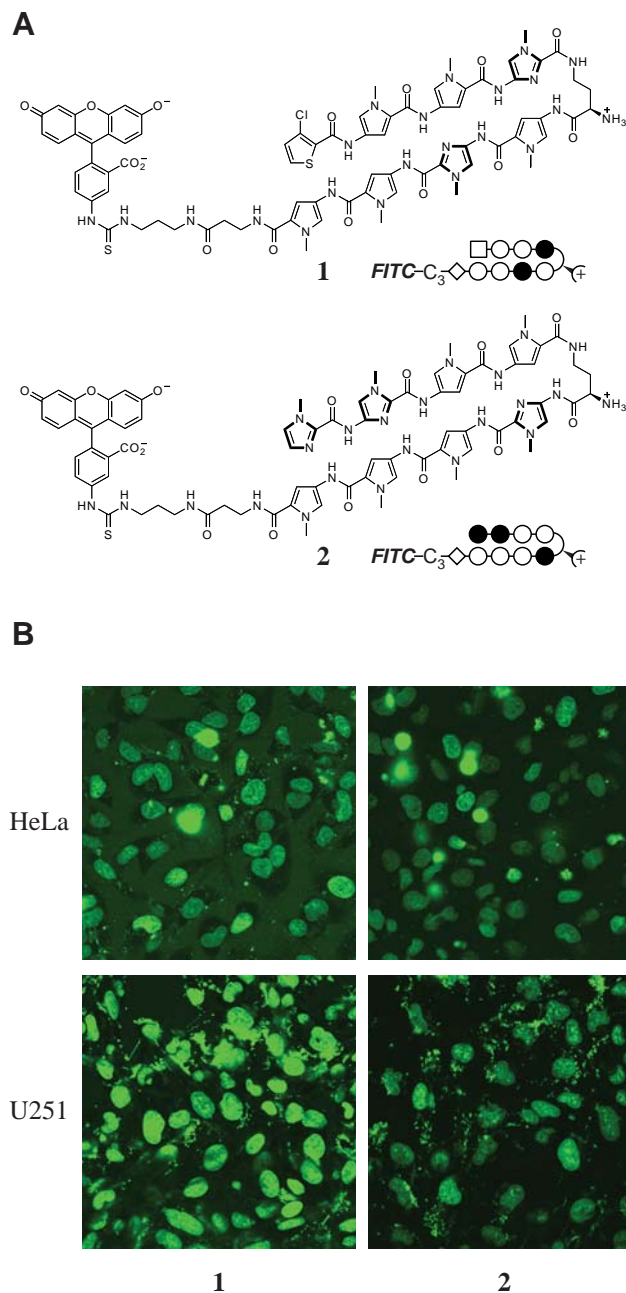
### 3.2. Polyamide synthesis

For these experiments, we wanted to examine polyamide-fluorescein conjugates that traffic strongly to the nucleus of live cells (Figure 3.1). Polyamide **1** specifically targets the DNA sequence 5'-WTWCGW-3' with an equilibrium association constant of  $3.8 \times 10^9 \text{ M}^{-1}$ , while polyamide **2** binds 5'-WGGWCW-3' with an affinity of  $6.3 \times 10^9 \text{ M}^{-1}$  (Figure 3.2). For these molecules synthesized with Boc- $\beta$ -Ala-PAM resin, the 1,3-diaminopropane ( $\text{C}_3$ ) linker is used to connect the polyamide and fluorescein moiety.

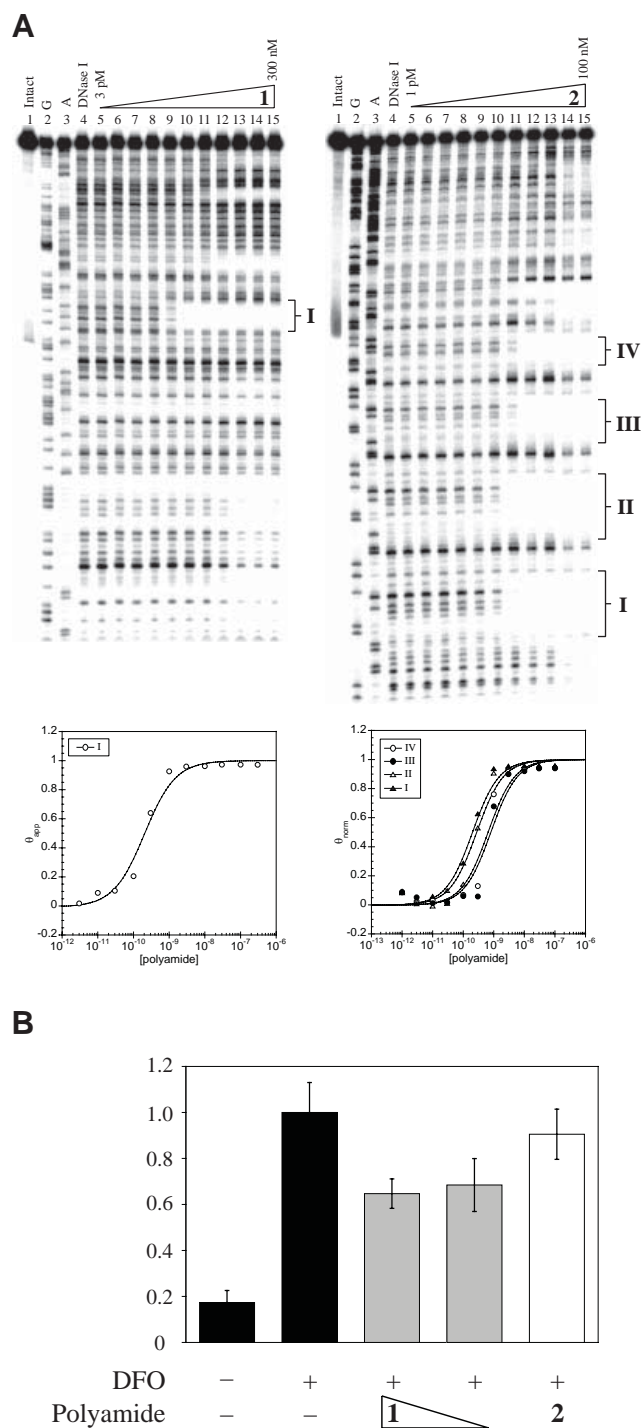
Polyamides **1** and **2** show positive nuclear localization in HeLa and U251 cell culture by confocal laser scanning microscopy (Figure 3.1). In each of these images, the nuclear staining exceeds that of the medium, and all are qualitatively rated as ++.<sup>5,6</sup> Furthermore, conjugate **1** binds to the hypoxia response element (HRE) and inhibits vascular endothelial growth factor (VEGF) expression in HeLa cells by quantitative RT-PCR, while the mismatch control compound **2** does not downregulate gene expression (Figure 3.2).<sup>8-10</sup>

### 3.3. Steady-state fluorimetry

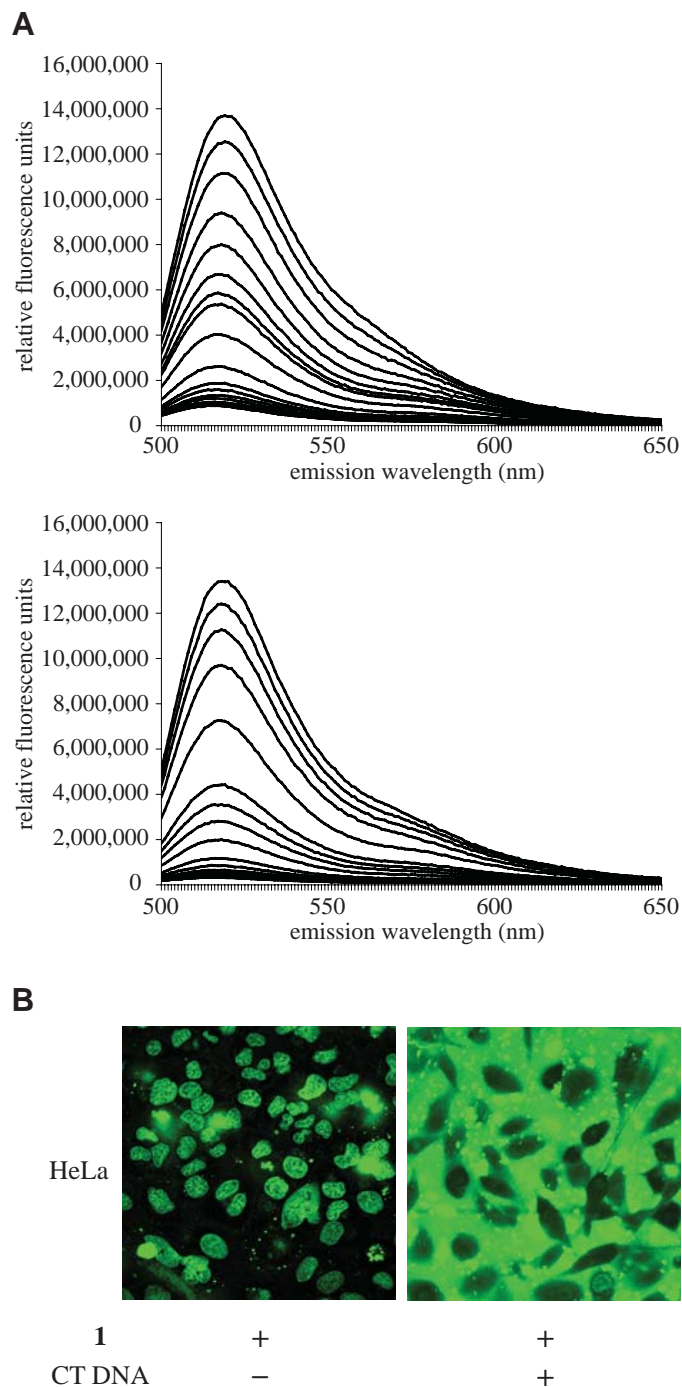
Fluorescence enhancement is observed when polyamide-fluorophore conjugates bind DNA.<sup>22-25</sup> When in solution, the fluorescence of the conjugate is largely quenched, but, upon DNA binding, sequestration of the polyamide in the minor groove restores fluorescence. The fluorescence enhancements of polyamides **1** and **2** were 14- and 36-fold, respectively, as determined by steady-state fluorimetry (Figure 3.3). The fluorescent nuclei observed in confocal microscopy images appear much brighter than the background medium. Does this mean that the nuclear polyamide concentration is very high, or does the fluorescence enhancement dominate?



**Figure 3.1.** (A) Structures of polyamide-fluorescein conjugates **1** and **2**. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square;  $\beta$ -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C<sub>3</sub>”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus. (B) Confocal laser scanning microscopy images of polyamide-fluorescein conjugates **1** and **2** in HeLa and U251 cell lines. Cells were incubated with 2  $\mu$ M polyamide for 12-14 hours at 37°C in a 5% CO<sub>2</sub> atmosphere.



**Figure 3.2.** (A) Quantitative DNase I footprint titration experiments for polyamides **1** and **2**; lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase standard; lanes 5-15, polyamide concentrations. Binding isotherms are shown below each footprinting gel;  $\theta_{\text{norm}}$  values were calculated according to published methods.<sup>28</sup> (B) Quantitative RT-PCR data for **1** and **2** in HeLa cells induced with deferoxamine (DFO).<sup>8-10</sup> The final concentration of compound **1** was 0.2  $\mu\text{M}$  or 1  $\mu\text{M}$ . The final concentration of compound **2** was 1  $\mu\text{M}$ .

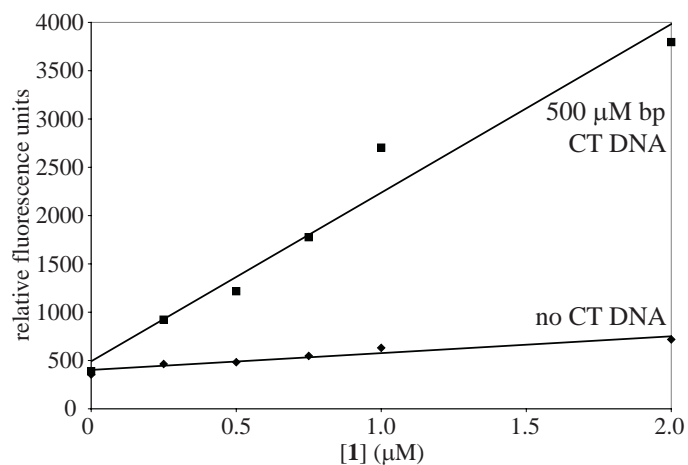


**Figure 3.3.** (A) Overlaid fluorescence emission spectra for polyamides **1** (top) and **2** (bottom). The polyamide concentration is 2  $\mu\text{M}$ . Calf thymus DNA concentrations are 0, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1, 2, 4, 6, 8, 10, 20, 40, 60, 80, 100, 200, 400, 600, 800, and 1000  $\mu\text{M}$  bp. (B) Confocal laser scanning microscopy images of polyamide-fluorescein conjugate **1** in HeLa cells. Cells were incubated with 2  $\mu\text{M}$  polyamide for 12-14 hours at 37°C in a 5%  $\text{CO}_2$  atmosphere. 500  $\mu\text{M}$  bp calf thymus DNA was added 1 hour prior to imaging.

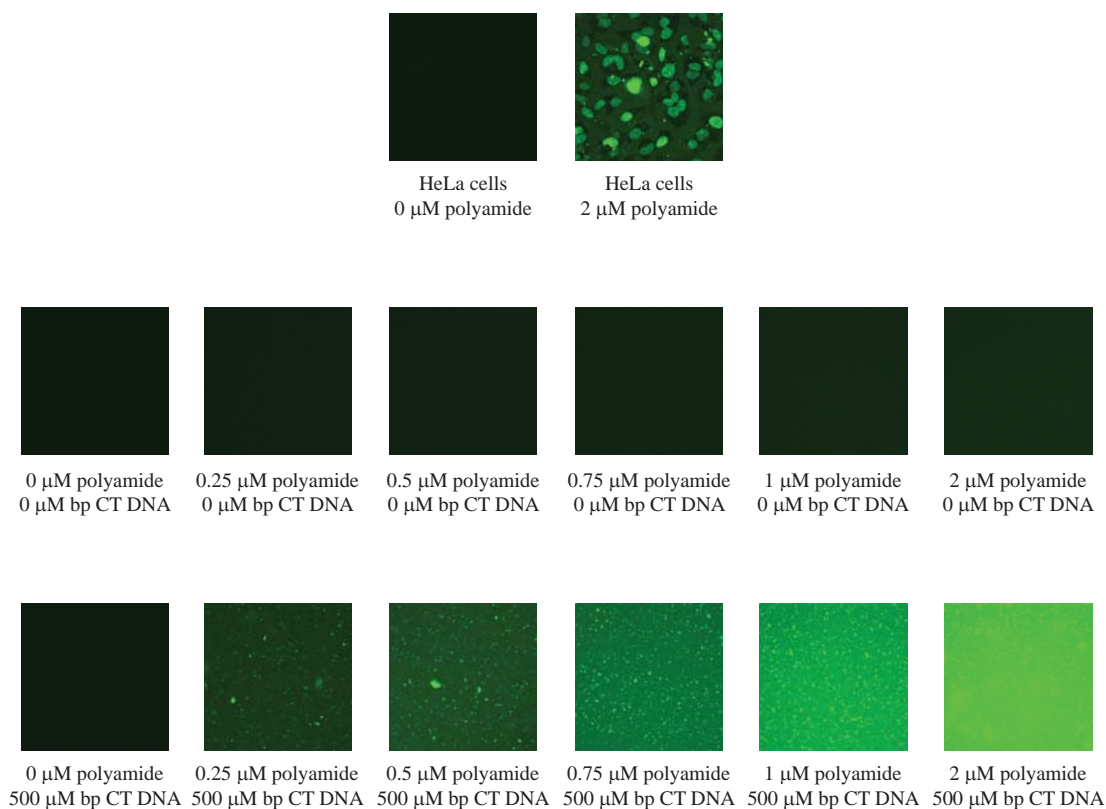
In order to control for this effect, calf thymus DNA was added to the medium prior to imaging. This allows the direct comparison of polyamide concentration inside the cell nucleus and outside the cell by increasing the fluorescence of both simultaneously. This experiment generates a picture of dark cells in a background of fluorescently enhanced polyamide solution (Figure 3.3). Overlay of the fluorescence and bright-field images indicates that the fluorescence in the cell nucleus is significantly lower than that of the medium. The image of the “dark” nucleus implies that the nuclear concentration is lower than the dosage concentration.

### **3.4. Confocal microscopy**

For the first quantitative approach, the fluorescence of cells was compared to the fluorescence of calibration standards containing known polyamide concentrations, using images taken with a confocal laser scanning microscope. After building a linear calibration curve with the known concentration values, the nuclear concentration in the cell was determined by interpolation (Figure 3.4 and Figure 3.5). Cells were incubated with 2  $\mu\text{M}$  polyamide for 12-14 hours in a 5%  $\text{CO}_2$  atmosphere at 37°C and imaged directly. Calibration standards contained both polyamide and 500  $\mu\text{M}$  bp calf thymus DNA. To investigate nuclear concentration, calf thymus DNA was included in the calibration standards to ensure that the fluorescence was turned “on.” For these experiments, the concentration of calf thymus DNA (500  $\mu\text{M}$  bp) is lower than the estimated DNA concentration in the nucleus (5 mM bp). Attempts to increase the concentration to the low mM bp range were limited by DNA solubility in medium. Similar experiments in TKMC buffer resulted in fluorescence quenching at the highest DNA concentrations. The actual DNA bp concentration in the nucleus could produce increased fluorescence enhancement and might bring the calculated polyamide concentration closer to the values observed in flow cytometry experiments. In addition, these experiments were performed on cells with mixed cell cycles, i.e., these cells were not grown in synchronized culture. A minimum of three images were taken per condition, with at least ten cells from each image used for analysis.



**Figure 3.4.** Sample calibration curve for confocal microscopy. Squares indicate equilibration with 500  $\mu\text{M}$  bp calf thymus DNA, while diamonds indicate no calf thymus DNA control. R-squared values are 0.9639 (500  $\mu\text{M}$  bp calf thymus DNA) and 0.9171 (no calf thymus DNA).



**Figure 3.5.** Confocal laser scanning microscopy images of calibration standards.



The experiments in HeLa cells indicated that the concentration in the cell nucleus was 0.5  $\mu\text{M}$  and 0.3  $\mu\text{M}$  for polyamides **1** and **2**, respectively (Table 3.1). These values were significantly lower than the 2  $\mu\text{M}$  dosage concentration, which was surprising given that the fluorescence images show the cell nucleus lighting up brightly against a dark background of cytoplasm and medium. The fluorescence signal from the nucleus exceeds the external signal even though the nuclear concentration is lower than the external concentration, as a result of fluorescence enhancement upon binding DNA.

When U251 cells were treated with 2  $\mu\text{M}$  of polyamides **1** and **2**, the interpolated nuclear concentration values were 1.6  $\mu\text{M}$  and 1.5  $\mu\text{M}$ , respectively. This observation suggests that the nuclear concentration is dependent on cell line, just as it varies for each polyamide. The quantitative approach indicates that U251 cells are more permissive than HeLa cells with respect to uptake of this type of molecule.

**Table 3.1.** Calculated nuclear concentration from confocal microscopy<sup>a</sup>

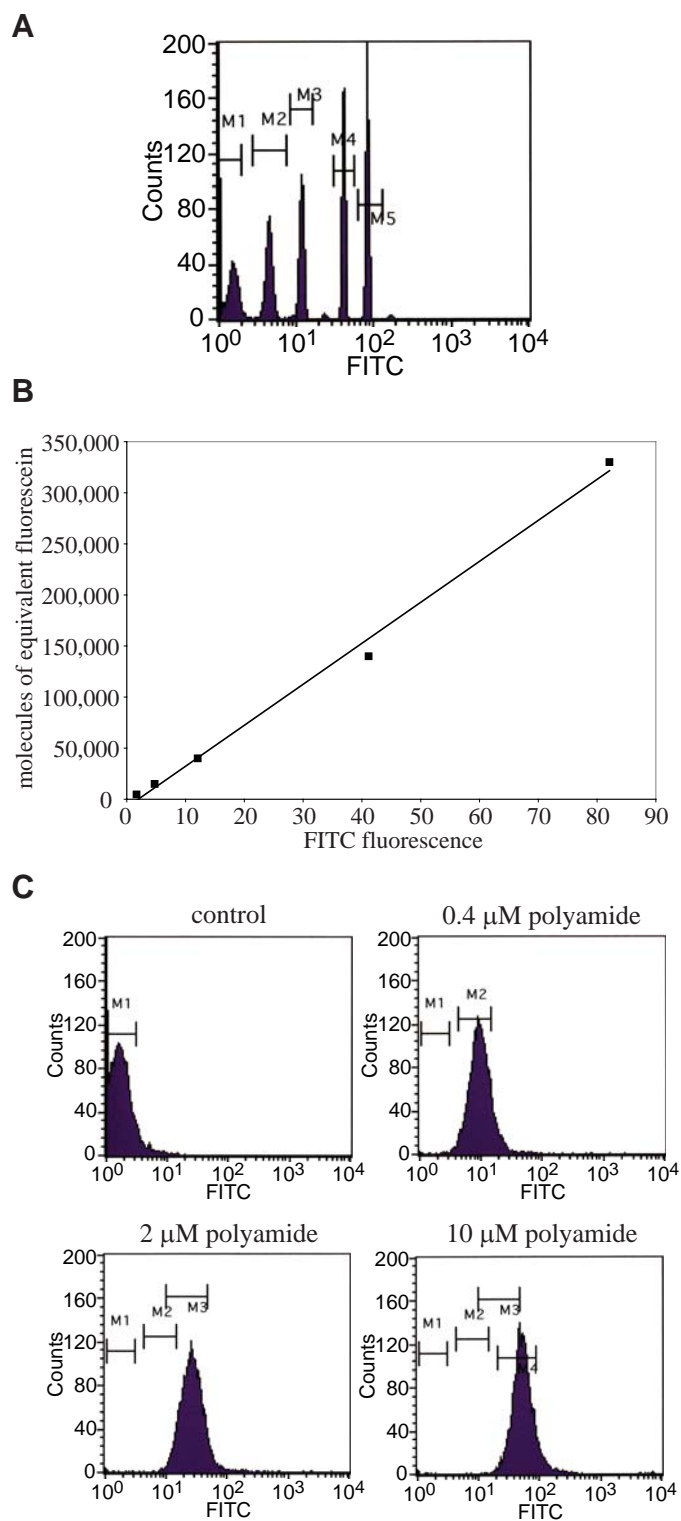
Polyamide	Cell line	2 $\mu\text{M}$
<b>1</b>	HeLa	0.5 ( $\pm 0.1$ )
<b>2</b>	HeLa	0.3 ( $\pm 0.1$ )
<b>1</b>	U251	1.6 ( $\pm 0.1$ )
<b>2</b>	U251	1.5 ( $\pm 0.3$ )

<sup>a</sup> Calculated nuclear concentration values reported in  $\mu\text{M}$ ; column headings indicate dosage concentration

### 3.5. Flow cytometry

The fluorescence of polyamide-containing cells was compared to fluorescent beads in the flow cytometer, and interpolation of the linear calibration curve produces the number of fluorophores per cell (Figure 3.6). Dividing this value by the volume of the nucleus yields the calculated nuclear concentration. Based on confocal microscopy images of HeLa and U251 cells, the cell nucleus was modeled as a cylinder with radius 10  $\mu\text{m}$  and height 5  $\mu\text{m}$  to give a calculated nuclear volume of  $1 \times 10^{-12}$  L. As above, cells were incubated with polyamide for 12-14 hours in a 5%  $\text{CO}_2$  atmosphere at  $37^\circ\text{C}$ . After washing, fluorescence was measured by analyzing 10,000 cells in the flow cytometer. Under identical instrument settings, SPHERO Rainbow Calibration Particles were analyzed in the fluorescein filter set. The functionalized particles contain eight different amounts of fluorophore per bead and are designed to produce eight separate peaks. Due to the high fluorescence levels of the cells, only the five highest fluorophore peaks were resolved within the fluorescence range and used for constructing the linear calibration curve. The fluorescence intensity of each peak is correlated with a known amount of fluorophore in solution, although the quantum yield of a fluorophore can change upon binding.<sup>26</sup> For analysis, 10,000 events were used in flow cytometry experiments and calibrations.

For flow cytometry experiments, cells were dosed with three different polyamide concentrations of 0.4  $\mu\text{M}$ , 2  $\mu\text{M}$ , and 10  $\mu\text{M}$ . In HeLa cells, the calculated nuclear concentrations for compound **1** were 0.02  $\mu\text{M}$ , 0.09  $\mu\text{M}$ , and 0.20  $\mu\text{M}$  (Table 3.2). While the concentration values increase with dosage, each 5-fold increase in dosage does not produce a proportional increase in nuclear concentration, and this plateau effect is more pronounced at higher dosage concentrations. In addition, these values were lower than those observed in confocal microscopy experiments. For the same dosage concentration of 2  $\mu\text{M}$ , the calculated concentration from flow cytometry was 0.09  $\mu\text{M}$ , which is 5-fold lower than the interpolated concentration from confocal microscopy (0.5  $\mu\text{M}$ ). The results from experiments with compound **2** in HeLa cells are calculated nuclear concentrations of



**Figure 3.6.** (A) Flow cytometry experiment with calibration particles. (B) Sample calibration curve for flow cytometry. R-squared value is 0.9944. (C) Flow cytometry experiments with HeLa cells incubated with 0  $\mu$ M (top left), 0.4  $\mu$ M (top right), 2  $\mu$ M (bottom left), and 10  $\mu$ M (bottom right) polyamide.

**Table 3.2.** Calculated cellular concentration from flow cytometry<sup>a</sup>

Polyamide	Cell line	0.4 $\mu$ M	2 $\mu$ M	10 $\mu$ M
<b>1</b>	HeLa	0.02 ( $\pm$ 0.01)	0.09 ( $\pm$ 0.03)	0.20 ( $\pm$ 0.06)
<b>2</b>	HeLa	0.04 ( $\pm$ 0.02)	0.14 ( $\pm$ 0.05)	0.35 ( $\pm$ 0.08)
<b>1</b>	U251	0.09 ( $\pm$ 0.04)	0.34 ( $\pm$ 0.23)	0.84 ( $\pm$ 0.41)
<b>2</b>	U251	0.04 ( $\pm$ 0.02)	0.14 ( $\pm$ 0.06)	0.34 ( $\pm$ 0.17)

<sup>a</sup> Calculated cellular concentration values reported in  $\mu$ M; column headings indicate dosage concentration

0.04  $\mu$ M, 0.14  $\mu$ M, and 0.35  $\mu$ M, indicating that uptake is polyamide core-dependent.

When U251 cells were treated with compound **1**, the calculated nuclear concentration values were 0.09  $\mu$ M, 0.34  $\mu$ M, and 0.84  $\mu$ M. The relatively high nuclear concentrations support the observation that U251 cells are more permissive than HeLa cells, as seen in the above confocal microscopy experiments.

### 3.6. Comparison of methods

For each polyamide and cell line explored here, the confocal microscopy method produces a higher calculated nuclear concentration value than the flow cytometry method. For instance, at the 2  $\mu$ M dosage level in HeLa cells, the polyamides exhibit 0.3-0.5  $\mu$ M nuclear concentration by confocal microscopy, as compared to 0.09-0.14  $\mu$ M by flow cytometry. However, both methods support the observation that U251 cells allow increased uptake of polyamide-fluorescein conjugates relative to HeLa cells. Overall, the calculations suggest that the concentration in the nucleus is much lower than that in the external medium. This is the opposite of what one might expect upon seeing brightly fluorescent cell nuclei with a dark background.

Both of the techniques explored here have strengths and limitations. Confocal microscopy allows data analysis from direct images of cell culture, requiring no further cell processing steps. This method also uses the same compound for calibration and imaging, as opposed to a generic fluorophore. On the other hand, flow cytometry lends itself well to the quantitative nature of this study and benefits from large sample sizes. However, the nuclear volume is not known with certainty, and the flow cytometer measures the fluorescence of the entire cell, not just the nucleus. Also, the calibration does not take into account the fluorescence enhancement effect; therefore, the calculated nuclear concentration from flow cytometry represents an upper bound for this value.

In this study, we elaborate methods for quantitating the concentration of polyamide-fluorescein conjugates in live cells, incorporating data gathered from confocal microscopy and flow cytometry techniques. When cultured HeLa cells are treated with 2  $\mu\text{M}$  polyamide, the calculated nuclear concentration is between 0.09-0.5  $\mu\text{M}$ , suggesting that the intracellular concentration is 4- to 25-fold lower than the external concentration. As observed in previous studies, cellular uptake is cell-line dependent.<sup>3,5,6</sup> Although qualitative images in HeLa and U251 cell lines both receive the same positive rating, it appears that the nuclear concentration in U251 cells is actually several-fold higher than in HeLa cells.

### 3.7. Experimental

**3.7.1 Polyamide synthesis.** Polyamides were synthesized with Boc- $\beta$ -Ala-PAM resin (Peptides International) according to published manual solid-phase synthesis protocols.<sup>27</sup> The protected  $\text{FmocHN}\gamma$ -turn amine was deprotected with 20% piperidine in DMF and reprotected as the Boc derivative with a solution of  $\text{Boc}_2\text{O}$  (Fluka) and DIEA in DMF. The Boc-protected resin was cleaved with 1 mL of 1,3-diaminopropane at 37°C with agitation for 16 h. Products were purified by preparatory reverse-phase HPLC on a Beckman Gold system using either a Waters Delta-Pak 25×100 mm, 15  $\mu\text{m}$  300 Å  $\text{C}_{18}$  PrepPak Cartridge reverse-phase column or a Varian Dynamax 21.4×250 mm Microsorb 8  $\mu\text{m}$  300 Å  $\text{C}_8$  reverse-phase column in 0.1% (w/v) TFA with acetonitrile as the eluent. The appropriate fractions were lyophilized after characterization by analytical HPLC, UV-visible spectroscopy, and MALDI-TOF or ESI mass spectrometry. Conjugates were formed by reacting fluorescein-5-isothiocyanate (FITC, Invitrogen) with the polyamide in a solution of DIEA (20 equiv) and DMF for 1 h at room temperature. Conjugates were deprotected with neat TFA (Halocarbon) and triethylsilane for 30 min at room temperature before purification by preparatory reverse-phase HPLC. Lyophilization of the appropriate fractions yielded the polyamide conjugates **1** and **2**, which were characterized as described above. Extinction coefficients were calculated according to standard protocols.<sup>28</sup> Chemicals not otherwise specified were from Aldrich.

**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -C<sub>3</sub>-FITC (1).** UV-vis ( $\text{H}_2\text{O}$ )  $\lambda_{\text{max}}$  310, 444 nm; ESI-MS  $m/z$  1633.6 ( $\text{C}_{76}\text{H}_{74}\text{ClN}_{22}\text{O}_{15}\text{S}_2^-$  calculated  $[\text{M-H}]^-$  1633.48)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy- $\beta$ -C<sub>3</sub>-FITC (2).** UV-vis ( $\text{H}_2\text{O}$ )  $\lambda_{\text{max}}$  312, 444 nm; ESI-MS  $m/z$  1597.5 ( $\text{C}_{76}\text{H}_{77}\text{N}_{24}\text{O}_{15}\text{S}^-$  calculated  $[\text{M-H}]^-$  1597.57)

**3.7.2. Steady-state fluorimetry.** For steady-state fluorimetry experiments, titration samples were prepared with sonicated calf thymus DNA (GE Healthcare) as serial dilutions with supplemented medium. Addition of 3  $\mu\text{L}$  of the 100  $\mu\text{M}$  polyamide solution

to each dilution yielded a final polyamide concentration of 2  $\mu$ M. Fluorescence spectra were measured with a Jobin Yvon/SPEX Fluorolog spectrofluorimeter (Model FL3-11) equipped with a Hamamatsu R928 PMT. Samples were excited at 490 nm using 2 nm emission and excitation slits. Fluorescence was measured from 500 nm to 650 nm at room temperature.

**3.7.3. Cell culture.** For cell culture experiments, human cancer cell lines HeLa and U251 were cultured in a 5% CO<sub>2</sub> atmosphere at 37°C in supplemented DMEM (HeLa, GIBCO) or RPMI medium 1640 (U251, GIBCO).<sup>5,6</sup> All media were supplemented with 10% fetal bovine serum (HeLa, Omega Scientific) or 5% fetal bovine serum (U251, Omega Scientific) and 1% penicillin/streptomycin solution (Mediatech). HeLa cells were purchased from ATCC. U251 cells were received as a gift from Dr. Giovanni Melillo of the National Cancer Institute.<sup>29</sup> Quantitative RT-PCR experiments were performed using HeLa cells according to published protocols.<sup>8,9</sup>

**3.7.4. Confocal microscopy.** For confocal microscopy experiments, cell lines were trypsinized (Mediatech) for 5 min at 37°C, centrifuged for 10 min at 4°C at 100 $\times$ g, and resuspended in fresh medium to a concentration of 1.33 $\times$ 10<sup>5</sup> cells/mL.<sup>5,6</sup> Incubations were performed by adding 150  $\mu$ L of cells into culture dishes equipped with glass bottoms for direct imaging (MatTek). Cells were grown in the glass-bottom dishes for 24 h. The medium was then removed and replaced with 147  $\mu$ L of fresh medium, followed by addition of 3  $\mu$ L of the 100  $\mu$ M polyamide solution for a final polyamide concentration of 2  $\mu$ M. Cells were incubated in a 5% CO<sub>2</sub> atmosphere at 37°C for 12-14 h. Imaging was performed with a 40 $\times$  oil-immersion objective lens on a Zeiss LSM 5 Pascal inverted laser scanning microscope. Polyamide-fluorescein conjugate fluorescence and visible-light images were obtained using standard filter sets for fluorescein.<sup>5,6</sup> 12-bit images were analyzed using Zeiss LSM and ImageJ software. Calibration standards were prepared from the 100  $\mu$ M

polyamide solution as serial dilutions with fresh medium and were directly imaged in glass-bottom dishes in a volume of 150  $\mu$ L. Sonicated calf thymus DNA solution (GE Healthcare) was added to appropriate calibration standards for a final DNA concentration of 500  $\mu$ M bp.

**3.7.5. Flow cytometry.** For flow cytometry experiments, cell lines were trypsinized (Mediatech) for 5 min at 37°C, centrifuged for 10 min at 4°C at 100×g, and resuspended in fresh medium to a concentration of  $1.33 \times 10^5$  cells/mL. Incubations were performed by adding 3 mL of cells into 6-well culture plates (BD Falcon). Cells were grown in the culture plates for 24 h. The medium was then removed and replaced with fresh medium, followed by addition of the appropriate volume of the 100  $\mu$ M polyamide solution for a final polyamide concentration of 0, 0.4, 2, or 10  $\mu$ M in a final volume of 3 mL. Cells were incubated in a 5% CO<sub>2</sub> atmosphere at 37°C for 12-14 h. Following incubation, cells were trypsinized for 5 min at 37°C, centrifuged for 10 min at 4°C at 100×g, and resuspended in HBSS solution (no Mg<sup>2+</sup>, no Ca<sup>2+</sup>, no phenol red, 2.5 mg/mL BSA fraction V, 10 mM HEPES, pH 7.0-7.2) (Sigma). Cell viability was checked with trypan blue stain, and >95% appeared to be viable. Cells were centrifuged again for 10 min at 4°C at 100×g and resuspended in the above HBSS solution to a concentration of  $5 \times 10^5$  cells/mL. The cell suspension was pipetted through a cell strainer (BD Falcon) into a 5 mL polystyrene round-bottom tube (BD Falcon). Samples were analyzed using a BD FACSCalibur System flow cytometer. SPHERO Rainbow Calibration Particles (Spherotech) were used as calibration standards.

**3.7.6. Quantitative DNase I footprinting.** For quantitative DNase I footprint titration experiments, plasmid pCFH6 was constructed according to standard protocols for DNA manipulation.<sup>30</sup> The two 87-mer oligonucleotides 5'-GATCGTGTAATCAAATGGTCATAGCTGTGTAATCATATGGTCATAGCTGT



GTAATCACATGGTCATAGCTGTGTAATCAGATGGTCATAGC-3' and 5'-  
AGCTGCTATGACCATCTGATTACACAGCTATGACCATGTGATTACACAGC

TATGACCATATGATTACACAGCTATGACCATTTGATTACAC-3' (Integrated DNA Technologies) were annealed and ligated into the BamHI/HindIII (Roche) restriction fragment of pUC19 (Sigma). DNA sequencing of the constructed plasmid was performed at the Sequence Analysis Facility at the California Institute of Technology. The forward primer 5'-AATTCGAGCTCGGTACCCGGG-3' was <sup>32</sup>P-labeled at the 5'-end. The reverse primer 5'-CTGGCACGACAGGTTTCCCGAC-3' was used to amplify plasmid pCFH6 as previously described.<sup>28</sup> pGL-VEGF-*Luc* was 5'-end-labeled and amplified as previously described.<sup>8,9</sup> PCR products (5'-end-labeled, 291 bp for pCFH6, 197 bp for pGL-VEGF-*Luc*) were isolated according to standard protocols.<sup>28</sup> Quantitative DNase I footprint titration experiments were performed on the 5'-<sup>32</sup>P-end-labeled PCR products of plasmids pGL-VEGF-*Luc* and pCFH6 with polyamides **1** and **2**, respectively, according to standard protocols.<sup>28</sup> Radiolabeled DNA was equilibrated with polyamide solutions for 14-16 h at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0 prior to DNase I cleavage. Chemical sequencing reactions were performed according to published methods.<sup>31,32</sup>

**Acknowledgments.** We thank Shelley Diamond and the Caltech Flow Cytometry Facility for instrumentation assistance. We also thank Scott Fraser and the Beckman Institute Biological Imaging Center for helpful discussions. We are grateful to the National Institutes of Health for research support. Mass spectrometry analyses were performed in the Mass Spectrometry Facility of the Division of Chemistry and Chemical Engineering at the California Institute of Technology. Fluorimetry measurements were performed in the Beckman Institute Laser Resource Center.

## References

1. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
2. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284.
3. Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 3313.
4. Crowley, K. S.; Phillion, D. P.; Woodard, S. S.; Schweitzer, B. A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J. K. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1565.
5. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12063.
6. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802.
7. Xiao, X. S.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *Angew. Chem. Int. Ed.* **2007**, *46*, 2865.
8. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16768.
9. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**, *35*, 363.
10. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *ACS Chem. Biol.* **2007**, *2*, 561.
11. Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10418.

12. Fukuda, N.; Ueno, T.; Tahira, Y.; Ayame, H.; Zhang, W.; Bando, T.; Sugiyama, H.; Saito, S.; Matsumoto, K.; Mugishima, H.; Serie, K. *J. Am. Soc. Nephrol.* **2006**, *17*, 422.
13. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559.
14. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468.
15. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104.
16. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, *282*, 111.
17. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B. *Bioorg. Med. Chem.* **2003**, *11*, 4333.
18. Wurtz, N. R.; Dervan, P. B. *Chem. Biol.* **2000**, *7*, 153.
19. Kwonj, Y.; Arndt, H. D.; Qian, M.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M. *J. Am. Chem. Soc.* **2004**, *126*, 15940.
20. Stafford, R. L.; Arndt, H. D.; Brezinski, M. L.; Ansari, A. Z.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2660.
21. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B. *Angew. Chem. Int. Ed.* **2007**, *46*, 7956.
22. Rucker, V. C.; Foister, S.; Melander, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 1195.
23. Rucker, V. C. Ph.D. Thesis, California Institute of Technology, Pasadena, CA, 2003.

24. Rucker, V. C.; Dunn, A. R.; Sharma, S.; Dervan, P. B.; Gray, H. B. *J. Phys. Chem. B* **2004**, *108*, 7490.
25. Chenoweth, D. M.; Viger, A.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2216.
26. SPHERO Technical Notes 8 and 9 (<http://www.spherotech.com/>).
27. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.
28. Trauger, J. W.; Dervan, P. B. *Methods Enzymol.* **2001**, *340*, 450.
29. Kong, D. H.; Park, E. J.; Stephen, A. G.; Calvani, M.; Cardellina, J. H.; Monks, A.; Fisher, R. J.; Shoemaker, R. H.; Melillo, G. *Cancer Res.* **2005**, *65*, 9047.
30. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd ed. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1989.
31. Maxam, A. M.; Gilbert, W. *Methods Enzymol.* **1980**, *65*, 499.
32. Iverson, B. L.; Dervan, P. B. *Nucleic Acids Res.* **1987**, *15*, 7823.

## **Chapter 4A**

### **Effects of Tail Linkages on the Sequence-Specific Recognition of DNA by Hairpin Polyamide Conjugates**

This project was done in collaboration with James Puckett (Dervan group, Caltech).

James Puckett synthesized polyamides **6-10** and prepared plasmid pJWP6.

## Abstract

Hairpin pyrrole-imidazole polyamides are synthetic ligands that bind the minor groove of DNA at subnanomolar concentrations with programmable sequence specificity. While the empirically derived “pairing rules” have been well characterized for unlinked polyamide cores, far less is known about the binding affinity and sequence specificity of hairpin polyamide conjugates, which are now commonly used for improved nuclear localization in biological experiments. To further our understanding of C-terminal tail linkage effects on sequence specificity, the equilibrium association constants of five hairpin polyamide conjugates  $\text{ImImPyPy-(R)}^{\text{H}_2\text{N}\gamma}\text{-ImPyPyPy-}\beta\text{-C}_3\text{-R}_1$  ( $\beta$  =  $\beta$ -alanine,  $\text{C}_3$  = 1,3-diaminopropane linker) were measured by quantitative DNase I footprint titration experiments on plasmids containing the eight possible 5'-ANTGGTCA-3' and 5'-NATGGTCA-3' binding sites ( $\text{N}$  = A, T, C, or G). In addition, the binding affinities of five hairpin polyamide conjugates  $\text{ImImPyPy-(R)}^{\text{H}_2\text{N}\gamma}\text{-ImPyPyPy-(+)-R}_2$  ((+) = 3,3'-diamino-*N*-methyldipropylamine linker) were analyzed on the same two plasmids. For both types of compounds, the linker moiety exhibited 3- to 8-fold specificity for A,T over G,C, and the R group showed a similar 2- to 9-fold preference for A,T, with amide linkages conferring greater specificity than thiourea linkages. These results indicate that linkers and functional R groups on the tails of hairpin polyamide conjugates have recognition properties that should be considered in the design of these molecules to target DNA binding sites.

#### 4.1. Introduction

Aberrant gene expression has been implicated as the cause of many diseases.<sup>1,2</sup> Small molecules that can predictably regulate gene transcription could hold promise in human medicine.<sup>3</sup> Hairpin pyrrole-imidazole polyamides are synthetic ligands that target predetermined DNA sequences with affinities comparable to those of naturally occurring DNA-binding proteins in the nanomolar range.<sup>4,5</sup> These cell-permeable small molecules have been shown to localize to the nucleus of living cells<sup>6-10</sup> and regulate endogenous gene expression.<sup>11</sup>

Polyamides selectively bind in the minor groove of B-form DNA according to a set of “pairing rules,” where each heterocyclic ring pair targets a specific Watson-Crick base pair.<sup>4,5</sup> The antiparallel pairing of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) aromatic rings targets the C•G base pair (bp), while Im/Py discriminates G•C.<sup>12,13</sup> The Py/Py pair recognizes A•T and T•A base pairs in the DNA minor groove.<sup>14</sup> *N*-methylhydroxypyrrole (Hp) can be paired with Py to distinguish T•A from A•T.<sup>15,16</sup> Structural data from NMR and X-ray crystallography studies have elucidated these binding preferences.<sup>12,14,17,18</sup>

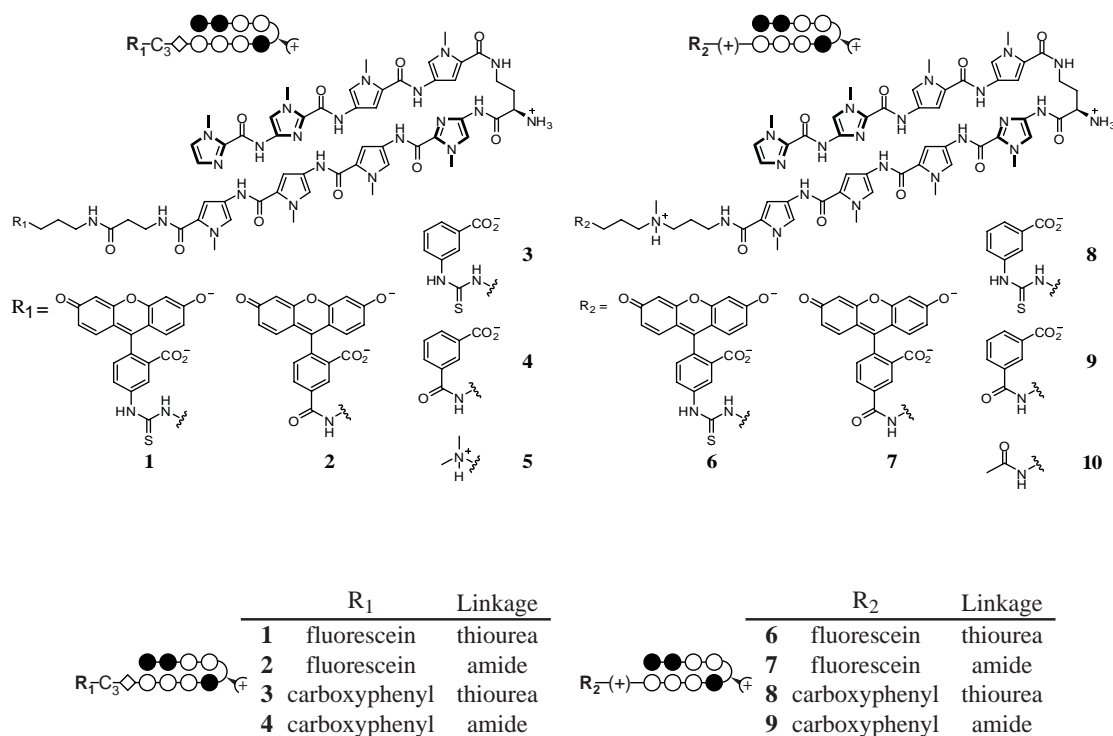
A covalent  $\gamma$ -aminobutyric acid ( $\gamma$ ) linkage pre-organizes two polyamide subunits in the DNA minor groove, prepaying the entropic cost of antiparallel binding.<sup>19</sup> Like the degenerate Py/Py pair, the  $\gamma$  residue at the hairpin “turn” is specific for A,T base pairs.<sup>20</sup> The use of chiral (*R*)-2,4-diaminobutyric acid to replace  $\gamma$  improves binding affinity and enforces a 5'→3' N→C binding orientation.<sup>21</sup> As with the hairpin turn, a variety of different “tails” can be designed. When polyamides are synthesized on solid phase with Boc- $\beta$ -Ala-PAM resin and cleaved with 3-dimethylamino-1-propylamine (Dp), the final compound contains a residual  $\beta$ -Dp tail at the C-terminus.<sup>22</sup> Quantitative DNase I footprint titration experiments using the six-ring hairpin polyamide ImPyPy-(*R*)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPy- $\beta$ -Dp have demonstrated that both the  $\beta$ -alanine residue and the Dp tail are W-specific recognition elements, where W = A or T, at the N-1 position.<sup>20</sup> Oxime resin can be used to synthesize hairpin polyamides that do not contain  $\beta$ -alanine at the C-terminus and allow

rational targeting of more sites.<sup>23</sup> Interestingly, polyamide-fluorophore conjugates without  $\beta$ -alanine tend to have better nuclear uptake properties than analogs with  $\beta$ -alanine.<sup>9,10</sup>

Hairpin polyamide cores can be functionalized at the turn and the tail to form polyamide conjugates suitable for a number of different applications. In particular, polyamide-fluorophore conjugates have been used to assay cellular uptake and nuclear localization in live cells using confocal microscopy.<sup>6-10</sup> In one notable instance, such a compound has been shown to downregulate endogenous gene expression in cell culture.<sup>11</sup> However, there is no systematic data comparing the binding affinities and specificities of polyamide conjugates with their parent molecules at the N-2 and N-3 positions relative to the Py/Im polyamide core recognition element. Increased understanding of the effects on DNA binding would enable rational polyamide conjugate design in future biological studies.

In this study, we examine the effects of substitution at the tail on DNA affinity, specificity, and cellular uptake (Figure 4.1). The ImImPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy-R polyamide core sequence was used in all ten compounds.<sup>16</sup> In the first series, five molecules of the type ImImPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy- $\beta$ -C<sub>3</sub>-R<sub>1</sub> ( $\beta$  =  $\beta$ -alanine, C<sub>3</sub> = 1,3-diaminopropane linker) were synthesized using Boc- $\beta$ -Ala-PAM resin. Fluorescein and 3-carboxyphenyl were conjugated to the polyamide using both thiourea and amide linkages. In the second series, five analogous conjugates of the type ImImPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy-(+)-R<sub>2</sub> ((+) = 3,3'-diamino-*N*-methyldipropylamine linker) were synthesized using oxime resin. The DNA binding energetics of all ten compounds were assayed by DNase I footprinting experiments on plasmids containing the eight possible 5'-ANTGGTCA-3' and 5'-NATGGTCA-3' binding sites (N = A, T, C, or G).





**Figure 4.1.** Structures of polyamide conjugates **1-10**. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C<sub>3</sub>”; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

## 4.2. Materials and methods

### 4.2.1. Polyamide synthesis

Polyamides were synthesized with Boc- $\beta$ -Ala-PAM resin (compounds **1-5**, Peptides International) or oxime resin (compounds **6-10**, Novabiochem) according to published manual solid-phase synthesis protocols.<sup>22,23</sup> The protected <sup>Fmoc</sup>HN $\gamma$ -turn amine was deprotected with 20% piperidine in DMF and reprotected as the Boc derivative with a solution of Boc<sub>2</sub>O (Fluka) and DIEA in DMF. The Boc-protected resin was cleaved with 1 mL of the appropriate amine (1,3-diaminopropane, compounds **1-4**; 3-dimethylamino-1-propylamine, compound **5**; 3,3'-diamino-*N*-methyldipropylamine, compounds **6-10**) at

37°C with agitation for 16 h. Products were purified by preparatory reverse-phase high-performance liquid chromatography (HPLC) on a Beckman Gold system using either a Waters Delta-Pak 25 × 100 mm, 15 μm 300 Å C<sub>18</sub> PrepPak Cartridge reverse-phase column or a Varian Dynamax 21.4 × 250 mm Microsorb 8 μm 300 Å C<sub>8</sub> reverse-phase column in 0.1% (w/v) TFA with acetonitrile as the eluent. The appropriate fractions were lyophilized after characterization by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF-MS) and/or electrospray ionization mass spectrometry (ESI-MS).

Conjugates were formed by reacting the isothiocyanate (compounds **1**, **3**, **6**, and **8**), succinimidyl ester (compounds **2** and **7**), or anhydride (compound **10**) with the polyamide in a solution of DIEA (20 equiv) and DMF for 1 h at room temperature. Conjugates **4** and **9** were formed by preactivating isophthalic acid (3.0 equiv) with PyBOP (2.9 equiv, Novabiochem) in a solution of DIEA (20 equiv) and DMF at 37°C for 30 min, followed by reaction of the activated solution with the polyamide for 1 h at room temperature. Conjugates were deprotected with neat TFA (Halocarbon) and triethylsilane for 30 min at room temperature before purification by preparatory reverse-phase HPLC. Lyophilization of the appropriate fractions yielded the polyamide conjugates **1-10**, which were characterized as described above. Extinction coefficients were calculated according to standard protocols.<sup>24</sup> Chemicals not otherwise specified were from Aldrich.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-β-C<sub>3</sub>-FITC (1).** Cleaved from β-Ala-PAM resin with neat 1,3-diaminopropane and conjugated with fluorescein-5-isothiocyanate (FITC, Invitrogen). UV-vis λ<sub>max</sub> = 312, 444 nm; ESI-MS *m/z* 1597.5 (C<sub>76</sub>H<sub>77</sub>N<sub>24</sub>O<sub>15</sub>S<sup>-</sup> calculated [M-H]<sup>-</sup> 1597.57)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-β-C<sub>3</sub>-FAM (2).** Cleaved from β-Ala-PAM resin with neat 1,3-diaminopropane and conjugated with 5-carboxyfluorescein, succinimidyl ester (5-FAM, SE, Invitrogen). UV-vis λ<sub>max</sub> = 315, 444 nm; MALDI-TOF-MS *m/z* 1568.74 (C<sub>76</sub>H<sub>78</sub>N<sub>23</sub>O<sub>16</sub><sup>+</sup> calculated [M+H]<sup>+</sup> 1568.60)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-β-C<sub>3</sub>-CPITC (3).** Cleaved from β-Ala-PAM resin with neat 1,3-diaminopropane and conjugated with 3-carboxyphenyl isothiocyanate (Trans World Chemicals). UV-vis  $\lambda_{\max}$  = 316 nm; ESI-MS  $m/z$  1387.6 (C<sub>63</sub>H<sub>71</sub>N<sub>24</sub>O<sub>12</sub>S<sup>-</sup> calculated [M-H]<sup>-</sup> 1387.54)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-β-C<sub>3</sub>-IPA (4).** Cleaved from β-Ala-PAM resin with neat 1,3-diaminopropane and conjugated with isophthalic acid. UV-vis  $\lambda_{\max}$  = 315 nm; MALDI-TOF-MS  $m/z$  1358.66 (C<sub>63</sub>H<sub>72</sub>N<sub>23</sub>O<sub>13</sub><sup>+</sup> calculated [M+H]<sup>+</sup> 1358.57)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-β-Dp (5).** Cleaved from β-Ala-PAM resin with neat 3-dimethylamino-1-propylamine. UV-vis  $\lambda_{\max}$  = 315 nm; MALDI-TOF-MS  $m/z$  1238.88 (C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> calculated [M+H]<sup>+</sup> 1238.58)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-(+)-FITC (6).** Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with fluorescein-5-isothiocyanate (FITC, Invitrogen). UV-vis  $\lambda_{\max}$  = 316, 444 nm; ESI-MS  $m/z$  1597.6 (C<sub>77</sub>H<sub>81</sub>N<sub>24</sub>O<sub>14</sub>S<sup>-</sup> calculated [M-H]<sup>-</sup> 1597.61)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-(+)-FAM (7).** Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with 5-carboxyfluorescein, succinimidyl ester (5-FAM, SE, Invitrogen). UV-vis  $\lambda_{\max}$  = 316, 444 nm; MALDI-TOF-MS  $m/z$  1568.52 (C<sub>77</sub>H<sub>82</sub>N<sub>23</sub>O<sub>15</sub><sup>+</sup> calculated [M+H]<sup>+</sup> 1568.64)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-(+)-CPITC (8).** Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with 3-carboxyphenyl isothiocyanate (Trans World Chemicals). UV-vis  $\lambda_{\max}$  = 316 nm; ESI-MS  $m/z$  1387.4 (C<sub>64</sub>H<sub>75</sub>N<sub>24</sub>O<sub>11</sub>S<sup>-</sup> calculated [M-H]<sup>-</sup> 1387.58)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-(+)-IPA (9).** Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with isophthalic acid. UV-vis  $\lambda_{\max}$  = 316 nm; MALDI-TOF-MS  $m/z$  1358.31 (C<sub>64</sub>H<sub>76</sub>N<sub>23</sub>O<sub>12</sub><sup>+</sup> calculated [M+H]<sup>+</sup> 1358.60)

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup><sub>γ</sub>-ImPyPyPy-(+)-Ac (10).** Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with acetic anhydride. UV-vis  $\lambda_{\max}$

= 316 nm; MALDI-TOF-MS  $m/z$  1252.55 ( $C_{58}H_{74}N_{23}O_{10}^+$  calculated  $[M+H]^+$  1252.60)

#### 4.2.2. Plasmid construction

Plasmids pCFH6 and pJWP6 were constructed according to standard protocols for DNA manipulation.<sup>25</sup> Plasmid pCFH6 was constructed by annealing the two 87-mer DNA oligonucleotides 5'-GATCGTGTAATCAATGGTCATAGCTGTGTAATCATTGGTCATAGCTGTGT AATCACTGGTCATAGCTGTGTAATCAGTGGTCATAGC-3' and 5'-AGCTGCTATGACCACTGATTACACAGCTATGACCAGTGATTACACAGCTA TGACCAATGATTACACAGCTATGACCATTGATTACAC-3' (Integrated DNA Technologies). Plasmid pJWP6 was constructed by annealing the two 91-mer DNA oligonucleotides 5'-GATCGTGTAATCAAATGGTCATAGCTGTGTAATCATATGGTCATAGCTGT GTAATCACATGGTCATAGCTGTGTAATCAGATGGTCATAGC-3' and 5'-AGCTGCTATGACCATCTGATTACACAGCTATGACCATGTGATTACACAGC TATGACCATATGATTACACAGCTATGACCATTGATTACAC-3'. Annealed oligonucleotides were ligated into the BamHI/HindIII (Roche) restriction fragment of pUC19 (Sigma). DNA sequencing of the constructed plasmids was performed at the Sequence Analysis Facility at the California Institute of Technology.

#### 4.2.3. Preparation of 5'-<sup>32</sup>P-end-labeled DNA

The forward primer 5'-AATTCGAGCTCGGTACCCGGG-3' was <sup>32</sup>P-labeled at the 5'-end. The reverse primer 5'-CTGGCACGACAGGTTTCCCGAC-3' was used to amplify plasmids pCFH6 and pJWP6 as previously described.<sup>24</sup> PCR products (5'-<sup>32</sup>P-end-labeled, 291 bp for pCFH6, 295 bp for pJWP6) were isolated according to standard protocols.<sup>24</sup>

#### 4.2.4. Quantitative DNase I footprint titrations

Quantitative DNase I footprint titration experiments were performed on the 5'-<sup>32</sup>P-end-labeled PCR products of plasmids pCFH6 and pJWP6 according to standard protocols.<sup>24</sup> Radiolabeled DNA was equilibrated with polyamide solutions for 14-16 h at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0 prior to DNase I cleavage. Chemical sequencing reactions were performed according to published methods.<sup>26,27</sup>

#### 4.2.5. Cell cultures

Human cancer cell lines HeLa, MCF7, and PC3 were cultured in a 5% CO<sub>2</sub> atmosphere at 37°C in supplemented DMEM (HeLa, GIBCO) or RPMI medium 1640 (MCF7 and PC3, GIBCO).<sup>9,10</sup> All media were supplemented with 10% fetal bovine serum (Omega Scientific) and 1% penicillin/streptomycin solution (Mediatech).

#### 4.2.6. Confocal microscopy

Cell lines were trypsinized (Mediatech) for 5 min at 37°C, centrifuged for 10 min at 4°C at 100 g, and resuspended in fresh medium to a concentration of  $1.33 \times 10^5$  cells/mL (HeLa) or  $3.33 \times 10^5$  cells/mL (MCF7 and PC3).<sup>9,10</sup> Incubations were performed by adding 150 µL of cells into culture dishes equipped with glass bottoms for direct imaging (MatTek). Cells were grown in the glass-bottom culture dishes for 24 h. The medium was then removed and replaced with 147 µL of fresh medium, followed by addition of 3 µL of the 100 µM polyamide solution for a final polyamide concentration of 2 µM. Cells were incubated in a 5% CO<sub>2</sub> atmosphere at 37°C for 12 h. Imaging was performed with a 40× oil-immersion objective lens on a Zeiss LSM 5 Pascal inverted laser scanning microscope. Polyamide-fluorescein conjugate fluorescence and visible light images were obtained using standard filter sets for fluorescein.<sup>9,10</sup>

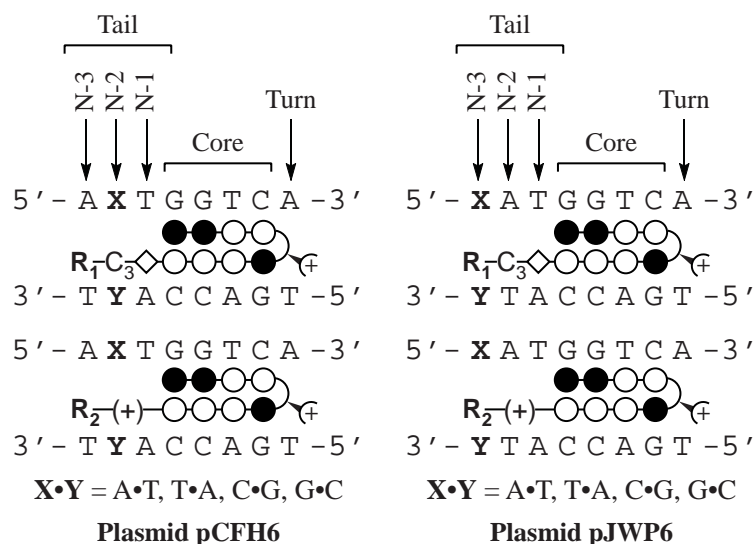
### 4.3. Results

#### 4.3.1. DNA binding affinity and sequence specificity of polyamide conjugates

Ten hairpin polyamide conjugates were synthesized by solid-phase methods (Figure 4.1). Compounds **1-5**, synthesized on  $\beta$ -Ala-PAM resin, were of the type ImImPyPy-(*R*)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy- $\beta$ -C<sub>3</sub>-R<sub>1</sub> ( $\beta$  =  $\beta$ -alanine, C<sub>3</sub> = 1,3-diaminopropane linker). Compounds **6-10**, synthesized on oxime resin, were of the type ImImPyPy-(*R*)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy-(+)-R<sub>2</sub> ((+) = 3,3'-diamino-*N*-methyldipropylamine linker). Each series of five compounds contained thiourea-linked fluorescein, amide-linked fluorescein, thiourea-linked 3-carboxyphenyl, amide-linked 3-carboxyphenyl, and control (Dp or Ac) conjugates.

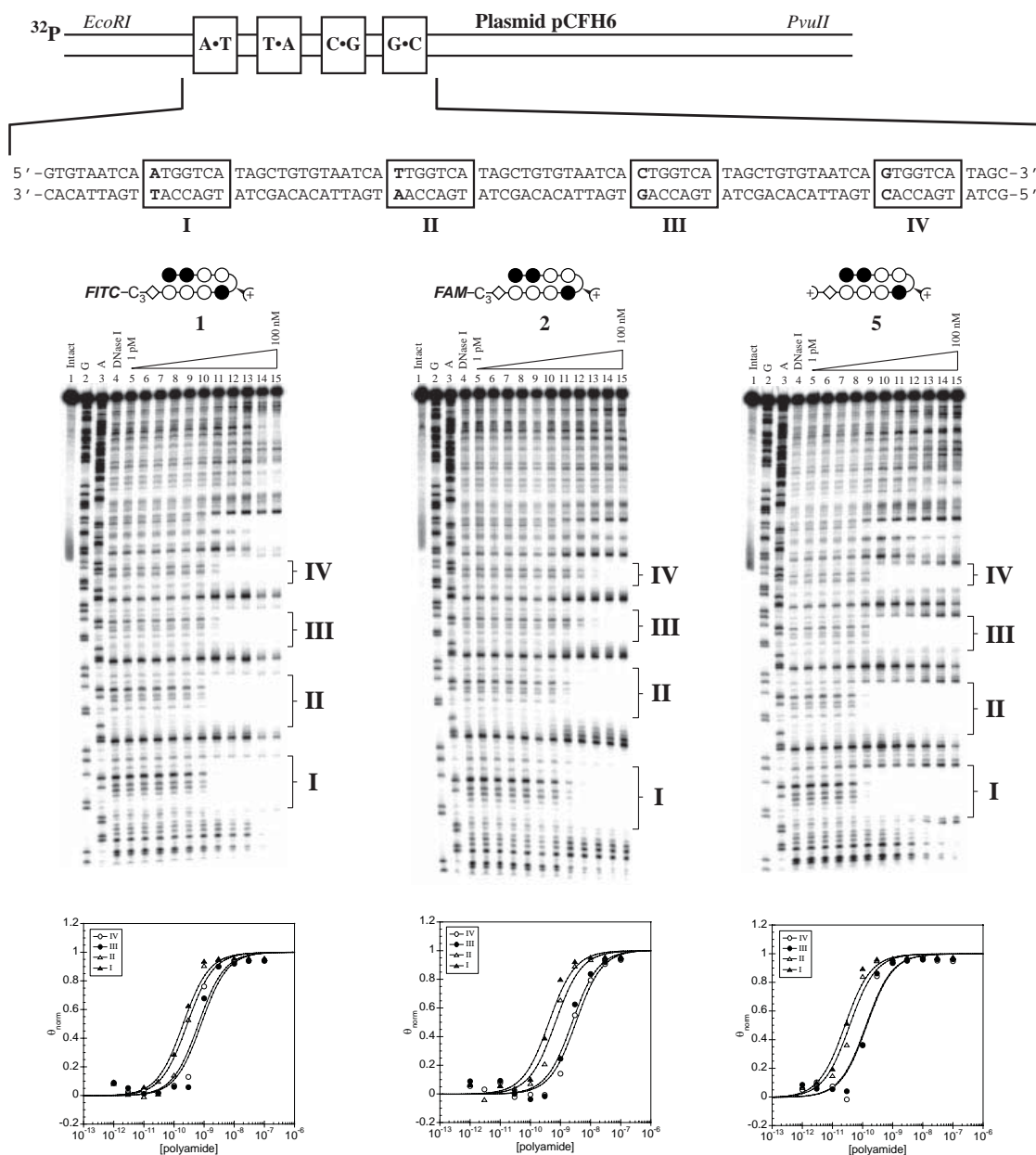
Plasmid pCFH6 was designed to contain four binding sites of the form 5'-ANTGGTCA-3' (N = A, T, C, or G) (Figure 4.2). The four Watson-Crick base pairs were permuted at the 5'-end of the 7-bp binding site to test sequence specificity at the N-2 position of the C-terminal tail. Following the results of footprinting experiments on plasmid pCFH6 (Figure 4.3 and Table 4.1), plasmid pJWP6 was designed to contain four binding sites of the form 5'-NATGGTCA-3' with permutation at the 5'-end of the 8-bp binding site to test sequence specificity at the N-3 position of the C-terminal tail (Figure 4.2).

Quantitative DNase I footprint titration experiments were performed on the 291 bp, 5'-<sup>32</sup>P-end-labeled PCR product of plasmid pCFH6 and the 295 bp, 5'-<sup>32</sup>P-end-labeled PCR product of pJWP6.<sup>24</sup> Each polyamide was assayed on a plasmid containing four potential match sites in which a single Watson-Crick base pair was permuted. On plasmid pCFH6, the thiourea-linked polyamide-fluorescein conjugate **1** bound the 7-bp binding sites 5'-ATGGTCA-3' and 5'-TTGGTCA-3' with equilibrium association constants of  $6.3 \times 10^9 \text{ M}^{-1}$  and  $5.1 \times 10^9 \text{ M}^{-1}$ , respectively (Figure 4.3 and Table 4.1). Compound **1** bound the other two designed sites (C•G and G•C) with  $K_a = 1.8 \times 10^9 \text{ M}^{-1}$ , indicating that this molecule exhibits 4-fold specificity for A•T and 3-fold specificity for T•A. The amide-linked polyamide-fluorescein conjugate **2** showed decreased binding affinity but strong sequence specificity, as its equilibrium association constant of  $3.2 \times 10^9 \text{ M}^{-1}$  displayed an



**Figure 4.2.** Designed binding sites on plasmids pCFH6 and pJWP6. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C<sub>3</sub>”; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

8-fold preference for A•T over C,G. Both of the carboxyphenyl conjugates **3** and **4** bound their match sites with a binding affinity of greater than  $1 \times 10^{10} \text{ M}^{-1}$  and exhibited similar 4-fold specificity. As expected, the smaller carboxyphenyl conjugates bound more tightly than fluorescein conjugates. The binding profiles of compounds **3** and **4** were very similar even though these conjugates were created with different linkages. The control compound **5** had the highest equilibrium association constant in the series, as its  $K_a = 3.2 \times 10^9 \text{ M}^{-1}$  for the A•T site was 5-fold higher than the  $K_a$  for the C,G sites. The A,T preference seen in **5** is consistent with the trend observed for the analogous six-ring polyamide ImImPy- $(R)^{H_2N}\gamma$ -ImPyPy- $\beta$ -Dp, which exhibited 23-fold and 11-fold specificity for A•T and T•A, respectively.<sup>20</sup> The binding energetics data for compounds **1-4** demonstrate that the 1,3-diaminopropane linker also displays A,T specificity.



**Figure 4.3.** Quantitative DNase I footprint titration experiments for polyamide conjugates 1, 2, and 5 on the 291 bp, 5'-<sup>32</sup>P-end-labeled PCR product of plasmid pCFH6: lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-15, 1 pM, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, and 100 nM polyamide, respectively. Binding isotherms for the four designed sites are shown below each footprinting gel;  $\theta_{\text{norm}}$  values were calculated according to published methods.<sup>24</sup> Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as "C<sub>3</sub>"; the dimethylaminopropylamine tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.



**Table 4.1.** Equilibrium association constants<sup>a</sup>

pCFH6	R	Linkage	5'-AATGGTCA-3'	5'-ATTGGTCA-3'	5'-ACTGGTCA-3'	5'-AGTGGTCA-3'
<b>1</b>	fluorescein	thiourea	6.3 ( $\pm 1.9$ ) $\times 10^9$ [4]	5.1 ( $\pm 1.7$ ) $\times 10^9$ [3]	1.7 ( $\pm 0.5$ ) $\times 10^9$ [1]	1.8 ( $\pm 0.5$ ) $\times 10^9$ [1]
<b>2</b>	fluorescein	amide	3.2 ( $\pm 1.1$ ) $\times 10^9$ [8]	2.1 ( $\pm 0.6$ ) $\times 10^9$ [5]	4.0 ( $\pm 1.2$ ) $\times 10^8$ [1]	4.1 ( $\pm 1.0$ ) $\times 10^8$ [1]
<b>3</b>	carboxyphenyl	thiourea	1.3 ( $\pm 0.5$ ) $\times 10^{10}$ [4]	1.2 ( $\pm 0.5$ ) $\times 10^{10}$ [3]	3.7 ( $\pm 1.8$ ) $\times 10^9$ [1]	4.1 ( $\pm 1.8$ ) $\times 10^9$ [1]
<b>4</b>	carboxyphenyl	amide	1.5 ( $\pm 0.4$ ) $\times 10^{10}$ [4]	1.2 ( $\pm 0.4$ ) $\times 10^{10}$ [3]	4.1 ( $\pm 1.4$ ) $\times 10^9$ [1]	4.8 ( $\pm 1.9$ ) $\times 10^9$ [1]
<b>5</b>	—	—	3.7 ( $\pm 0.5$ ) $\times 10^{10}$ [5]	2.5 ( $\pm 0.1$ ) $\times 10^{10}$ [3]	8.7 ( $\pm 1.4$ ) $\times 10^9$ [1]	8.0 ( $\pm 0.7$ ) $\times 10^9$ [1]
<b>6</b>	fluorescein	thiourea	1.7 ( $\pm 0.2$ ) $\times 10^{10}$ [3]	1.6 ( $\pm 0.2$ ) $\times 10^{10}$ [3]	6.8 ( $\pm 0.3$ ) $\times 10^9$ [1]	6.0 ( $\pm 0.7$ ) $\times 10^9$ [1]
<b>7</b>	fluorescein	amide	6.5 ( $\pm 1.0$ ) $\times 10^9$ [7]	6.6 ( $\pm 1.2$ ) $\times 10^9$ [7]	2.0 ( $\pm 0.3$ ) $\times 10^9$ [2]	9.7 ( $\pm 2.0$ ) $\times 10^8$ [1]
<b>8</b>	carboxyphenyl	thiourea	8.8 ( $\pm 1.3$ ) $\times 10^9$ [3]	8.8 ( $\pm 1.0$ ) $\times 10^9$ [3]	3.5 ( $\pm 1.1$ ) $\times 10^9$ [1]	2.6 ( $\pm 0.5$ ) $\times 10^9$ [1]
<b>9</b>	carboxyphenyl	amide	1.0 ( $\pm 0.2$ ) $\times 10^{10}$ [3]	1.0 ( $\pm 0.2$ ) $\times 10^{10}$ [3]	5.1 ( $\pm 0.4$ ) $\times 10^9$ [1]	3.3 ( $\pm 0.4$ ) $\times 10^9$ [1]
<b>10</b>	acetyl	amide	3.1 ( $\pm 0.5$ ) $\times 10^{10}$ [3]	2.9 ( $\pm 0.3$ ) $\times 10^{10}$ [3]	1.5 ( $\pm 0.1$ ) $\times 10^{10}$ [1]	1.1 ( $\pm 0.1$ ) $\times 10^{10}$ [1]

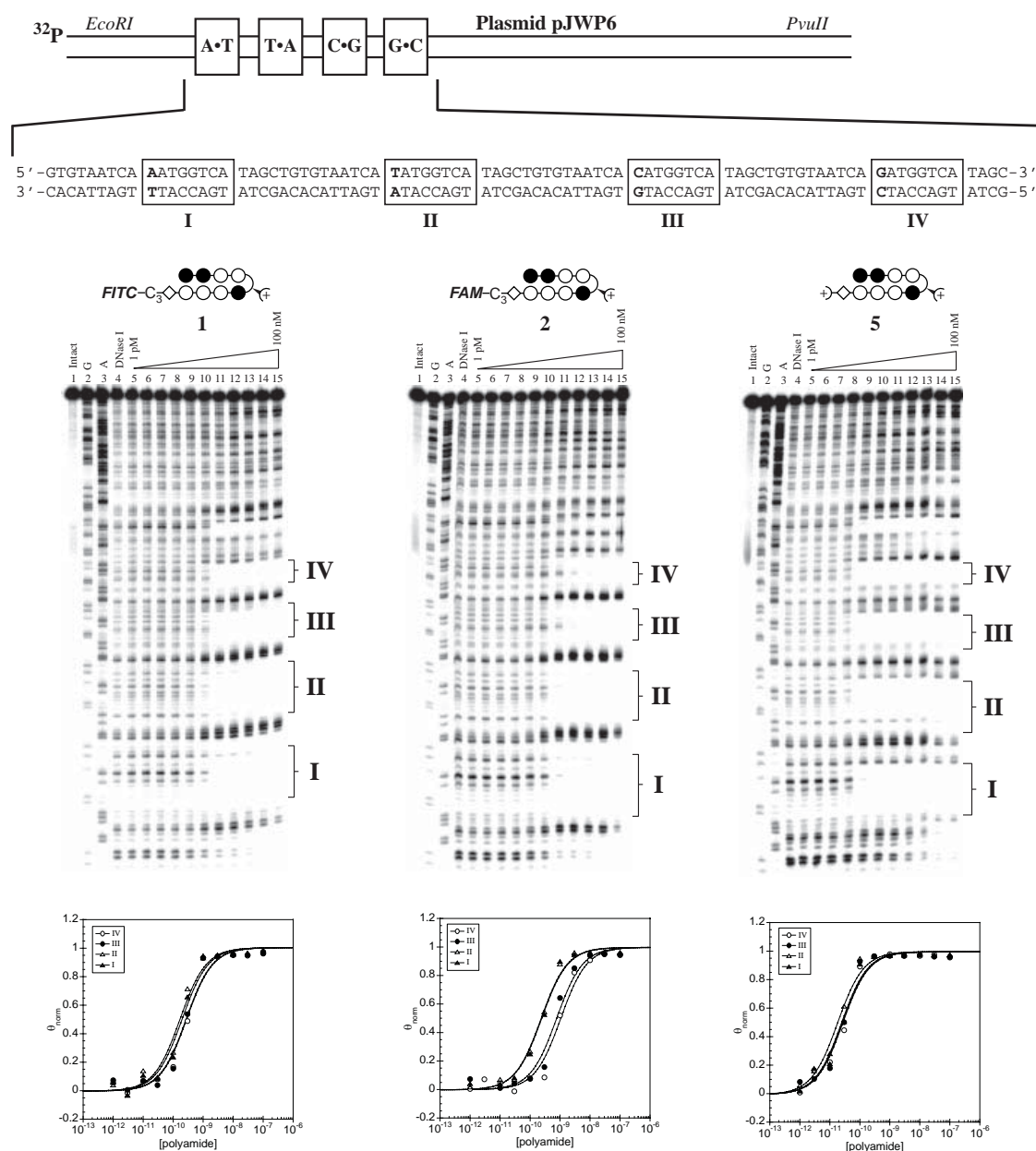
<sup>a</sup>  $K_a$  ( $M^{-1}$ ) values reported are the mean values from at least three DNase I footprint titration experiments. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0. Specificity is indicated in brackets and calculated as  $K_a$  (site) /  $K_a$  (min).

Compounds **6-10** have an additional positive charge due to the protonation of the 3,3'-diamino-*N*-methyldipropylamine linker at physiological pH. The thiourea-linked fluorescein conjugate **6** bound the A•T site with  $K_a = 1.7 \times 10^{10} M^{-1}$  (3-fold specificity), and the amide-linked fluorescein conjugate **7** bound the T•A site with  $K_a = 6.6 \times 10^9 M^{-1}$  (7-fold specificity). When compared with **1** and **2**, oxime compounds **6** and **7** bound DNA more tightly but with comparable specificity. Compounds **8** and **9** bound A,T sites with  $K_a = 8.8 \times 10^9 M^{-1}$  and  $1.0 \times 10^{10} M^{-1}$ , respectively. The acetylated control compound **10** displayed an equilibrium association constant of  $3.1 \times 10^{10} M^{-1}$ , and, for all three of these compounds **8-10**, their 3-fold preference for A,T was slightly less than that shown by compounds **3-5**. These experiments indicate that the 3,3'-diamino-*N*-methyldipropylamine linker is A,T-specific; however, the linker likely spans the two base pairs 5'-WT-3'.

Based on the conclusion that both linkers specify for A,T base pairs, plasmid pJWP6 was designed to test whether the conjugated R group can serve as a recognition element, by placing an A•T base pair at the N-2 position and varying the N-3 base pair across the four Watson-Crick base pairs. Control polyamide **5** exhibited the lowest specificity for T•A ( $4.5 \times 10^{10} \text{ M}^{-1}$ ) over G•C ( $2.7 \times 10^{10} \text{ M}^{-1}$ ) (Figure 4.4 and Table 4.2). The other nine conjugates showed at least 2-fold specificity for A,T, and, as with the previous plasmid, amide-linked conjugates displayed stronger specificity than thiourea-linked compounds. One highlight was the amide-linked fluorescein compound **7**, whose equilibrium association constant of  $6.0 \times 10^9 \text{ M}^{-1}$  for A•T was nearly an order of magnitude greater than the  $K_a$  of  $6.8 \times 10^8 \text{ M}^{-1}$  for G•C. In general, rational design of the linker and R group leads to the recognition of an eighth base pair at the N-3 position.

#### **4.3.2. Nuclear localization of polyamide-fluorescein conjugates**

Human cancer cell lines HeLa, MCF7, and PC3 were cultured for cell uptake studies using confocal laser scanning microscopy. For compounds **1**, **6**, and **7**, the nuclear staining exceeded that of the medium in all cell lines (Table 4.3). For compound **2**, there was some nuclear staining in HeLa cells, although it was similar to that of the medium, and little nuclear staining was observed in MCF7 and PC3 cells. The cellular uptake data for compound **6** were previously reported.<sup>10</sup> Because FITC and FAM were considered to be interchangeable previously, the results for compound **7** were expected. However, the positive nuclear localization data for the  $\beta$ -alanine-linked fluorescein conjugate **1** was a pleasant surprise, suggesting that 1,3-diaminopropane could be a viable linker for  $\beta$ -alanine-linked conjugates synthesized on PAM resin. The mixed results for compound **2** are a reminder that thiourea and amide linkages are often different in terms of cellular uptake as well as sequence specificity.



**Figure 4.4.** Quantitative DNase I footprint titration experiments for polyamide conjugates **1**, **2**, and **5** on the 295 bp, 5’-<sup>32</sup>P-end-labeled PCR product of plasmid pJWP6: lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-15, 1 pM, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, and 100 nM polyamide, respectively. Binding isotherms for the four designed sites are shown below each footprinting gel;  $\theta_{\text{norm}}$  values were calculated according to published methods.<sup>24</sup> Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C<sub>3</sub>”; the dimethylaminopropylamine tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

**Table 4.2.** Equilibrium association constants<sup>a</sup>





pJWP6	R	Linkage	5'-AATGGTCA-3'	5'-TATGGTCA-3'	5'-CATGGTCA-3'	5'-GATGGTCA-3'
<b>1</b>	fluorescein	thiourea	5.1 ( $\pm 0.8$ ) $\times 10^9$ [2]	5.4 ( $\pm 1.1$ ) $\times 10^9$ [2]	3.6 ( $\pm 0.7$ ) $\times 10^9$ [1]	3.2 ( $\pm 0.6$ ) $\times 10^9$ [1]
<b>2</b>	fluorescein	amide	3.7 ( $\pm 1.0$ ) $\times 10^9$ [4]	3.5 ( $\pm 1.4$ ) $\times 10^9$ [4]	1.2 ( $\pm 0.2$ ) $\times 10^9$ [1]	8.8 ( $\pm 1.1$ ) $\times 10^8$ [1]
<b>3</b>	carboxyphenyl	thiourea	1.4 ( $\pm 0.6$ ) $\times 10^{10}$ [2]	1.5 ( $\pm 0.6$ ) $\times 10^{10}$ [2]	8.3 ( $\pm 2.3$ ) $\times 10^9$ [1]	7.1 ( $\pm 2.3$ ) $\times 10^9$ [1]
<b>4</b>	carboxyphenyl	amide	1.5 ( $\pm 0.2$ ) $\times 10^{10}$ [3]	1.5 ( $\pm 0.1$ ) $\times 10^{10}$ [3]	6.1 ( $\pm 1.1$ ) $\times 10^9$ [1]	4.7 ( $\pm 1.9$ ) $\times 10^9$ [1]
<b>5</b>	–	–	3.1 ( $\pm 1.5$ ) $\times 10^{10}$ [1]	4.5 ( $\pm 1.9$ ) $\times 10^{10}$ [2]	3.0 ( $\pm 1.2$ ) $\times 10^{10}$ [1]	2.7 ( $\pm 1.2$ ) $\times 10^{10}$ [1]
<b>6</b>	fluorescein	thiourea	1.6 ( $\pm 0.4$ ) $\times 10^{10}$ [2]	2.3 ( $\pm 0.4$ ) $\times 10^{10}$ [3]	1.2 ( $\pm 0.2$ ) $\times 10^{10}$ [1]	8.9 ( $\pm 1.1$ ) $\times 10^9$ [1]
<b>7</b>	fluorescein	amide	6.0 ( $\pm 3.9$ ) $\times 10^9$ [9]	4.9 ( $\pm 3.1$ ) $\times 10^9$ [7]	2.1 ( $\pm 0.6$ ) $\times 10^9$ [3]	6.8 ( $\pm 0.9$ ) $\times 10^8$ [1]
<b>8</b>	carboxyphenyl	thiourea	1.9 ( $\pm 0.4$ ) $\times 10^{10}$ [2]	2.2 ( $\pm 0.4$ ) $\times 10^{10}$ [3]	1.1 ( $\pm 0.1$ ) $\times 10^{10}$ [1]	8.3 ( $\pm 1.5$ ) $\times 10^9$ [1]
<b>9</b>	carboxyphenyl	amide	3.2 ( $\pm 0.2$ ) $\times 10^{10}$ [6]	2.2 ( $\pm 0.5$ ) $\times 10^{10}$ [4]	1.1 ( $\pm 0.1$ ) $\times 10^{10}$ [2]	5.1 ( $\pm 1.2$ ) $\times 10^9$ [1]
<b>10</b>	acetyl	amide	3.6 ( $\pm 0.5$ ) $\times 10^{10}$ [3]	3.2 ( $\pm 0.3$ ) $\times 10^{10}$ [3]	1.4 ( $\pm 0.1$ ) $\times 10^{10}$ [1]	1.1 ( $\pm 0.1$ ) $\times 10^{10}$ [1]

<sup>a</sup>  $K_a$  ( $M^{-1}$ ) values reported are the mean values from at least three DNase I footprint titration experiments. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM  $MgCl_2$ , and 5 mM  $CaCl_2$  at pH 7.0. Specificity is indicated in brackets and calculated as  $K_a(\text{site}) / K_a(\text{min})$ .

#### 4.4. Discussion

DNase I footprinting experiments indicate that hairpin polyamide conjugates **1-4** and **6-9** have decreased binding affinity relative to their parent polyamide compounds **5** and **10**, respectively. This trend becomes more severe as the steric bulk of the conjugated R group increases, so fluorescein conjugates are impacted more than 3-carboxyphenyl conjugates. Control polyamides **5** and **10** bind their preferred match sites with an equilibrium association constant of  $3 \times 10^{10} M^{-1}$ , while 3-carboxyphenyl and fluorescein conjugates bind in the mid- $10^9 M^{-1}$  range. The subnanomolar binding affinities of these conjugates, coupled with good sequence specificity, make these compounds strong leads for continued biological applications.

**Table 4.3.** Nuclear localization of polyamide-fluorescein conjugates<sup>a</sup>

	Compound	Linkage	HeLa	MCF7	PC3
	1	thiourea	++	++	++
	2	amide	+	–	–
	6	thiourea	++	++	++
	7	amide	++	++	++

<sup>a</sup> Uptake profile of polyamide-fluorescein conjugates in three cell lines. ++, Nuclear staining exceeds that of the medium; +, nuclear staining less than or equal to that of the medium, but still prominent; –, very little nuclear staining, with the most fluorescence seen in the cytoplasm and/or medium. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C<sub>3</sub>”; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

For conjugates synthesized on Boc- $\beta$ -Ala-PAM resin, it appears that the 1,3-diaminopropane linker is A,T-specific, with the C<sub>3</sub> linker extending out only one additional Watson-Crick base pair, as evidenced by the footprinting of control compound **5**. For oxime compounds, the 3,3'-diamino-*N*-methyldipropylamine linker prefers 5'-WT-3' base pair steps over 5'-GT-3' or 5'-CT-3'. In general,  $\beta$ -C<sub>3</sub> and 3,3'-diamino-*N*-methyldipropylamine linkers lead to compounds with similar binding affinities, although they might contain different charges at physiological pH. Notably, amide-linked compounds have higher sequence specificity, and amide-linked fluorescein conjugates have an especially strong A,T preference. Thus, it is now possible to take advantage of the inherent W-specific recognition properties of linkers and R groups to design high-affinity, high-specificity DNA-binding compounds taking into account binding preferences at the N-2 and N-3 positions.

The positive cellular uptake data is highly encouraging as hairpin polyamide conjugates are used in cell culture experiments. The nuclear localization of compound **1** suggests that the  $\beta$ -alanine residue is not lethal when the appropriate amine linker is utilized, so PAM resin can be used in the synthesis of this class of conjugates going forward. The uptake of compounds **6** and **7** is consistent with the observed trend that conjugates synthesized on oxime resin (without the  $\beta$ -alanine residue) tend to localize well in human cancer cell lines. Having the ability to cleave from either resin with different linkers broadens the range of cell-permeable hairpin polyamide conjugates that are suitable for regulation of gene transcription.

**Acknowledgements.** We are grateful to the National Institutes of Health for research support. James Puckett synthesized polyamides **6-10** and prepared plasmid pJWP6.

## References

1. Pandolfi, P.P. (2001) Transcription therapy for cancer. *Oncogene*, **20**, 3116-3127.
2. Darnell, J.E. (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 740-749.
3. Arkin, M.R. and Wells, J.A. (2004) Small-molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nat. Rev. Drug Discov.*, **3**, 301-317.
4. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.*, **9**, 2215-2235.
5. 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.
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. Gygi, M.P., Ferguson, M.D., Mefford, H.C., Lund, K.P., O'Day, C., Zhou, P., Friedman, C., van den Engh, G., Stolowitz, M.L. and Trask, B.J. (2002) Use of fluorescent sequence-specific polyamides to discriminate human chromosomes by microscopy and flow cytometry. *Nucleic Acids Res.*, **30**, 2790-2799.
8. 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. *Bioorg. Med. Chem. Lett.*, **13**, 1565-1570.
9. 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. USA*, **100**, 12063-12068.
  10. 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.
  11. 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. USA*, **101**, 16768-16773.
  12. Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B. (1992) 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. *Proc. Natl. Acad. Sci. USA*, **89**, 7586-7590.
  13. Wade, W.S., Mrksich, M. and Dervan, P.B. (1992) Design of Peptides That Bind in the Minor Groove of DNA at 5'-(a,T)G(a,T)C(a,T)-3' Sequences by a Dimeric Side-by-Side Motif. *J. Am. Chem. Soc.*, **114**, 8783-8794.



14. Pelton, J.G. and Wemmer, D.E. (1989) Structural Characterization of a 2-1 Distamycin A.D(Cgcaaattggc) Complex by Two-Dimensional Nmr. *Proc. Natl. Acad. Sci. USA*, **86**, 5723-5727.
15. Trauger, J.W., Baird, E.E. and Dervan, P.B. (1996) Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature*, **382**, 559-561.
16. 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.
17. Kielkopf, C.L., Baird, E.E., Dervan, P.D. and Rees, D.C. (1998) Structural basis for G center dot C recognition in the DNA minor groove. *Nat. Struct. Biol.*, **5**, 104-109.
18. Kielkopf, C.L., White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) A structural basis for recognition of A center dot T and T center dot A base pairs in the minor groove of B-DNA. *Science*, **282**, 111-115.
19. Mrksich, M., Parks, M.E. and Dervan, P.B. (1994) Hairpin Peptide Motif - a New Class of Oligopeptides for Sequence-Specific Recognition in the Minor-Groove of Double-Helical DNA. *J. Am. Chem. Soc.*, **116**, 7983-7988.
20. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1999) Effects of gamma-turn and beta-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *J. Am. Chem. Soc.*, **121**, 1113-1120.

21. Herman, D.M., Baird, E.E. and Dervan, P.B. (1998) Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.*, **120**, 1382-1391.
22. 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.
23. 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.
24. Trauger, J.W. and Dervan, P.B. (2001) Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. *Methods Enzymol.*, **340**, 450-466.
25. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26. Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.*, **65**, 499-560.
27. Iverson, B.L. and Dervan, P.B. (1987) Adenine Specific DNA Chemical Sequencing Reaction. *Nucleic Acids Res.*, **15**, 7823-7830.

## **Chapter 4B**

### **Inhibition of VEGF with $\beta$ -Alanine- $C_3$ -Linked Hairpin Polyamide Conjugates**

This project was done in collaboration with Nicholas Nickols (Dervan group, Caltech).

Nicholas Nickols synthesized polyamide **1**.

**Abstract**

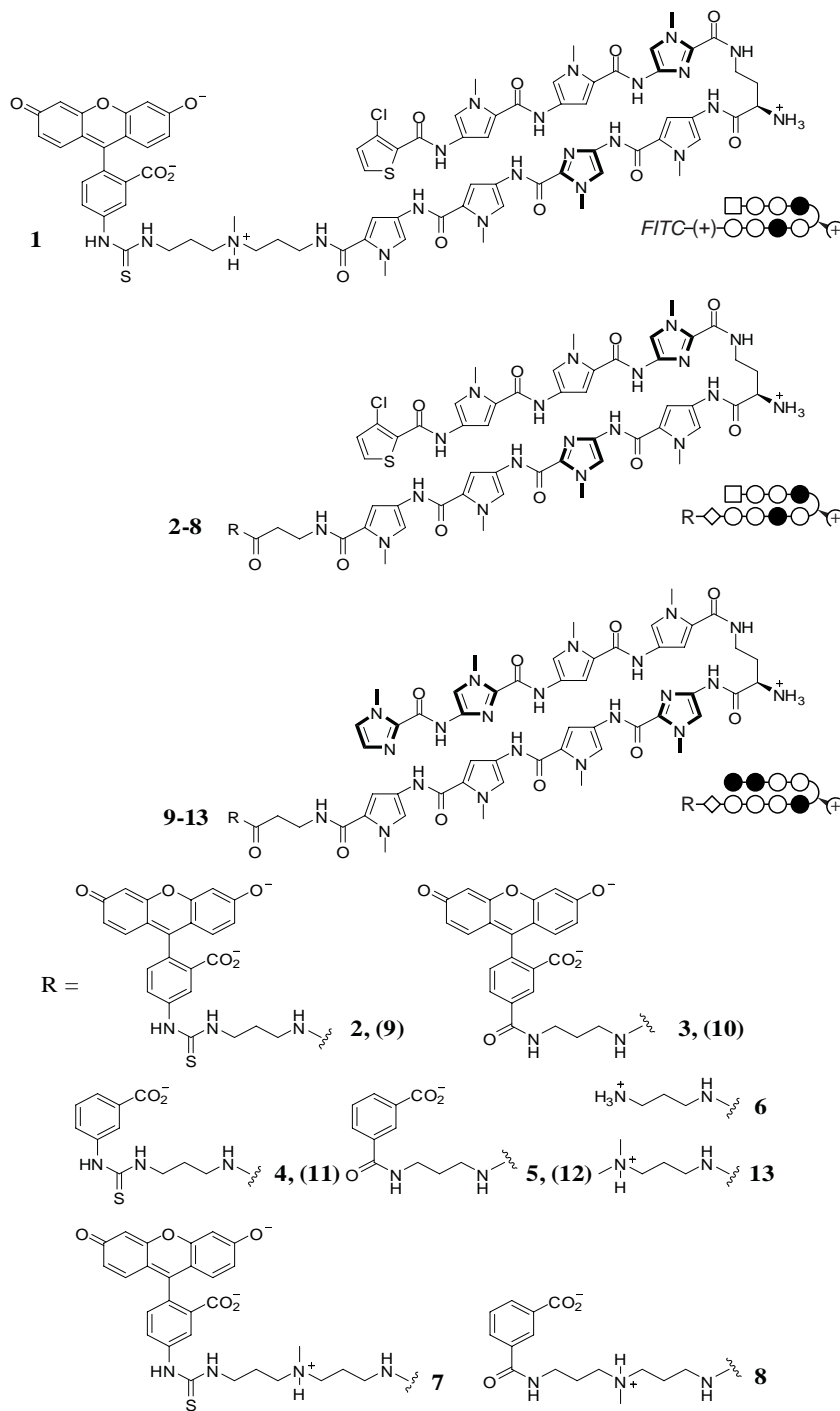
Hairpin polyamide conjugates synthesized on Boc- $\beta$ -Ala-PAM resin and cleaved with 1,3-diaminopropane have been observed to exhibit favorable nuclear uptake properties in human cancer cell lines. In this study, DNase I footprint titrations indicate that these small molecules bind the hypoxia response element 5'-TACGTG-3' with affinities of  $K_a > 3 \times 10^9 \text{ M}^{-1}$ . Polyamide-fluorescein conjugates in this class of compounds traffic to the nucleus of live HeLa cells. A series of hairpin polyamide conjugates with the  $\beta$ -alanine- $\text{C}_3$  linker was assayed by quantitative RT-PCR for effects on hypoxia-inducible transcription of vascular endothelial growth factor (VEGF). Fluorescein and isophthalic acid  $\beta$ -alanine- $\text{C}_3$ -linked polyamide conjugates **2**, **4**, and **5** are shown to decrease VEGF mRNA expression.

#### 4.5. Introduction

Pyrrole-imidazole polyamides are synthetic ligands that bind the DNA minor groove at subnanomolar concentrations in a sequence-specific manner.<sup>1,2</sup> Polyamide-fluorophore conjugates have been imaged by confocal laser scanning microscopy to determine the extent of nuclear localization.<sup>3-6</sup> These experiments showed that the use of a fluorescein dye on the amine tail improved uptake.<sup>4</sup> Another trend that emerged was that conjugates containing the  $\beta$ -alanine linkage exhibited poor cellular uptake across a number of cell lines.<sup>5,6</sup> For instance, ImImPyPy- $\gamma$ -ImPyPyPy-(+)-FITC localized to the nucleus of all cell lines tested, while ImImPyPy- $\gamma$ -ImPyPyPy- $\beta$ -(+)-FITC was excluded from the nucleus.<sup>5</sup>

However, these studies also included several examples of  $\beta$ -alanine-linked polyamide-fluorescein conjugates that traffic unaided to the nucleus.<sup>5,6</sup> These observations suggested that the composition of the tail linker could have a positive effect on uptake. Recent studies on polyamide-fluorescein conjugates synthesized with Boc- $\beta$ -Ala-PAM resin using a 1,3-diaminopropane linker showed improved nuclear localization. Specifically, imaging experiments with  $\beta$ -alanine- $C_3$ -linked FAM conjugates containing a WM moiety showed that these polyamides localized to the nucleus of HeLa, MCF-7, and PC3 human cancer cell lines.<sup>7</sup> Furthermore, in Chapter 4A, the hairpin polyamide conjugate ImImPyPy-(*R*)<sup>H<sub>2</sub>N</sup> $\gamma$ -ImPyPyPy- $\beta$ - $C_3$ -FITC was observed to localize to the nucleus of the same three cell lines. Interestingly, in this case, the FAM conjugate showed poorer uptake. These promising examples indicate that an uncharged linker could allow the use of Boc- $\beta$ -Ala-PAM resin in the synthesis of cell-permeable hairpin polyamide conjugates suitable for gene regulation.

A series of hairpin polyamide conjugates **2-5** was synthesized on Boc- $\beta$ -Ala-PAM resin, and 1,3-diaminopropane was used to link a fluorescein or isophthalic acid moiety to the polyamide tail (Figure 4.5). The DNA binding affinity of these compounds for the hypoxia response element 5'-TACGTG-3' was determined by DNase I footprint titrations. Nuclear localization of the polyamide-fluorescein conjugates was imaged in HeLa cells. Targeting of polyamides to the hypoxia response element has been shown to disrupt the binding of



**Figure 4.5.** Structures of polyamide conjugates **1-13**. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square;  $\beta$ -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C<sub>3</sub>”; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

hypoxia-inducible factor as measured by reduction in vascular endothelial growth factor (VEGF) gene expression.<sup>8-10</sup> The activity of  $\beta$ -alanine- $C_3$ -linked polyamide conjugates **2-5** was assayed by quantitative RT-PCR to determine the level of VEGF mRNA transcript.

## 4.6. Materials and methods

### 4.6.1. Polyamide synthesis

Polyamides were synthesized with oxime resin (compound **1**, Novabiochem) or Boc- $\beta$ -Ala-PAM resin (compounds **2-13**, Peptides International) according to published manual solid-phase synthesis protocols.<sup>11,12</sup> For compounds **1-13**, *N*- $\alpha$ -9-fluorenylmethoxycarbonyl-*N*- $\gamma$ -*tert*-butoxycarbonyl-D-2,4-diaminobutyric acid (Fmoc-D-Dab(Boc)-OH, Peptides International) was used for the turn linkage. The protected <sup>FmocHN</sup> $\gamma$ -turn amine was deprotected with 20% piperidine in DMF and reprotected as the Boc derivative with a solution of Boc<sub>2</sub>O (Fluka) and DIEA in DMF. The Boc-protected resin was cleaved with 1 mL of the appropriate amine (3,3'-diamino-*N*-methyldipropylamine, compounds **1**, **7**, **8**; 1,3-diaminopropane, compounds **2-6**, **9-12**; 3-dimethylamino-1-propylamine, compound **13**) at 37°C with agitation for 16 h. Products were purified by preparatory reverse-phase high-performance liquid chromatography (HPLC) on a Beckman Gold system using either a Waters Delta-Pak 25  $\times$  100 mm, 15  $\mu$ m 300 Å C<sub>18</sub> PrepPak Cartridge reverse-phase column or a Varian Dynamax 21.4  $\times$  250 mm Microsorb 8  $\mu$ m 300 Å C<sub>8</sub> reverse-phase column in 0.1% (w/v) TFA with acetonitrile as the eluent. The appropriate fractions were lyophilized after characterization by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF-MS) and/or electrospray ionization mass spectrometry (ESI-MS). HPLC analysis was performed on a Beckman Gold system equipped with a diode-array detector using a Phenomenex Gemini 4.6 mm  $\times$  250 mm, 5  $\mu$ m 110 Å C<sub>18</sub> reverse-phase column. UV spectra were measured on an Agilent Technologies 8453 UV-vis ChemStation spectrophotometer. MALDI-TOF-MS was carried out on an Applied Biosystems Voyager DE-PRO.

Conjugates were formed by reacting the isothiocyanate (compounds **1**, **2**, **4**, **7**, **9**, and **11**) or succinimidyl ester (compounds **3** and **10**) with the polyamide in a solution of DIEA (20 equiv) and DMF for 1 h at room temperature. Conjugates **5**, **8**, and **12** were formed by preactivating isophthalic acid (3.0 equiv) with PyBOP (2.9 equiv, Novabiochem) in a solution of DIEA (20 equiv) and DMF at 37°C for 30 min, followed by reaction of the activated solution with the polyamide for 1 h at room temperature. Conjugates were deprotected with neat TFA (Halocarbon) and triethylsilane for 30 min at room temperature before purification by preparatory reverse-phase HPLC. Lyophilization of the appropriate fractions yielded the polyamide conjugates **1-13**, which were characterized as described above. Extinction coefficients were calculated according to standard protocols.<sup>13</sup> Chemicals not otherwise specified were from Aldrich.

**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -C<sub>3</sub>-FITC (2).** Cleaved from  $\beta$ -Ala-PAM resin with neat 1,3-diaminopropane and conjugated with fluorescein-5-isothiocyanate (FITC, Invitrogen). UV-vis  $\lambda_{\text{max}} = 315, 443 \text{ nm}$ ; ESI-MS  $m/z$  1633.6 ( $\text{C}_{76}\text{H}_{74}\text{ClN}_{22}\text{O}_{15}\text{S}_2^-$  calculated [M-H]<sup>-</sup> 1633.48)

**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -C<sub>3</sub>-FAM (3).** Cleaved from  $\beta$ -Ala-PAM resin with neat 1,3-diaminopropane and conjugated with 5-carboxyfluorescein, succinimidyl ester (5-FAM, SE, Invitrogen). UV-vis  $\lambda_{\text{max}} = 313, 446 \text{ nm}$ ; MALDI-TOF-MS  $m/z$  1605.40 ( $\text{C}_{76}\text{H}_{75}\text{ClN}_{21}\text{O}_{16}\text{S}^+$  calculated [M+H]<sup>+</sup> 1604.51)

**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -C<sub>3</sub>-CPITC (4).** Cleaved from  $\beta$ -Ala-PAM resin with neat 1,3-diaminopropane and conjugated with 3-carboxyphenyl isothiocyanate (Trans World Chemicals). UV-vis  $\lambda_{\text{max}} = 309 \text{ nm}$ ; ESI-MS  $m/z$  1423.4 ( $\text{C}_{63}\text{H}_{68}\text{ClN}_{22}\text{O}_{12}\text{S}_2^-$  calculated [M-H]<sup>-</sup> 1423.45)

**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -C<sub>3</sub>-IPA (5).** Cleaved from  $\beta$ -Ala-PAM resin with neat 1,3-diaminopropane and conjugated with isophthalic acid. UV-vis  $\lambda_{\text{max}} = 314 \text{ nm}$ ; MALDI-TOF-MS  $m/z$  1394.72 ( $\text{C}_{63}\text{H}_{69}\text{ClN}_{21}\text{O}_{13}\text{S}^+$  calculated [M+H]<sup>+</sup> 1394.48)



**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -C<sub>3</sub>-NH<sub>2</sub> (6).** Cleaved from  $\beta$ -Ala-PAM resin with neat 1,3-diaminopropane. UV-vis  $\lambda_{\text{max}} = 315$  nm; MALDI-TOF-MS  $m/z$  1246.68 (C<sub>55</sub>H<sub>65</sub>ClN<sub>21</sub>O<sub>10</sub>S<sup>+</sup> calculated [M+H]<sup>+</sup> 1246.46)

**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -(+)-FITC (7).** Cleaved from  $\beta$ -Ala-PAM resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with fluorescein-5-isothiocyanate (FITC, Invitrogen). UV-vis  $\lambda_{\text{max}} = 313, 442$  nm; ESI-MS  $m/z$  1704.3 (C<sub>80</sub>H<sub>83</sub>ClN<sub>23</sub>O<sub>15</sub>S<sub>2</sub><sup>-</sup> calculated [M-H]<sup>-</sup> 1704.56)

**CtPyPyIm-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy- $\beta$ -(+)-IPA (8).** Cleaved from  $\beta$ -Ala-PAM resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with isophthalic acid. UV-vis  $\lambda_{\text{max}} = 316$  nm; MALDI-TOF-MS  $m/z$  1465.59 (C<sub>67</sub>H<sub>78</sub>ClN<sub>22</sub>O<sub>13</sub>S<sup>+</sup> calculated [M+H]<sup>+</sup> 1465.55)

The synthesis and characterization of polyamide conjugate **1** has been reported previously.<sup>8</sup> The synthesis and characterization of polyamide conjugates **9-13** are reported in Chapter 4A.

#### 4.6.2. Quantitative DNase I footprint titrations

pGL2-VEGF-*Luc* was 5'-end-labeled and amplified as previously described.<sup>8,9</sup> PCR products (5'-end-labeled, 197 bp for pGL2-VEGF-*Luc*) were isolated according to standard protocols.<sup>13</sup> Quantitative DNase I footprint titration experiments were performed on the 5'-<sup>32</sup>P-end-labeled PCR products of plasmid pGL2-VEGF-*Luc* with polyamides **2-8** according to standard protocols.<sup>13</sup> Radiolabeled DNA was equilibrated with polyamide solutions for 14-16 h at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0 prior to DNase I cleavage. Chemical sequencing reactions were performed according to published methods.<sup>14,15</sup> Storage phosphor autoradiography was performed on a Molecular Dynamics Typhoon 8600 phosphorimager. 18 M $\Omega$  water was obtained from an AquaMAX Ultra water purification system, and all buffers were 0.2  $\mu$ m filtered.

#### **4.6.3. Cell cultures**

The human cervical cancer cell line HeLa was cultured in a 5% CO<sub>2</sub> atmosphere at 37°C in supplemented DMEM medium (GIBCO).<sup>5,6</sup> DMEM medium was supplemented with 10% fetal bovine serum (Omega Scientific) and 1% penicillin/streptomycin solution (Mediatech).

#### **4.6.4. Confocal microscopy**

Confocal microscopy experiments were performed according to published protocols.<sup>5,6</sup> HeLa cells were trypsinized (Mediatech) for 5 min at 37°C, centrifuged for 10 min at 4°C at 100 × g, and resuspended in fresh medium to a concentration of  $1.33 \times 10^5$  cells/mL.<sup>5,6</sup> Incubations were performed by adding 150 µL of cells into culture dishes equipped with glass bottoms for direct imaging (MatTek). Cells were grown in the glass-bottom culture dishes for 24 h. The medium was then removed and replaced with 147 µL of fresh medium, followed by addition of 3 µL of the 100 µM solution for a final polyamide concentration of 2 µM. Cells were incubated in a 5% CO<sub>2</sub> atmosphere at 37°C for 12-14 h. Imaging was performed with a 40× oil-immersion objective lens on a Zeiss LSM 5 Pascal inverted laser scanning microscope. Polyamide-fluorescein conjugate fluorescence and visible light images were obtained using standard filter sets for fluorescein.<sup>5,6</sup>

#### **4.6.5. Quantitative RT-PCR experiments**

Quantitative RT-PCR experiments were performed using HeLa cells according to published protocols.<sup>8,9</sup> For cell culture experiments, HeLa cells were plated in a 24-well format with a concentration of  $3 \times 10^4$  cells/mL in a volume of 500 µL. Polyamides were added to the medium at this time. Hypoxia was induced with 300 µM desferrioxamine mesylate (DFO) after 48 h. RNA was harvested after 16 h of chemical induction using an RNeasy Mini Kit (Qiagen) and reverse transcribed using PowerScript Reverse Transcriptase

(Clontech). Quantitative RT-PCR was performed using SYBR GREEN PCR Master Mix (Applied Biosystems) on an Applied Biosystems 7300 Real Time PCR System. VEGF gene expression levels were measured relative to GUSB.

## 4.7. Results and discussion

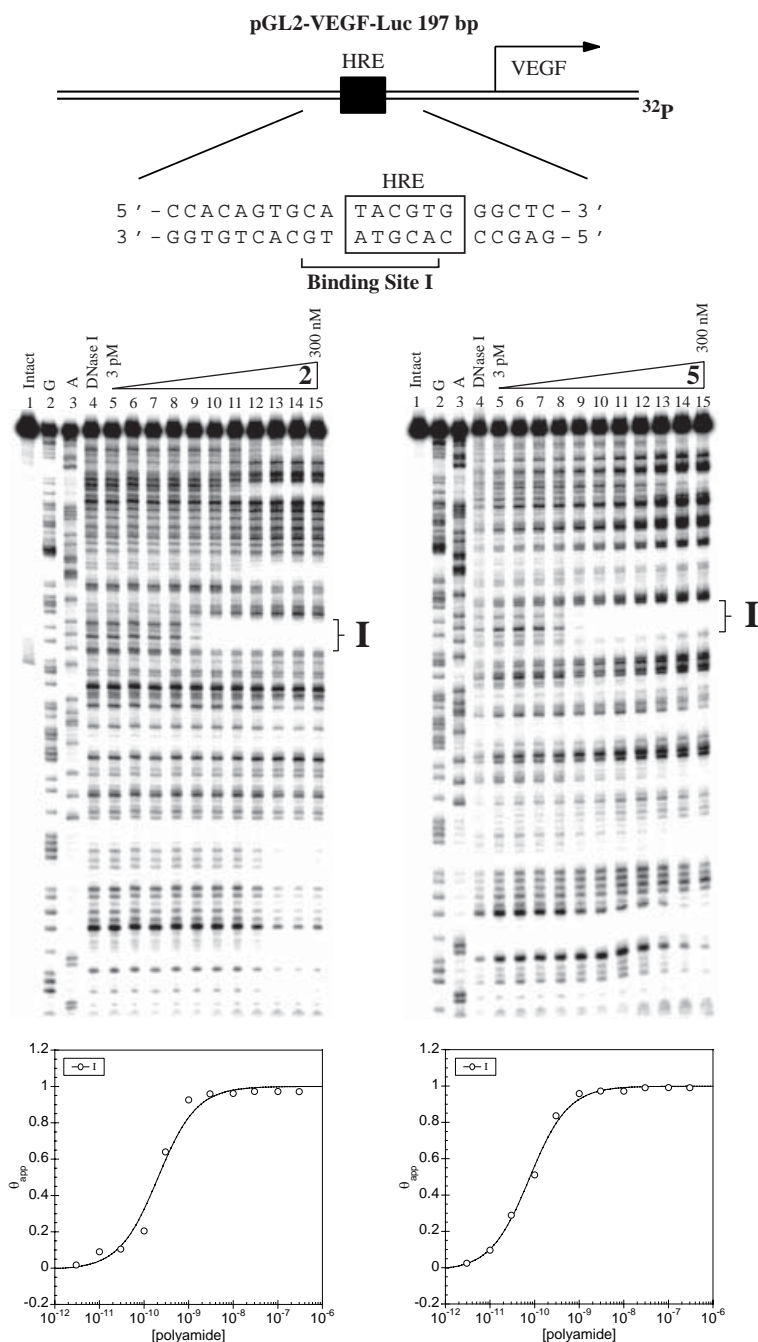
### 4.7.1. Polyamide design and synthesis

The hairpin polyamide conjugates shown in Figure 4.5 comprise two polyamide cores for match and mismatch experiments. The synthesis and activity of polyamide conjugate **1** have been previously characterized.<sup>8-10</sup> Polyamides **2-8** were designed to target the hypoxia response element 5'-TACGTG-3' binding site. Polyamides **2-6** were synthesized with Boc- $\beta$ -Ala-PAM resin and cleaved with 1,3-diaminopropane to give the  $\beta$ -alanine-C<sub>3</sub> linker. Conjugates **2-5** were formed with a fluorescein or isophthalic acid moiety containing a thiourea or amide linkage. An unconjugated  $\beta$ -alanine-C<sub>3</sub> polyamide **6** was included in the series as a control compound. Conjugates **7** and **8** incorporate a fluorescein or isophthalic acid moiety with a positively charged  $\beta$ -alanine-triamine linker.

The synthesis of polyamides **9-13** has been described in Chapter 4A. These compounds bind the sequence 5'-WGGWCW-3' with affinities of  $K_a > 2 \times 10^9 \text{ M}^{-1}$ . The nuclear localization of the fluorescein conjugate **9** is excellent across the HeLa, MCF-7, and PC3 cell lines tested, whereas the uptake of the FAM conjugate **10** is more limited but still nuclear.

### 4.7.2. DNase I footprint titrations

The binding affinities of polyamides **2-8** on the hypoxia response element 5'-TACGTG-3' on plasmid pGL2-VEGF-*Luc* were measured by DNase I footprint titration experiments (Figure 4.6 and Table 4.4). With the exception of polyamide **3**, all of the compounds in the series have equilibrium association constants of  $K_a > 3 \times 10^9 \text{ M}^{-1}$ . As previously observed, the fluorescein conjugates **2** and **3** exhibit decreased binding affinities.



**Figure 4.6.** Quantitative DNase I footprint titration experiments for polyamides **2** and **5** on the 197 bp, 5'-<sup>32</sup>P-end-labeled PCR product of pGL2-VEGF-*Luc*; lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-15, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, 100 nM, 300 nM polyamide, respectively. Binding isotherms are shown below each footprinting gel;  $\theta_{\text{norm}}$  values were calculated according to published methods.<sup>13</sup>

**Table 4.4.** Equilibrium association constants for polyamides **2-8**<sup>a</sup>

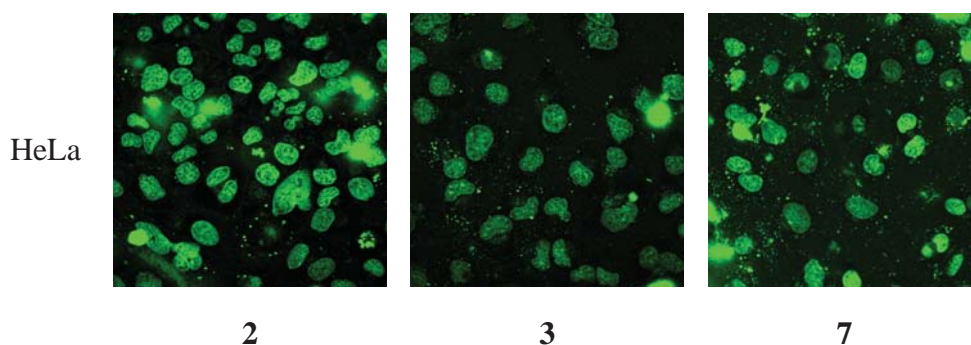
Polyamide	5'-TACGTG-3'
<b>2</b>	$3.8 (\pm 1.2) \times 10^9$
<b>3</b>	$1.0 (\pm 0.3) \times 10^8$
<b>4</b>	$1.2 (\pm 0.2) \times 10^{10}$
<b>5</b>	$1.4 (\pm 0.1) \times 10^{10}$
<b>6</b>	$1.3 (\pm 0.2) \times 10^{11}$
<b>7</b>	$3.6 (\pm 0.9) \times 10^{10}$
<b>8</b>	$2.1 (\pm 0.6) \times 10^{10}$

<sup>a</sup>  $K_a$  ( $M^{-1}$ ) values reported are the mean values from at least three DNase I footprint titration experiments; standard deviations are shown in parentheses. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM  $MgCl_2$ , and 5 mM  $CaCl_2$  at pH 7.0.

In Chapter 4A, the difference between the affinities of the FITC and FAM conjugates was less than two-fold. In this case, FITC conjugate **2** binds the match site with  $K_a = 3.8 \times 10^9 M^{-1}$ , while the FAM conjugate **3** has an affinity of  $1.0 \times 10^8 M^{-1}$ . This 38-fold difference is surprising and may be sequence-dependent.

#### 4.7.3. Confocal microscopy

The cellular uptake of the  $\beta$ -alanine- $C_3$ -linked polyamide-fluorescein conjugates **2**, **3**, and **7** was directly imaged on a confocal laser scanning microscope (Figure 4.7). HeLa cells were incubated with 2  $\mu M$  polyamide for 12-14 h in a 5%  $CO_2$  atmosphere at 37°C prior to imaging. All three polyamides showed nuclear staining that exceeds that of the medium. This result is encouraging in the case of the FAM conjugate **3** because the nuclear localization of the  $\beta$ -alanine- $C_3$ -linked FAM conjugate **10** was modest, as described in Chapter 4A. The positive uptake of the  $\beta$ -alanine-triamine-linked fluorescein conjugate **7** suggests that this polyamide core could be more amenable than others in terms of localization.

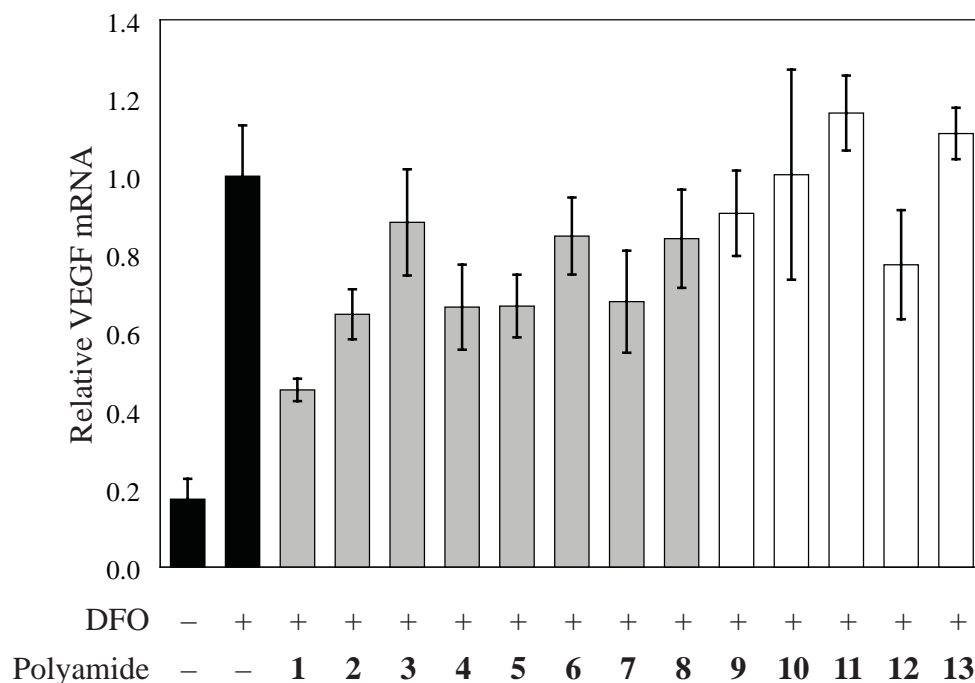


**Figure 4.7.** Nuclear localization of polyamides in HeLa cells after incubation at 2  $\mu$ M concentration for 12-14 h at 37°C in standard culture media.

#### 4.7.4. Quantitative RT-PCR experiments in HeLa cells

The binding of polyamide **1** to the hypoxia response element results in the downregulation of hypoxia-inducible genes, including vascular endothelial growth factor (VEGF).<sup>8-10</sup> HeLa cells were incubated with 0.2  $\mu$ M or 1  $\mu$ M polyamide for 48 h. After the incubation period, hypoxia was induced with 300  $\mu$ M DFO. RNA was harvested 16 h after induction, and quantitative RT-PCR analysis was performed (Figure 4.8 and Figure 4.9).

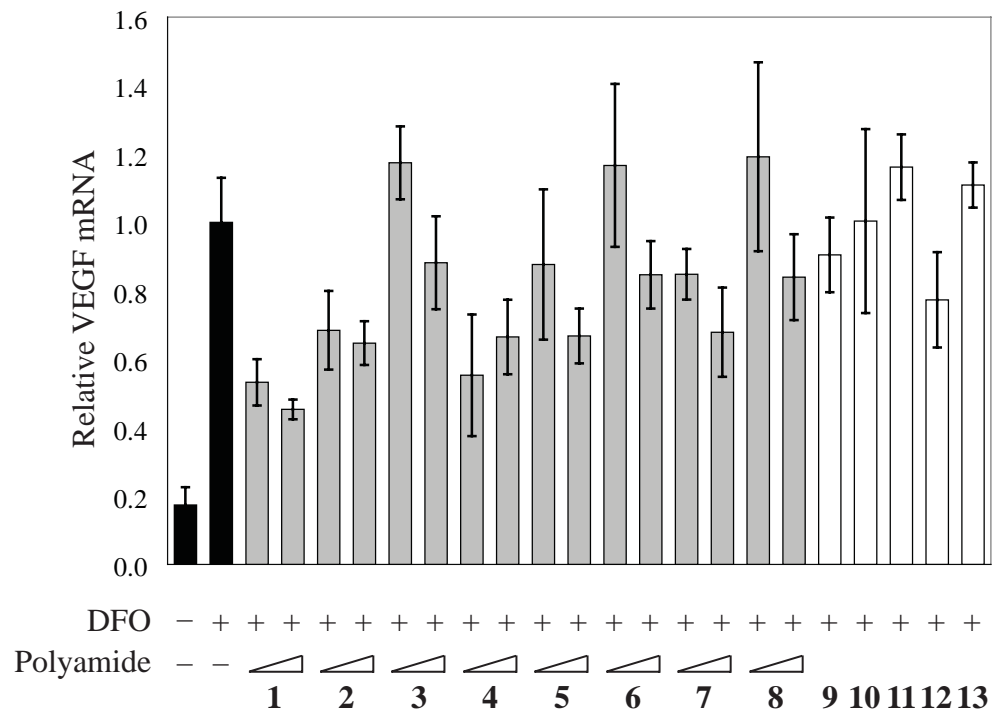
The  $\beta$ -alanine- $C_3$ -linked hairpin polyamide conjugates **2**, **4**, and **5** showed comparable levels of VEGF mRNA that were decreased relative to the untreated control. However, the lead compound **1** remained superior in downregulating the VEGF gene. The FAM conjugate **3**, which had shown good nuclear localization but decreased binding affinity, showed little activity. The high-affinity unconjugated polyamide **6** also displayed little effect, consistent with the model that polyamide conjugates exhibit improved nuclear localization. The  $\beta$ -alanine-triamine-linked fluorescein conjugate **7** showed moderate results, while the analogous isophthalic acid conjugate **8** was less effective. Mismatch polyamides **9-13** did not downregulate VEGF expression at the 1  $\mu$ M concentration.



**Figure 4.8.** Relative levels of VEGF mRNA expression as measured by real-time quantitative RT-PCR. HeLa cells were treated with final concentration of 1  $\mu$ M polyamide, and hypoxia was induced with 300  $\mu$ M DFO.

Taken together, these experiments indicate that  $\beta$ -alanine- $C_3$ -linked hairpin polyamide conjugates targeted to the hypoxia response element downregulate the hypoxia-inducible VEGF gene in a sequence-specific manner. Polyamide conjugates synthesized on Boc- $\beta$ -Ala-PAM resin should remain viable candidates for cell culture studies. Judicious linker selection is important in determining nuclear localization.

**Acknowledgements.** We are grateful to the National Institutes of Health for research support. Nicholas Nickols synthesized polyamide **1**.



**Figure 4.9.** Relative levels of VEGF mRNA expression as measured by real-time quantitative RT-PCR. HeLa cells were treated with final concentration of 0.2 or 1  $\mu$ M polyamide, and hypoxia was induced with 300  $\mu$ M DFO.



## References

1. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.*, **9**, 2215-2235.
2. 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.
3. 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.
4. 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. *Bioorg. Med. Chem. Lett.*, **13**, 1565-1570.
5. 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.
6. 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.

7. Stafford, R.L. (2008) Ph.D. Thesis, California Institute of Technology, Pasadena, CA.
8. 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.
9. 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.
10. Nickols, N.G., Jacobs, C.S., Farkas, M.E. and Dervan, P.B. (2007) Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.*, **2**, 561-571.
11. 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.
12. 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.

13. Trauger, J.W. and Dervan, P.B. (2001) Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. *Methods in Enzymology*, **340**, 450-466.
14. Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods in Enzymology*, **65**, 499-560.
15. Iverson, B.L. and Dervan, P.B. (1987) Adenine Specific DNA Chemical Sequencing Reaction. *Nucleic Acids Res.*, **15**, 7823-7830.

## Chapter 5

### **Disrupting the Oct4 Octamer DNA Element with Pyrrole-Imidazole Polyamide Conjugates**

This project was done in collaboration with John Phillips, Daniel Harki, and James Puckett (Dervan group, Caltech).

John Phillips performed quantitative DNase I footprint titration experiments. Daniel Harki synthesized polyamides **6**, **7**, and **9-14**. James Puckett synthesized polyamides **7** and **8**.

## Abstract

Oct4 is a member of the POU family of transcription factors that is expressed in the early stages of the developing embryo and whose activity is exquisitely important in the determination of cell fate. In embryonic stem cells, small changes in Oct4 levels mark the boundary between remaining in an undifferentiated state and differentiating into one of the cell layers. The Oct4 octamer binding site, located adjacent to the Sox2 heptamer binding site, has been mapped in the enhancer regions of several developmental genes, including Sox2, Nanog, and Pou5f1, the gene that encodes Oct4. Pyrrole-imidazole polyamides are synthetic ligands that bind the DNA minor groove at subnanomolar concentrations with programmable sequence specificity. A series of hairpin polyamide conjugates was designed to target the Oct4 octamer DNA element, and quantitative RT-PCR was used to assay the mRNA expression levels of Oct4 target genes. In the P19 mouse embryonal carcinoma cell line, polyamide treatment resulted in a decrease in Utf1 mRNA that was comparable to that of Oct4 RNA interference-treated cells. In addition, the Sox2 mRNA level increased two-fold when P19 cells were treated with Oct4-targeted polyamide, suggesting that de-repression had occurred. When the R1 mouse embryonic stem cell line was incubated with polyamide, a decrease in the Utf1 mRNA level and a two-fold increase in the Nanog mRNA level were measured by quantitative RT-PCR assay. These results indicate that polyamide conjugates modulate the expression levels of Oct4-driven genes that control the determination of cell fate.

## 5.1. Introduction

Embryonic stem (ES) cells are derived from the inner cell mass of the blastocyst, one of the earliest stages of embryonic development.<sup>1</sup> They are characterized by the capacity for symmetrical self-renewal and the potential to differentiate into all three germ layers: ectoderm, endoderm, and mesoderm.<sup>1,2</sup> Oct4, Sox2, and Nanog are transcription factors that contribute to the pluripotency of ES cells.<sup>3</sup> In particular, Oct4 plays a central role in ES self-renewal.<sup>4-6</sup> Control of Oct4 expression within a tight range is essential to maintain pluripotency.<sup>7</sup> A less than two-fold increase results in differentiation into primitive endoderm and mesoderm, while a 50% decrease causes ES cells to revert to trophectoderm.<sup>7</sup> The Oct4/Sox2 complex is an important target in regenerative medicine as researchers look to reprogram somatic cells by defined factors.<sup>8-12</sup>

Oct4 (encoded by the Pou5f1 gene) belongs to the POU (Pit, Oct, UNC) family of transcription factors and binds to an octamer DNA sequence 5'-ATGCAAAT-3'.<sup>13</sup> These DNA-binding proteins contain two structurally independent domains connected by a variable linker of 15 to 56 amino acids in length.<sup>13</sup> The POU-specific domain is roughly 75 amino acids in length and binds the 5'-end ATGC, while the POU-homeo domain contains approximately 60 amino acids and binds the 3'-end AAAT.<sup>13</sup> Oct4 is highly expressed in embryonic stem (ES) cells and is downregulated when these cells differentiate.<sup>14</sup>

Oct4 octamer binding sites are often found in close proximity to Sox2 heptamer elements.<sup>14</sup> Oct4 and Sox2 can form a ternary complex with DNA (Figure 5.1).<sup>15</sup> The two proteins can also bind as a heterodimer in the absence of DNA.<sup>14</sup> The enhancers of target genes have been mapped in mice, and these regulatory regions contain an Oct4 binding site located near Sox2 binding elements. These genes include Pou5f1 (the gene that encodes Oct4), Sox2, Nanog, Utf1, Fbx15, and Fgf4 (Figure 5.1).<sup>14</sup> These DNA elements are generally found adjacent to one another, but can be separated by 3 bp, as in the Fgf4 enhancer.



	Sox2	Oct4
i) Pou5f1	CTTTGTT---	ATGCATCT
ii) Sox2	CATTGTG---	ATGCATAT
iii) Nanog	CATTGTA---	ATGCAAAA
iv) Utf1	CATTGTT---	ATGCTAGT
v) Fbx15	CATTGTT---	ATGATAAA
vi) Fgf4	CTTTGTTTGG	ATGCTAAT

**Figure 5.1.** (top) The 2.6 Å X-ray crystal structure of a POU/HMG/DNA ternary complex (PDB accession code 1GT0).<sup>15</sup> The POU domain of Oct1 is shown in green, and the HMG domain of Sox2 is shown in blue. These two transcription factors interact on the FGF4 enhancer. (bottom) The alignment of Sox and Oct elements from mouse target genes known to bind Sox2 and Oct4.<sup>14</sup> Image rendered by John Phillips using Chimera software.

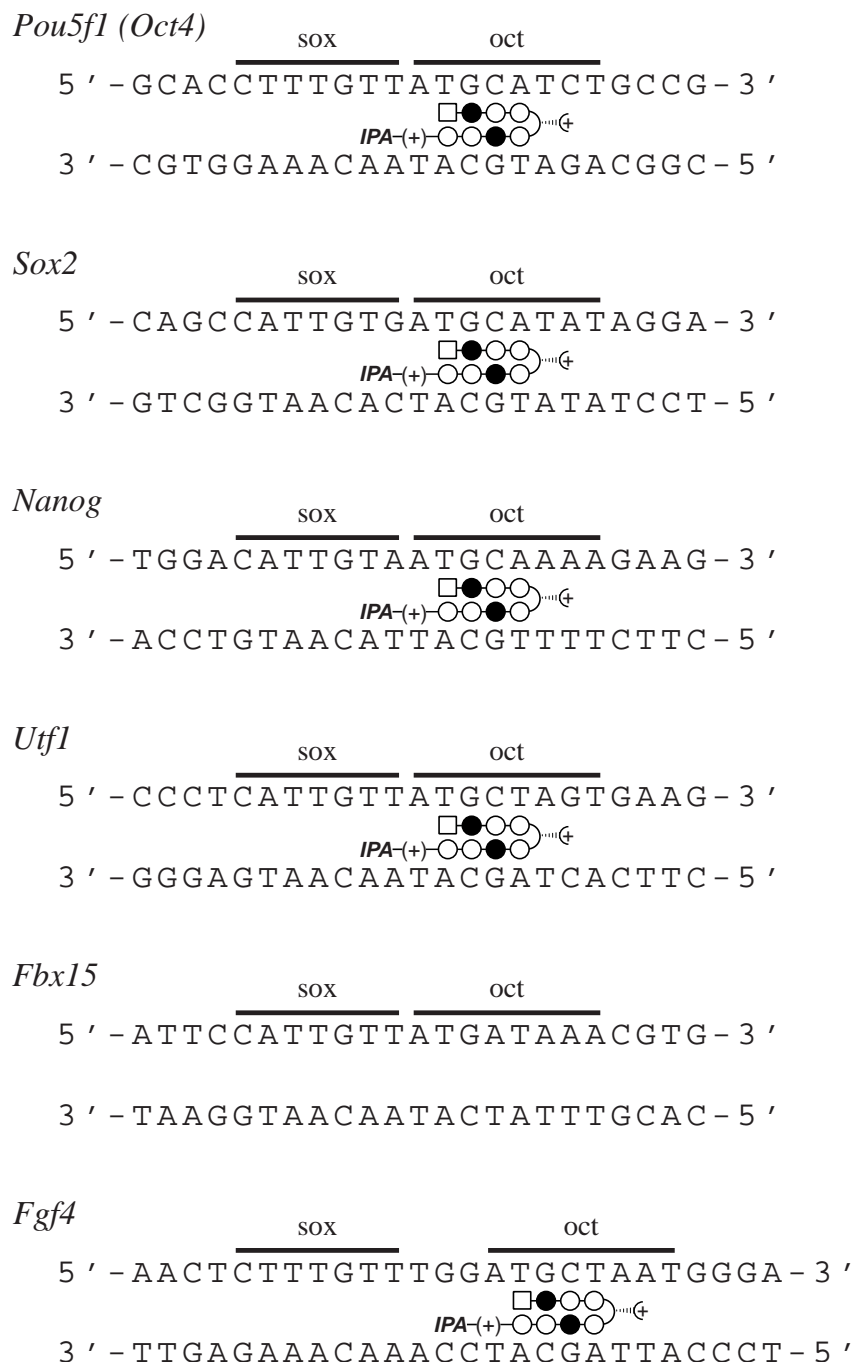
The regulatory elements that control the expression of the *Pou5f1* gene include a proximal enhancer located about 1.2 kb upstream of the transcription start site and a distal enhancer located about 2 kb upstream.<sup>14</sup> The distal enhancer controls ES cell-specific expression of *Pou5f1* and contains a composite oct-sox element such that Oct4 regulates its own expression.<sup>14</sup> The *Sox2* regulatory region also contains two enhancers.<sup>16</sup> It has been proposed that the essential function of *Sox2* is to stabilize ES cells in a pluripotent state by maintaining the requisite level of Oct4 expression.<sup>17</sup> It has been shown that increasing the level of *Sox2* inhibits the endogenous expression of several Oct4-*Sox2* target genes.<sup>18</sup>

*Nanog* is an Oct4 target gene that works together with Oct4 and *Sox2* to control ES cell pluripotency.<sup>19</sup> The binding sites of Oct4 and *Nanog* have been globally mapped in the mouse ES cell genome by the ChIP-PET method and were found to exhibit substantial overlap.<sup>20</sup> The oct-sox composite element in the *Nanog* proximal promoter is centered 175 bp upstream of the transcription start site.<sup>3</sup> Downregulation of Oct4 and *Nanog* by RNA interference was shown to induce differentiation in mouse ES cells.

The enhancers of the *Utf1*, *Fbx15*, and *Fgf4* genes also contain Oct4 octamer binding elements. The *Utf1* regulatory element selectively binds the Oct4-*Sox2* complex but not those containing other members of the Oct family.<sup>21</sup> The *Fbx15* enhancer octamer binding element is centered 530 bp upstream of the transcription start site and contains a noncanonical 5'-ATGATAAA-3' sequence.<sup>22</sup> In the *Fgf4* enhancer, the Oct4 and *Sox2* binding sites are separated by 3 bp.<sup>23</sup> The crystal structure in Figure 5.1 contains the Oct4-*Sox2* complex bound on the *Fgf4* enhancer.<sup>15</sup>

Pyrrole-imidazole polyamides bind the minor groove of DNA in a sequence-specific manner at subnanomolar concentrations.<sup>24,25</sup> The empirically derived “pairing rules” allow for shape-selective recognition of DNA sequences.<sup>24,25</sup> Polyamides designed to target 5'-WTGCWW-3' are shown with their putative binding sites in the enhancers of Oct4 target genes (Figure 5.2). For the *Pou5f1*, *Sox2*, *Nanog*, and *Utf1* genes, the Oct4 and *Sox2* binding sites are found directly adjacent to one another, but these elements are separated by





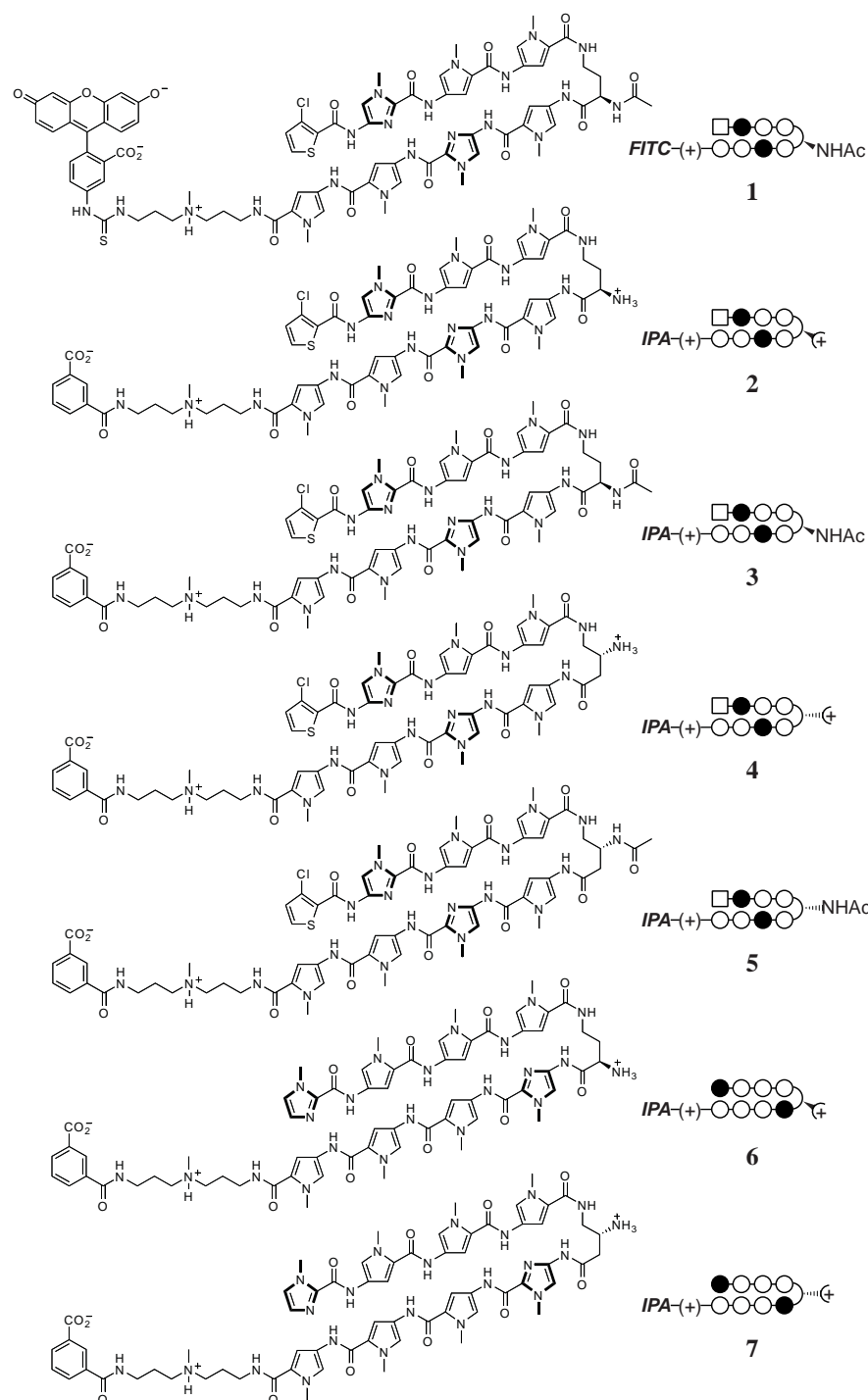
**Figure 5.2.** Ball-and-stick binding model for the hairpin motif with the polyamide bound to its target DNA sequence within the enhancer regions of Oct4 target genes, including *Pou5f1* (the gene that encodes Oct4), *Sox2*, *Nanog*, *Utf1*, *Fbx15*, and *Fgf4*. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as "(+)"; and the chiral 3,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a dashed wedge linking a half-circle with a plus. The *Sox2* heptamer element and the Oct4 octamer element are indicated as "sox" and "oct" above the DNA binding site.

3 bp in the Fgf4 enhancer. The Fbx15 element differs from the 5'-ATGCAAAT-3' octamer sequence by one base pair and would be a mismatch for the designed polyamide.

## 5.2. Results and discussion

### 5.2.1. Polyamide design and synthesis

Hairpin pyrrole-imidazole polyamide conjugates are synthetic ligands that bind DNA in a sequence-specific manner with subnanomolar affinity.<sup>24,25</sup> Polyamides **1-5** were designed to bind the DNA sequence 5'-WTGCWW-3', and their structures are indicated in Figure 5.3. Conjugate **1** contains fluorescein connected to the polyamide with a thiourea linkage to allow direct imaging of cellular uptake by confocal laser scanning microscopy. Conjugates **2-5** contain an isophthalic acid moiety connected to the polyamide with an amide linkage. This class of polyamide conjugates has been shown to exhibit improved nuclear localization.<sup>26</sup> Polyamides **2** and **3** contain a chiral 2,4-diaminobutyric acid turn linkage that improves the binding affinity over the achiral variant without loss of sequence specificity.<sup>27</sup> Polyamides **4** and **5** contain a chiral 3,4-diaminobutyric acid turn linkage that has been shown to increase the DNA binding affinity further with some, but not all, polyamide cores.<sup>28</sup> The acetylated turn linkage used in the design of compounds **1**, **3**, and **5** has been shown to improve nuclear localization in some human cancer cell lines.<sup>29,30</sup> Polyamides **1-5** were synthesized on oxime resin according to published manual solid-phase synthesis protocols.<sup>31,32</sup> Fluorescein and isophthalic acid conjugations, as well as the subsequent deprotection and acetylation steps, were performed according to published methods.<sup>26,28-30</sup> Polyamides **6** and **7** were synthesized as mismatch control compounds for cell culture experiments and have been previously characterized.<sup>28,33</sup>



**Figure 5.3.** Structures of polyamide conjugates 1-7. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as "(+)"; the chiral 2,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a solid wedge linking a half-circle with a plus; and the chiral 3,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a dashed wedge linking a half-circle with a plus.

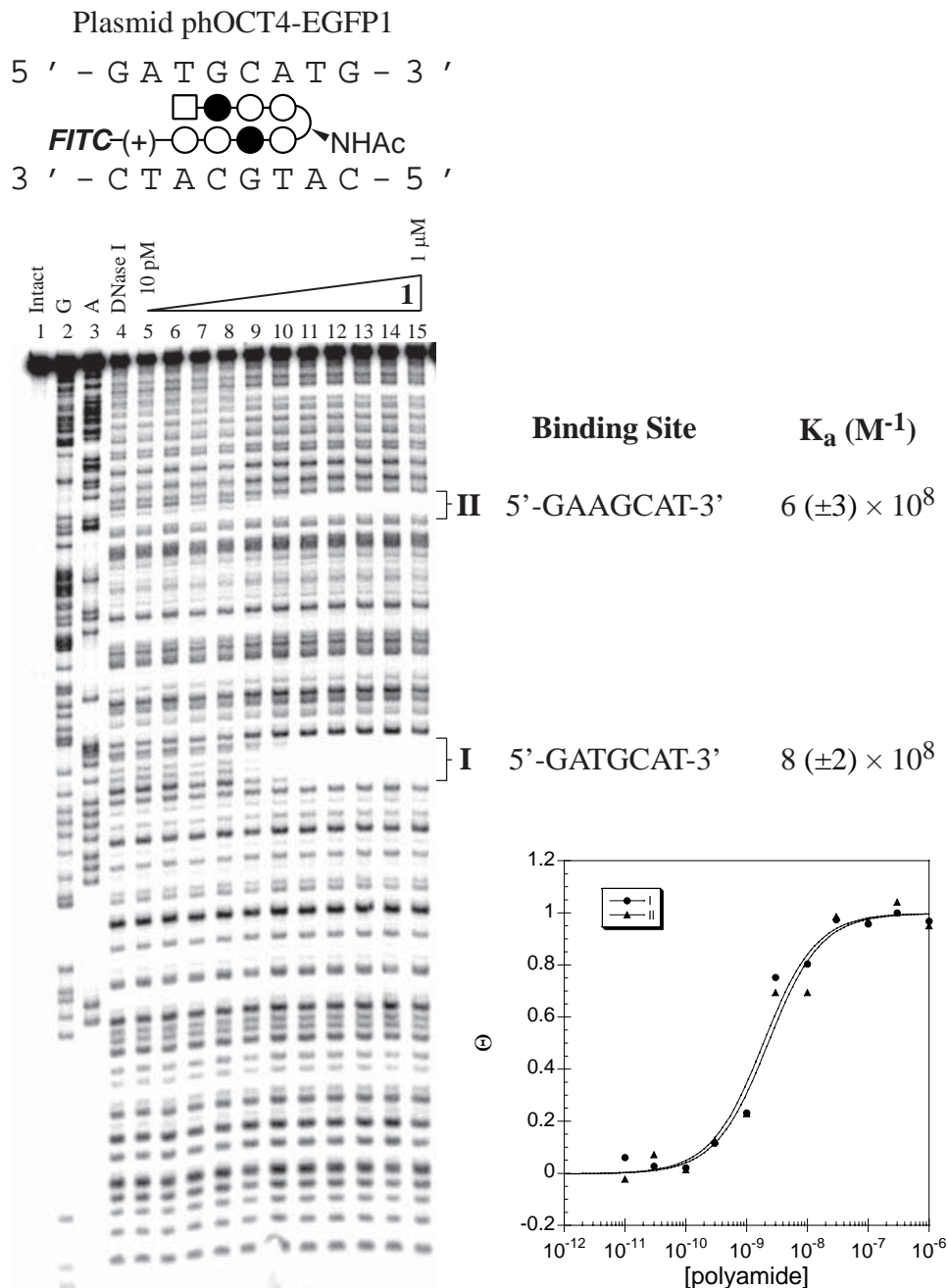
### 5.2.2. DNase I footprint titrations and melting temperature analysis

The equilibrium association constant ( $K_a$ ) of polyamide **1** was determined by DNase I footprint titration experiments on the Oct4 binding site of the phOCT4-EGFP1 plasmid obtained from Dr. Wei Cui of the Roslin Institute.<sup>34</sup> A partial sequence of this plasmid is shown in Figure 5.18. Polyamide **1** exhibited a binding affinity of  $K_a = 8 \times 10^8 \text{ M}^{-1}$  for the DNA sequence 5'-ATGCAT-3' and  $K_a = 6 \times 10^8 \text{ M}^{-1}$  for 5'-ATGCTT-3' (Figure 5.4). These experiments indicate that the acetylated polyamide-fluorescein conjugate **1** is a high-affinity DNA binder that exhibits good selectivity, as no other mismatch binding sites were evident in the footprinting gels. As shown in the footprinting data presented in Chapters 4A and 4B, fluorescein conjugation decreases the equilibrium association constant by an order of magnitude.

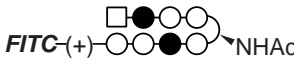
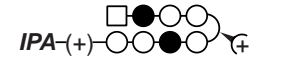
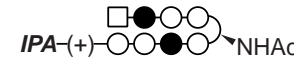
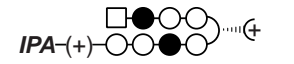
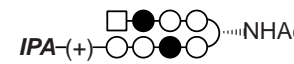
Melting temperature analysis by UV absorption spectrometry was used to assay the DNA binding affinity of polyamides **1-5**. The 12 bp DNA sequence containing the match site observed in the DNase I footprint titrations was used for this experiment (Figure 5.5). The  $T_m$  value for 5'-GAGATGCATGAC-3' and its complement sequence alone was 53.5°C. Polyamide **1** raised the  $T_m$  value by 6.0°C, and the reduced steric bulk of the tail moiety in polyamide **3** increased the  $T_m$  value by 11.4°C. As expected, the free amine compound **2** exhibited a higher  $\Delta T_m$  value of 15.5°C. Polyamides containing the 3,4-diaminobutyric acid turn subunit had the highest  $\Delta T_m$  values. Perhaps surprisingly, acetylated polyamide **5** exhibited a  $\Delta T_m$  value of 16.7°C that was higher than that of the free amine compound **2**. Polyamide **4** raised the melting temperature by 19.1°C, indicating that its DNA binding affinity was the largest in the series.

### 5.2.3. Cell culture and confocal microscopy

P19 mouse embryonal carcinoma cells were selected as a model system to measure the expression of Oct4 target genes in cell culture. This cell line was obtained from ATCC and cultured at 37°C in a 5% CO<sub>2</sub> atmosphere according to recommended protocols.

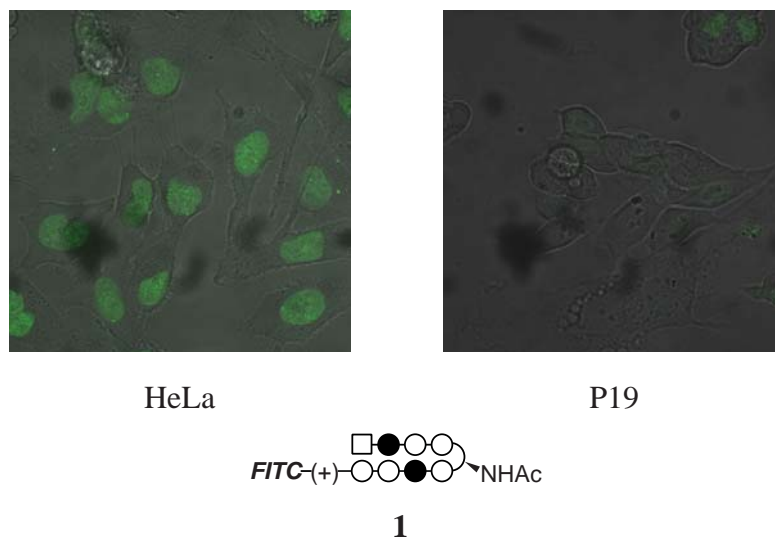


**Figure 5.4.** Quantitative DNase I footprint titration experiment for polyamide **1** on the 5'-end-labeled PCR product of plasmid phOCT4-EGFP1: lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-15, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, 100 nM, 300 nM, 1  $\mu$ M polyamide, respectively. Binding isotherms are shown, and  $\theta_{\text{norm}}$  values were calculated according to published methods.<sup>37</sup> Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square; the 3,3'-diamino-N-methyldipropylamine linker is shown as "(+)"; and the chiral 2,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a solid wedge linking a half-circle with a plus.

$  \begin{array}{rcl}  12 \text{ bp DNA} & 5' - \text{GAG} \boxed{\text{ATGCA T}} \text{GAC} - 3' \\  & 3' - \text{CTC} \boxed{\text{TACGTA}} \text{CTG} - 5'  \end{array}  $			
	Polyamide	$T_m$ (°C)	$\Delta T_m$ (°C)
	DNA only	53.5 (0.3)	
1		59.5 (0.4)	6.0
2		69.0 (0.2)	15.5
3		64.9 (0.2)	11.4
4		72.6 (0.4)	19.1
5		70.2 (0.3)	16.7

**Figure 5.5.** DNA melting temperature analysis. Experiments were performed with a final concentration of 2  $\mu\text{M}$  of DNA oligonucleotides and polyamides **1-5** in a buffer of 10 mM sodium cacodylate, 10 mM KCl, 10 mM  $\text{MgCl}_2$ , and 5 mM  $\text{CaCl}_2$  at pH 7.0. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; the chiral 2,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a solid wedge linking a half-circle with a plus; and the chiral 3,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a dashed wedge linking a half-circle with a plus. Error values in parentheses indicate standard deviations of quadruplicate experiments.

P19 cells were incubated for 12 h with polyamide **1** to determine the extent of nuclear localization. Direct imaging by confocal laser scanning microscopy was used to image cellular uptake. Polyamide **1** was observed to localize to the nucleus of P19 cells (Figure 5.6). However, the uptake profile in HeLa cells was better than that in P19 cells, suggesting that high polyamide concentrations would be necessary to see changes in gene expression. In Chapter 3, it was observed that a polyamide-fluorescein conjugate showed higher nuclear concentration in U251 cells than in HeLa cells.

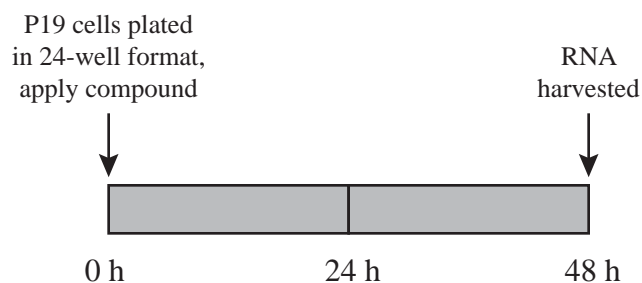


**Figure 5.6.** Confocal laser scanning microscopy images of polyamide-fluorescein conjugate **1** in HeLa and P19 cell lines. Cells were incubated with 2  $\mu\text{M}$  polyamide for 12 h at 37°C in a 5%  $\text{CO}_2$  atmosphere. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as "(+)"; and the chiral 2,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a solid wedge linking a half-circle with a plus.

#### 5.2.4. Quantitative RT-PCR experiments in P19 cells

For cell culture experiments, P19 mouse embryonal carcinoma cells were plated in a 24-well format and incubated with polyamides for 48 h. The time course for this protocol is illustrated in Figure 5.7. At the conclusion of the experiment, RNA was harvested, and quantitative RT-PCR was performed to measure the gene expression levels of Oct4 target genes relative to the housekeeping gene Gapdh. Polyamides **2** and **6** were assayed at 10  $\mu\text{M}$  and 20  $\mu\text{M}$ . Lower concentrations (2.5  $\mu\text{M}$  and 5  $\mu\text{M}$ ) did not change gene expression levels. Polyamide **6** was previously shown to bind the prostate-specific antigen (PSA) androgen response element and downregulate androgen-induced PSA genes in cultured prostate cancer cells.<sup>33</sup> Stealth RNAi duplexes from Invitrogen designed to silence the Pou5f1 gene, which encodes Oct4, were used as a positive control. All three duplexes were effective in reducing Pou5f1 gene expression levels greater than 80%, and duplex

### Time course for 48-hour experiment (P19 mouse embryonal carcinoma cells)



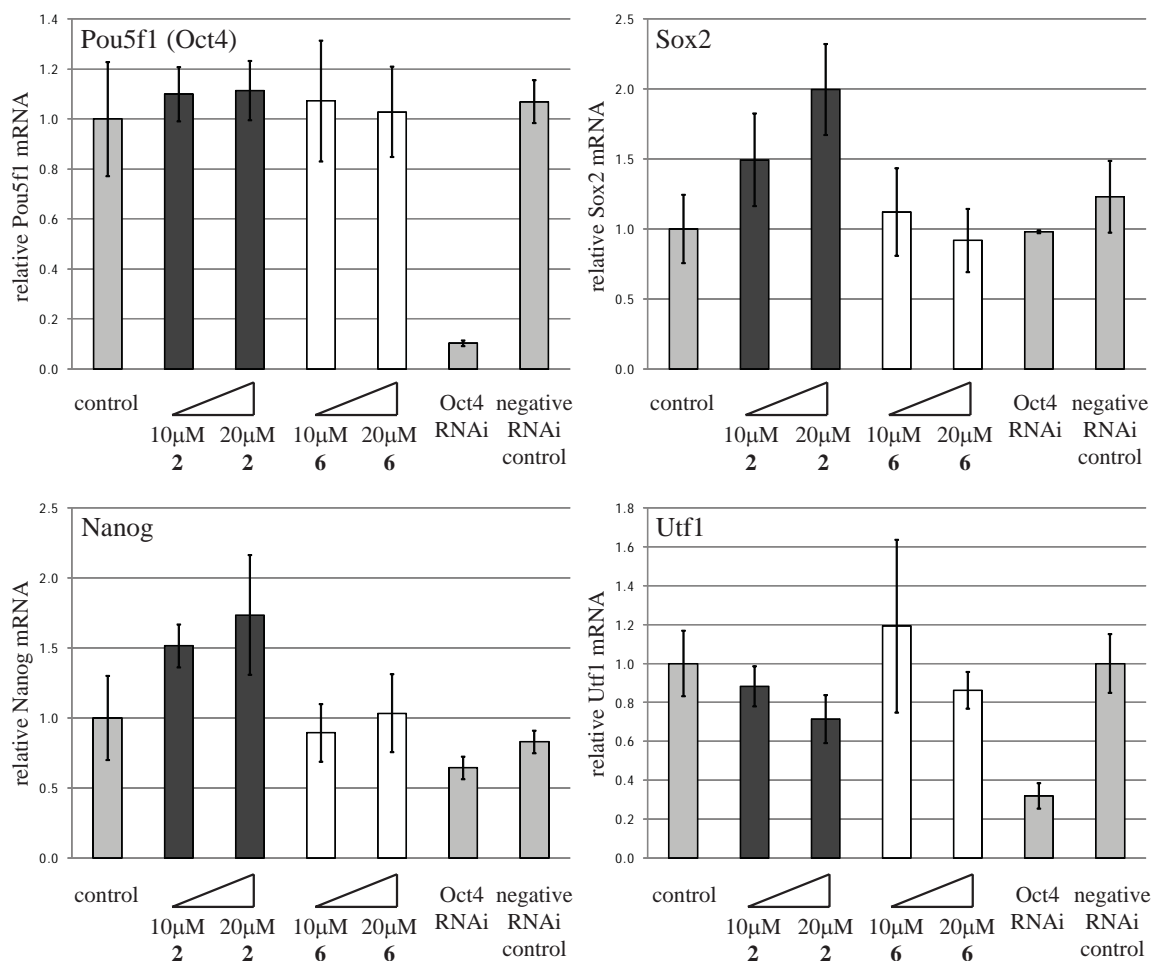
**Figure 5.7.** Time course for 48-hour incubation experiments with P19 mouse embryonal carcinoma cells. P19 cells were plated in 24-well format at  $t = 0$  h, and compounds were applied at this time. RNA was harvested at the conclusion of the experiment ( $t = 48$  h).

#2 achieved the greatest knockdown effect. A mismatch Stealth RNAi with medium GC content was used as a negative control.

Oct4 mRNA levels were unchanged by polyamides **2** and **6**, while the RNAi decreased mRNA levels by 90%, as shown in Figure 5.8. Utf1 mRNA levels decreased in the presence of match polyamide **2**, as would be expected by the disruption of Oct4 transcription factor binding at the octamer DNA element in the enhancer region of the Utf1 gene. Sox2 and Nanog mRNA levels increased in a dose-dependent manner in the presence of polyamide **2**. All four target genes assayed displayed a sequence-specific effect, as mismatch polyamide **6** at 10  $\mu$ M and 20  $\mu$ M concentrations did not change expression levels.

Polyamides **2-5** were assayed at 20  $\mu$ M to determine whether compounds with increased  $\Delta T_m$  values and acetylated turn linkages would show larger changes in gene expression levels (Figure 5.9). Gene expression was largely unchanged in the presence of acetylated compounds **3** and **5**, suggesting that this class of compounds does not exhibit good cellular uptake in P19 cells. On the other hand, Utf1 mRNA levels decreased in the presence of 20  $\mu$ M of polyamides **2** and **4**, which contain a free chiral amine on the turn subunit. Sox2 mRNA increased three-fold when P19 cells were incubated with 20  $\mu$ M of

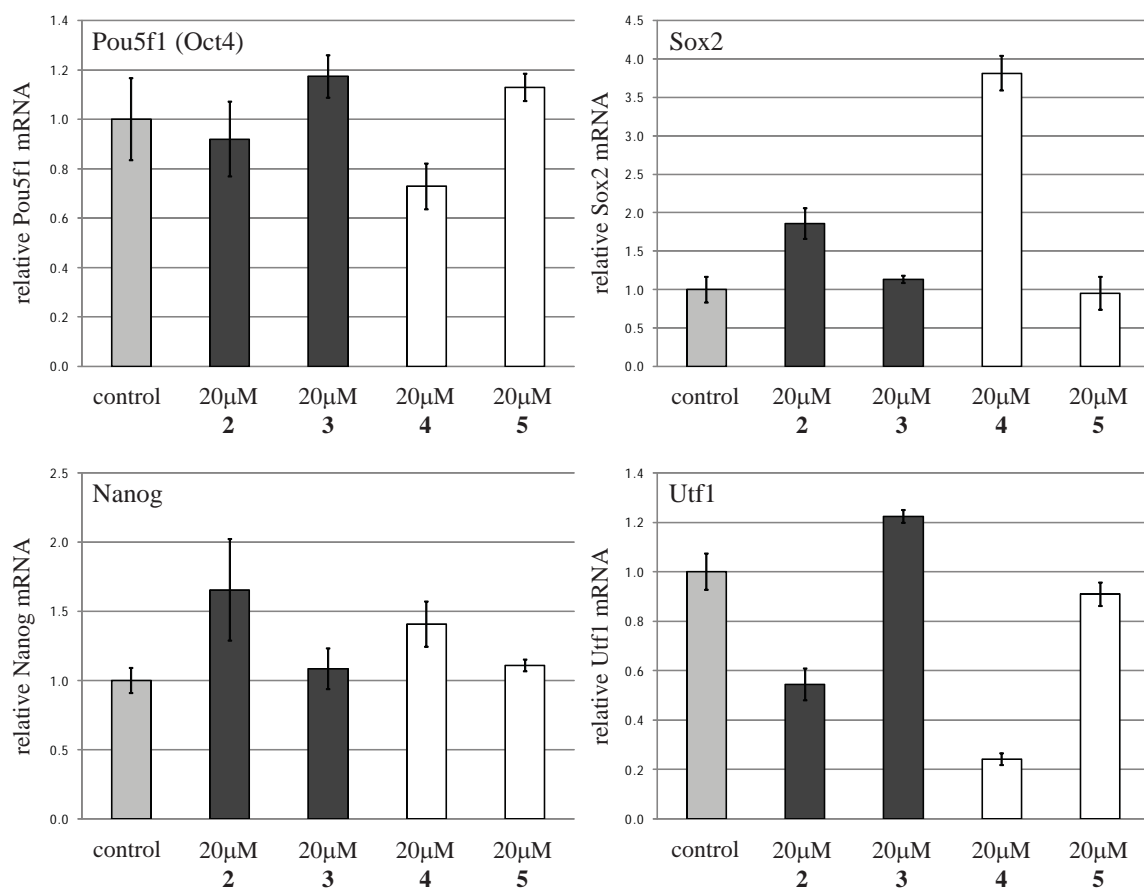




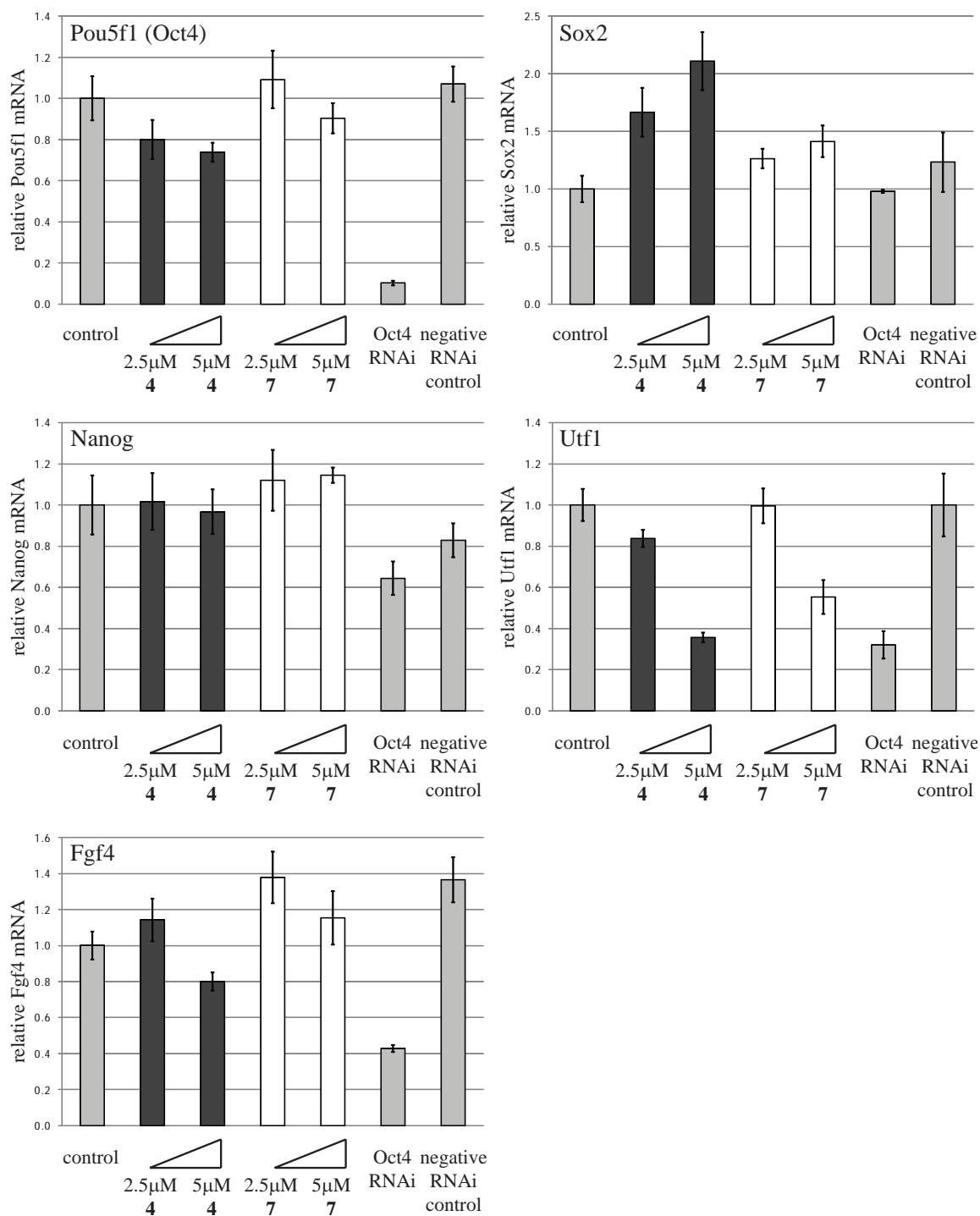
**Figure 5.8.** Quantitative RT-PCR experiments with polyamides **2** and **6** at 10  $\mu$ M and 20  $\mu$ M final concentration with 48 h incubation in P19 mouse embryonal carcinoma cells. Stealth RNAi targeted to Pou5f1 (Oct4) and a negative Stealth RNAi control are also shown. The mRNA levels of Pou5f1 (Oct4), Sox2, Nanog, and Utf1 are calculated relative to Gapdh, and error bars indicate standard deviations.

polyamide **4**. However, at these high concentrations of polyamide **4**, some dead P19 cells were evident at the conclusion of the experiment.

Polyamides **4** and **7** were assayed at lower concentrations of 2.5  $\mu$ M and 5  $\mu$ M to avoid cell toxicity (Figure 5.10). Polyamide **7** was previously shown to inhibit androgen receptor-mediated PSA expression in a prostate cancer cell line.<sup>28</sup> Changes in gene expression levels at these concentration ranges were similar to those observed with match polyamide **2** at 10  $\mu$ M and 20  $\mu$ M. Notably, Utf1 mRNA was decreased to a level comparable to that of the RNAi control, and the Sox2 mRNA level increased two-fold.



**Figure 5.9.** Quantitative RT-PCR experiments with polyamides **2-5** at 20  $\mu$ M final concentration with 48 h incubation in P19 mouse embryonal carcinoma cells. The mRNA levels of Pou5f1 (Oct4), Sox2, Nanog, and Utf1 are calculated relative to Gapdh, and error bars indicate standard deviations.



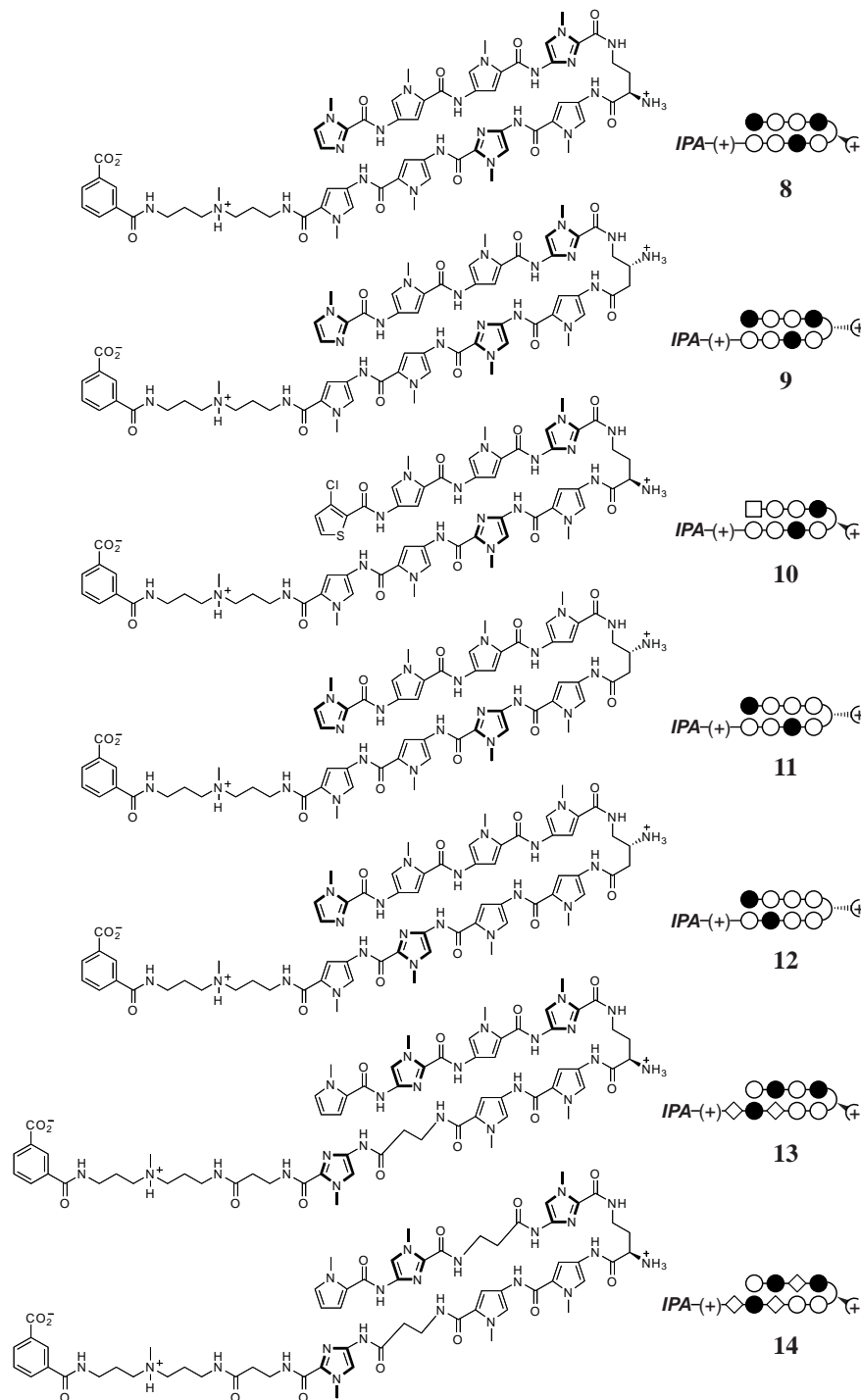
**Figure 5.10.** Quantitative RT-PCR experiments with polyamides **4** and **7** at 2.5  $\mu$ M and 5  $\mu$ M final concentration with 48 h incubation in P19 mouse embryonal carcinoma cells. Stealth RNAi targeted to Pou5f1 (Oct4) and a negative Stealth RNAi control are also shown. The mRNA levels of Pou5f1 (Oct4), Sox2, Nanog, Utf1, and Fgf4 are calculated relative to Gapdh, and error bars indicate standard deviations.

### 5.2.5. Quantitative RT-PCR experiments in R1 mouse ES cells

Based on these results in the model P19 cell line, similar qRT-PCR experiments were performed in R1 mouse embryonic stem (ES) cells. R1 mouse ES cells were maintained according to standard protocols by Caltech Genetically Engineered Mouse Services. In addition to compounds **4** and **7**, two polyamide cocktails were administered (Figure 5.11). Cocktail 1 contained 5  $\mu$ M each of polyamides **6**, **8**, **9**, and **10**. Cocktail 2 contained 5  $\mu$ M each of polyamides **11-14**. The effects of polyamides **8-10** in cell culture have been previously reported.<sup>28,33,35</sup> The time course for the 48-hour incubation experiment is illustrated in Figure 5.12. At  $t = 0$  h, R1 cells were plated in a 24-well format. After 24 hours, the medium was changed, and compounds were applied. After 48 hours, the medium was changed, and compounds were re-applied. After 72 hours, RNA was harvested.

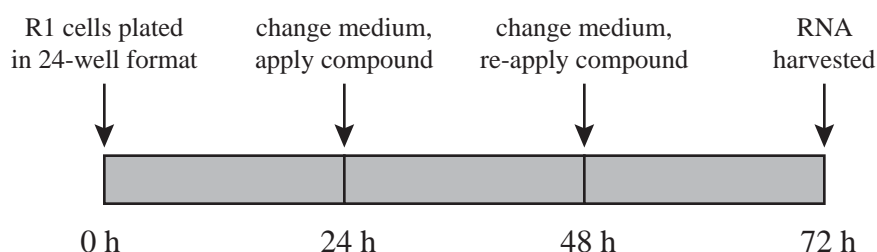
At the conclusion of the 48-hour incubation, all of the experimental samples appeared morphologically similar to the control untreated cells. Gene expression levels were measured by quantitative RT-PCR (Figure 5.13). *Utf1* mRNA levels decreased in the presence of the match polyamide **4** but showed a similar effect in the presence of mismatch polyamide **7**. *Sox2* and *Nanog* mRNA levels increased in the presence of the tested compounds.

An additional experiment with R1 mouse ES cells was designed to increase the incubation period to 72 hours, allowing more time for potential phenotypic changes.<sup>36</sup> Instead of harvesting after 72 hours, the medium was changed and compound was re-applied for a third time (Figure 5.12). RNA was harvested at  $t = 96$  hours after plating. Match polyamide **4** was assayed at 5  $\mu$ M and 10  $\mu$ M concentrations to determine whether higher polyamide concentration and increased incubation time would modulate gene expression further. Retinoic acid has been shown to chemically induce the differentiation of mouse ES cells. A final concentration of 2  $\mu$ M retinoic acid was administered to a set of wells in a solution of absolute ethanol. An analogous set of ethanol control wells was prepared for comparison.

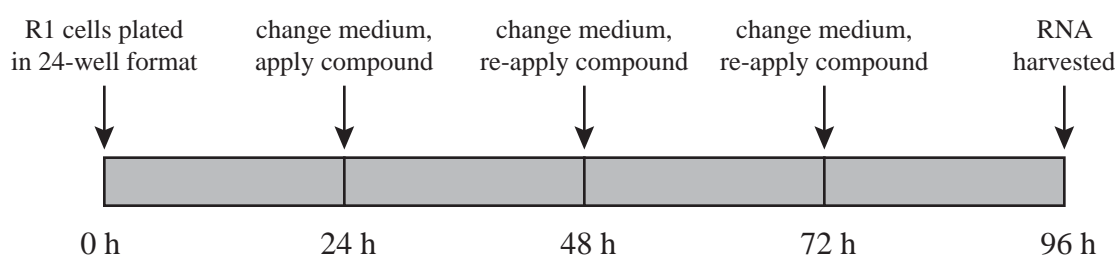


**Figure 5.11.** Structures of polyamide conjugates **8-14**. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; chlorothiophene is shown as a square;  $\beta$ -alanine is shown as a diamond; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; the chiral 2,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a solid wedge linking a half-circle with a plus; the chiral 3,4-diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits with a dashed wedge linking a half-circle with a plus.

### Time course for 48-hour experiment (R1 mouse embryonic stem cells)

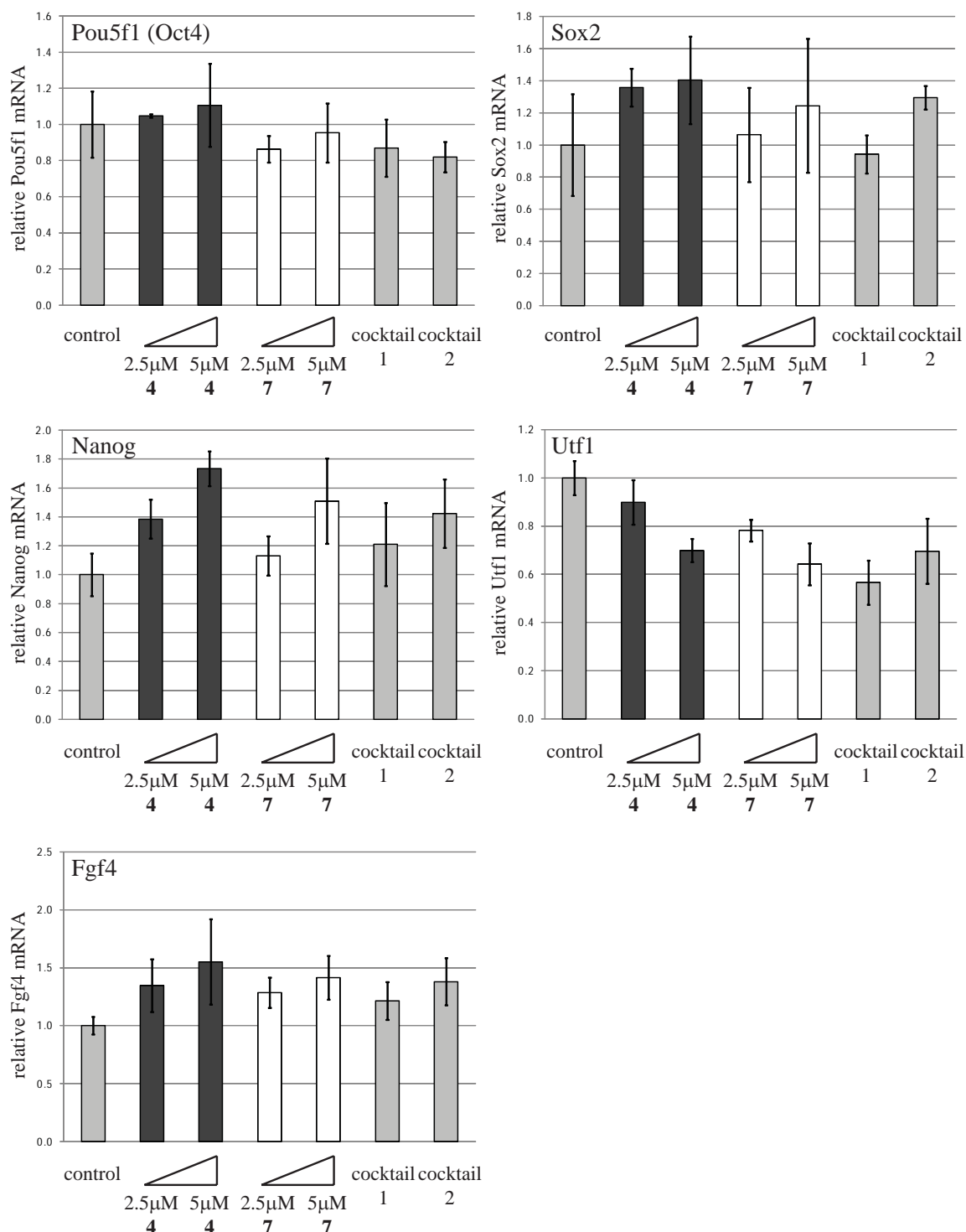


### Time course for 72-hour experiment (R1 mouse embryonic stem cells)

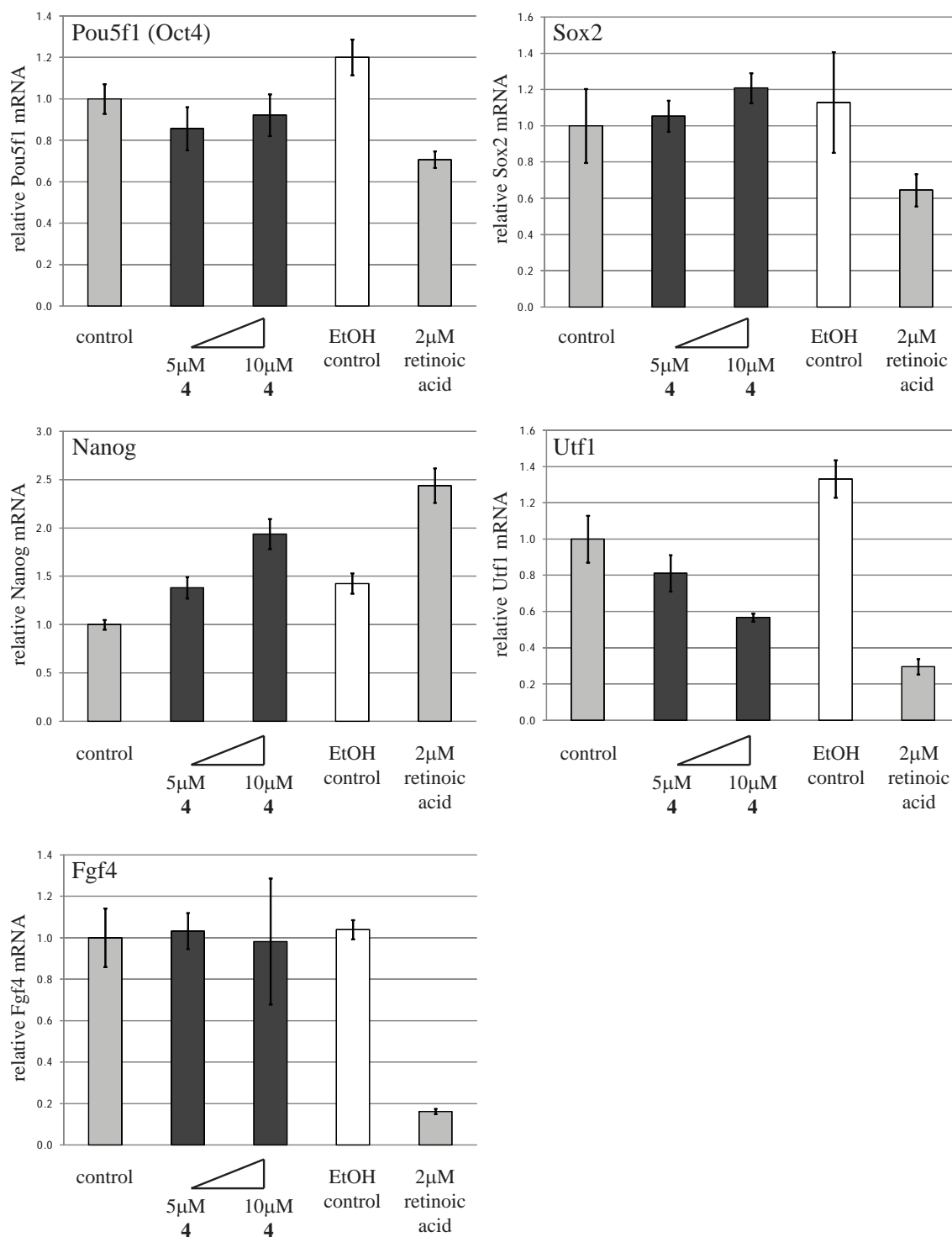


**Figure 5.12.** Time course for 48-hour and 72-hour incubation experiments with R1 mouse embryonic stem cells. R1 cells were plated in 24-well format at  $t = 0$  h, and compounds were first applied at  $t = 24$  h. Medium was changed every 24 hours, and compounds were re-applied at this time. RNA was harvested at the conclusion of the experiment ( $t = 72$  h for the 48-hour incubation experiment;  $t = 96$  h for the 72-hour incubation experiment).

As observed in the previous experiment, Oct4 mRNA levels were unchanged in the presence of match polyamide (Figure 5.14). After 72 hours of incubation with 10  $\mu$ M of polyamide **4**, Utf1 mRNA levels decreased over 40%. A comparison of 48-hour and 72-hour incubation periods for 5  $\mu$ M of polyamide **4** indicates that a similar level of downregulation was observed at both time points. Nanog mRNA levels increased two-fold after 72 hours of incubation with 10  $\mu$ M of polyamide **4**. Based on a comparison of experiments with 5  $\mu$ M of polyamide **4**, Nanog mRNA levels were higher after 48 hours of incubation than after 72 hours. Treatment with 2  $\mu$ M retinoic acid yielded a 70% decrease in Utf1 mRNA levels and a 2.5-fold increase in Nanog mRNA levels.



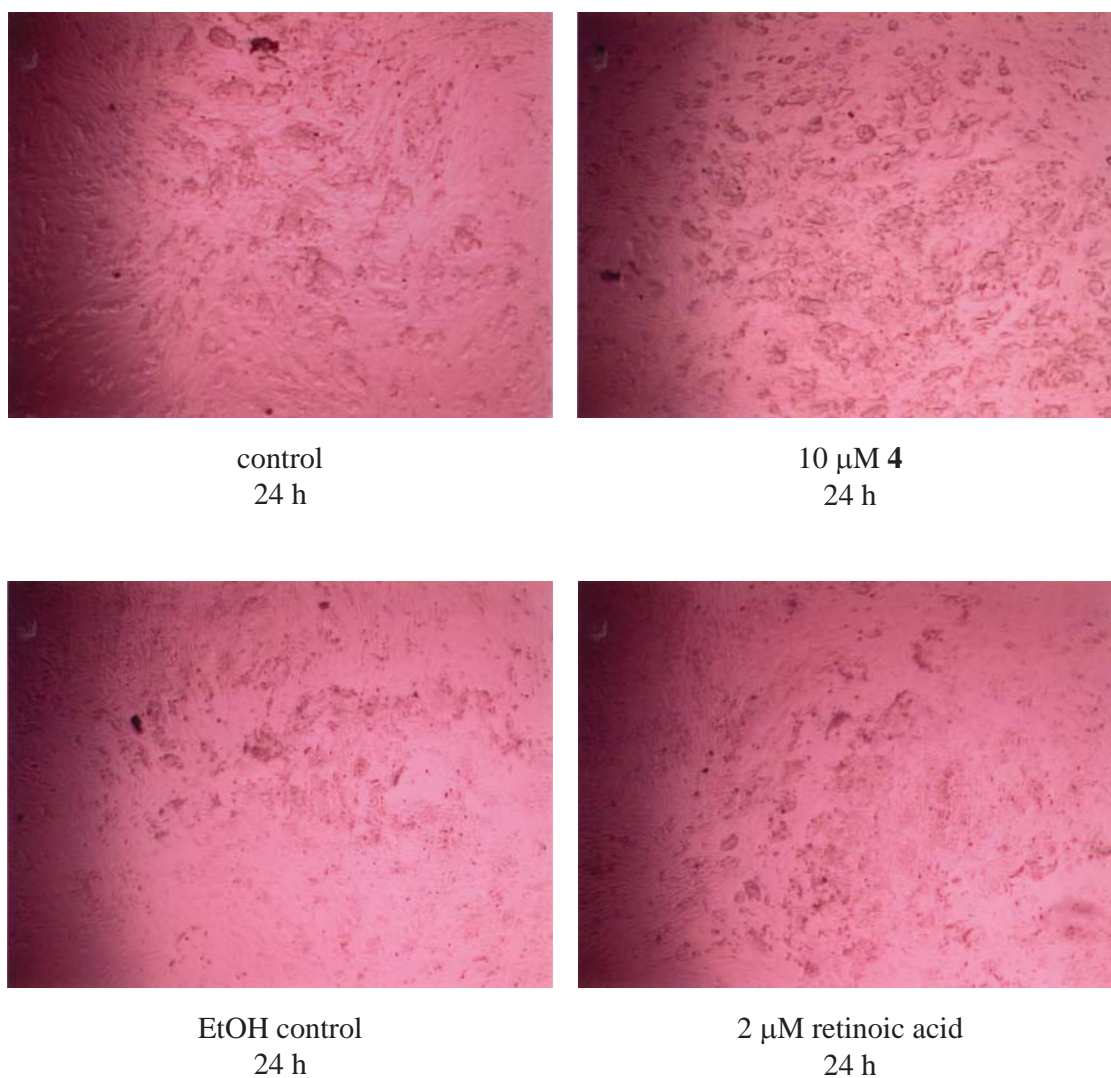
**Figure 5.13.** Quantitative RT-PCR experiments with polyamides **4** and **7** at 2.5  $\mu$ M and 5  $\mu$ M final concentration with 48 h incubation in R1 mouse embryonic stem cells. Cocktail 1 contains 5  $\mu$ M each of polyamides **6**, **8**, **9**, and **10**. Cocktail 2 contains 5  $\mu$ M each of polyamides **11-14**. The mRNA levels of Pou5f1 (Oct4), Sox2, Nanog, Utf1, and Fgf4 are calculated relative to Gapdh, and error bars indicate standard deviations.



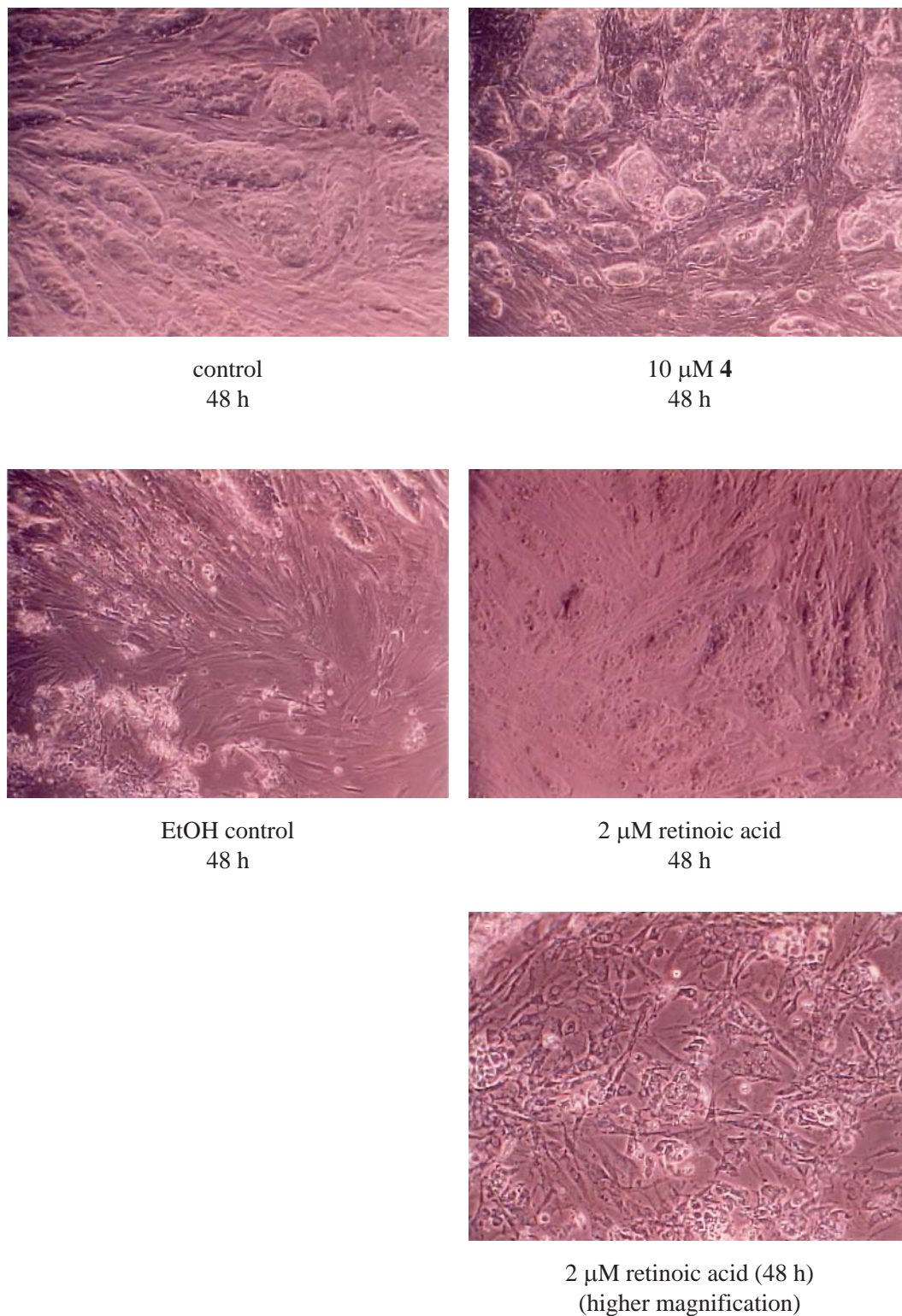
**Figure 5.14.** Quantitative RT-PCR experiments with polyamide **4** at 5  $\mu$ M and 10  $\mu$ M final concentration with 72 h incubation in R1 mouse embryonic stem cells. Retinoic acid was added as a solution in ethanol to a final concentration of 2  $\mu$ M. The mRNA levels of Pou5f1 (Oct4), Sox2, Nanog, Utf1, and Fgf4 are calculated relative to Gapdh, and error bars indicate standard deviations.



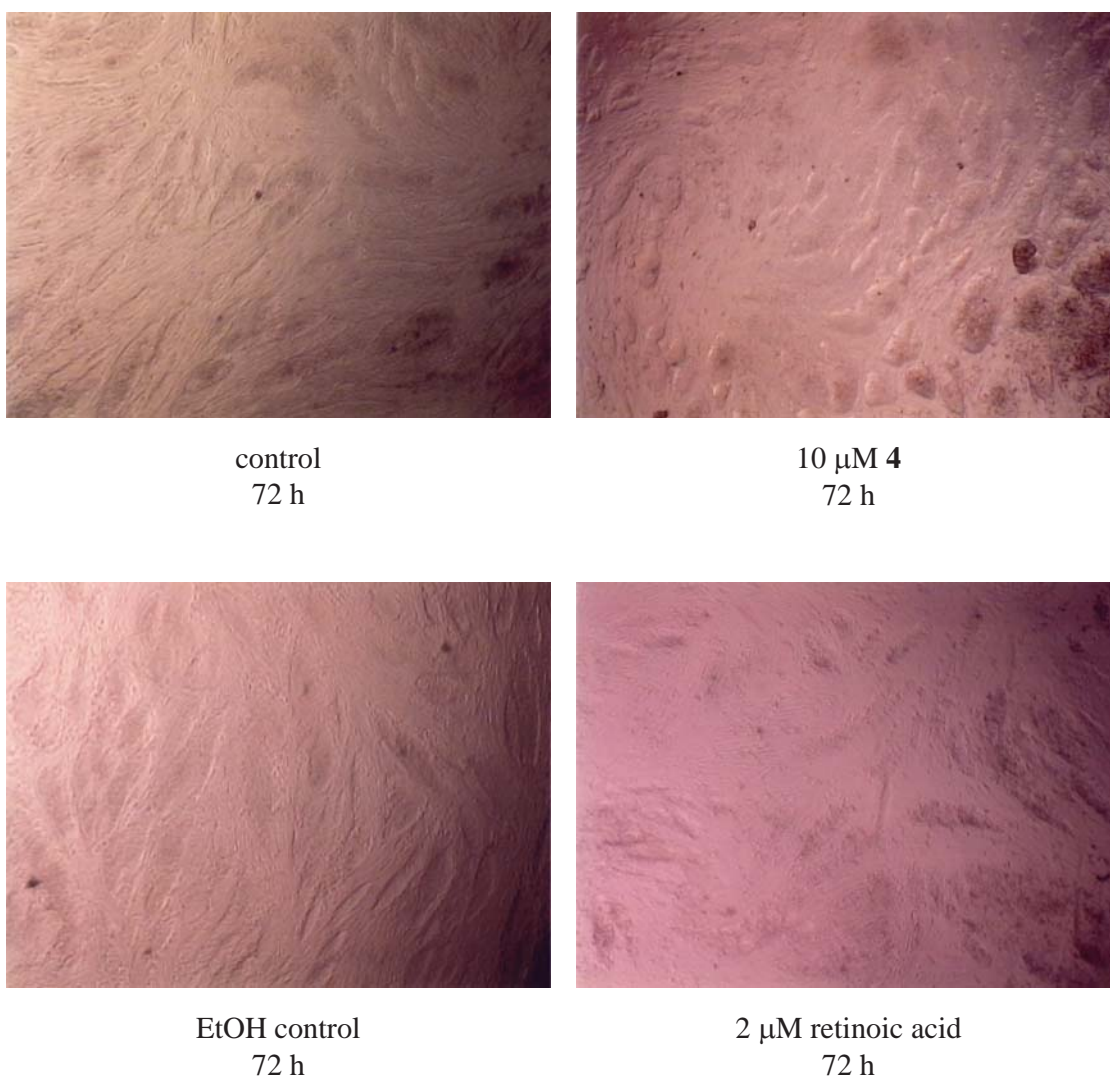
The cell morphology of polyamide-treated and EtOH-treated cells was similar to that of untreated control R1 mouse ES cells. However, the cell morphology of the retinoic acid-treated cells diverged from that of control cells after 24 hours of incubation (Figure 5.15). Retinoic acid-treated cells were healthy after 48 hours and remained divergent in appearance when compared with control cells (Figure 5.16). Based on a picture of these cells at higher magnification, it appeared that these cells were differentiating. After 72 hours of incubation, the polyamide-treated and EtOH-treated cells were healthy and nearly confluent (Figure 5.17). The morphology of these experimental samples was similar to that of the untreated cells. Retinoic acid-treated cells continued to show a different morphology. However, at the conclusion of the 72-hour incubation period, these cells began to look unhealthy, and some colonies had begun to lift off the bottom of the plate.



**Figure 5.15.** Images of R1 mouse embryonic stem cells after 24 h treatment. Untreated cell samples, as well as samples treated with 10  $\mu$ M **4**, EtOH only, and 2  $\mu$ M retinoic acid dosed in EtOH, are shown following 24 hours of incubation with the compound indicated.



**Figure 5.16.** Images of R1 mouse embryonic stem cells after 48 h treatment. Untreated cell samples, as well as samples treated with 10  $\mu\text{M}$  **4**, EtOH only, and 2  $\mu\text{M}$  retinoic acid dosed in EtOH, are shown following 48 hours of incubation with the compound indicated. A higher magnification image of retinoic acid-treated sample is shown at bottom.



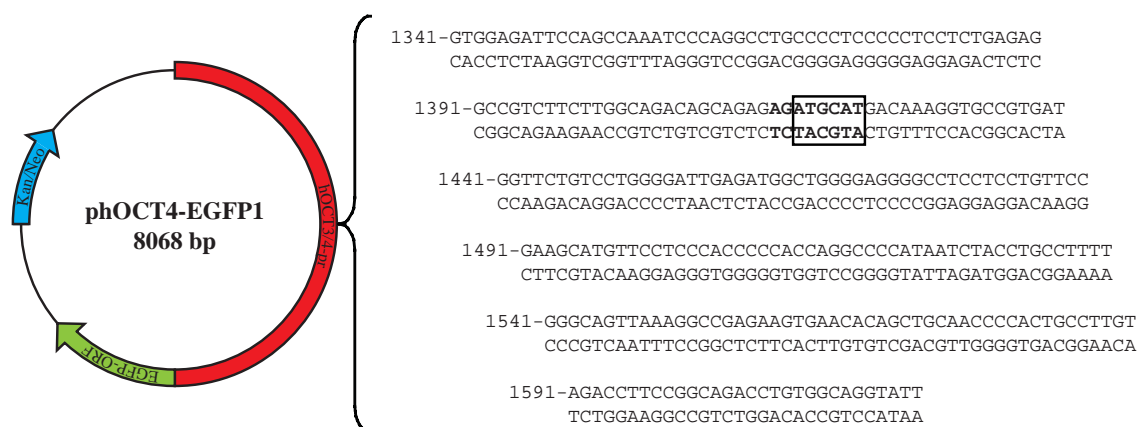
**Figure 5.17.** Images of R1 mouse embryonic stem cells after 72 h treatment. Untreated cell samples, as well as samples treated with 10  $\mu\text{M}$  4, EtOH only, and 2  $\mu\text{M}$  retinoic acid dosed in EtOH, are shown following 72 hours of incubation with the compound indicated.

### 5.3. Conclusion and future directions

The gene expression levels of Oct4 target genes that contain an Oct4 transcription factor binding site in the enhancer region were assayed by quantitative RT-PCR. Based on melting temperature analysis of a series of polyamides designed to target the DNA sequence 5'-WTGCWW-3', polyamide **4**, which contains a free chiral amine on the 3,4-diaminobutyric turn linkage, exhibits the highest binding affinity. After 48 h incubation time with 5  $\mu$ M of polyamide **4** in a P19 mouse embryonal carcinoma cell line, Utf1 mRNA levels were decreased by 60%. However, it was surprising to observe that Sox2 mRNA levels were increased two-fold. After 72 h incubation time with 10  $\mu$ M of polyamide **4** in a R1 mouse embryonic stem cell line, Utf1 mRNA levels were decreased by 40%. In contrast, Sox2 mRNA levels only increased modestly, and Nanog mRNA levels were increased two-fold. The cell morphology of polyamide-treated samples appeared similar to untreated samples by visual inspection.

The treatment of R1 embryonic stem cells with 2  $\mu$ M retinoic acid chemically induces differentiation that was evident after 24 h and 48 h. While Utf1 downregulation and Nanog de-repression were comparable in the retinoic acid-treated and polyamide-treated samples, no changes in cell morphology were observed after incubation with polyamide **4** for 72 h. Polyamide-induced differentiation of mouse embryonic stem cells may require longer time-course experiments. However, the healthy growth of mouse embryonic stem cells when incubated with 10  $\mu$ M of polyamide **4** is an encouraging sign that higher concentrations may be tolerated in future experiments.





**Figure 5.18.** The partial sequence of plasmid phOCT4-EGFP1. The equilibrium association constant ( $K_a$ ) of polyamide **1** was determined by DNase I footprint titration experiments on the Oct4 binding site of the phOCT4-EGFP1 plasmid obtained from Dr. Wei Cui of the Roslin Institute.<sup>34</sup>

## 5.4. Materials and methods

### 5.4.1. Polyamide synthesis

Polyamides were synthesized with oxime resin (compounds **1-12**, Novabiochem) or Boc- $\beta$ -Ala-PAM resin (compounds **13-14**, Peptides International) according to published manual solid-phase synthesis protocols.<sup>31,32</sup> For compounds **1-3**, **6**, **8**, **10**, **13**, and **14**, *N*- $\alpha$ -9-fluorenylmethoxycarbonyl-*N*- $\gamma$ -*tert*-butoxycarbonyl-D-2,4-diaminobutyric acid (Fmoc-D-Dab(Boc)-OH, Peptides International) was used for the turn linkage. The protected <sup>FmocHN</sup> $\gamma$ -turn amine was deprotected with 20% piperidine in DMF and reprotected as the Boc derivative with a solution of Boc<sub>2</sub>O (Fluka) and DIEA in DMF. For compounds **1** and **3**, the protected <sup>BocHN</sup> $\gamma$ -turn amine was deprotected with 20% trifluoroacetic acid (TFA, Halocarbon) in DCM and reprotected as the acetylated derivative with a solution of acetic anhydride and DIEA in NMP. For compounds **2**, **6**, **8**, **10**, **13**, and **14**, the Boc-protected amine was left untreated. For compounds **4**, **5**, **7**, **9**, **11**, and **12**, *N*- $\beta$ -benzyloxycarbonyl-*N*- $\gamma$ -*tert*-butoxycarbonyl-R-3,4-diaminobutyric acid (Z-Dbu(Boc)-OH, Senn Chemicals AG) was used for the turn linkage. For compounds **1-5**, the chlorothiophene-imidazole cap was added on solid phase as a dimer.

The acetylated or Boc-protected or Cbz-protected resin was cleaved with 1 mL of 3,3'-diamino-*N*-methyldipropylamine at 37°C with agitation for 16 h. Products were purified by preparatory reverse-phase high-performance liquid chromatography (HPLC) on a Beckman Gold system using a Phenomenex Gemini 21.2 mm  $\times$  250 mm, 5  $\mu$ m 110 Å C<sub>18</sub> reverse-phase column in 0.1% (w/v) TFA with acetonitrile as the eluent. The appropriate fractions were lyophilized after characterization by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF-MS). HPLC analysis was performed on a Beckman Gold system equipped with a diode-array detector using a Phenomenex Gemini 4.6 mm  $\times$  250 mm, 5  $\mu$ m 110 Å C<sub>18</sub> reverse-phase column. UV spectra were measured on an Agilent Technologies 8453 UV-vis ChemStation spectrophotometer. MALDI-TOF-MS was carried

out on an Applied Biosystems Voyager DE-PRO.

FITC conjugate **1** was formed by reacting fluorescein-5-isothiocyanate (Invitrogen) with the acetylated polyamide in a solution of DIEA (20 equiv) and DMF for 1 h at room temperature. The product was purified by preparatory reverse-phase HPLC, and lyophilization of the appropriate fractions yielded the polyamide-fluorescein conjugate **1**, which was characterized as described above.

IPA conjugates **2-14** were formed by pre-activating isophthalic acid (IPA, 3.0 equiv) with PyBOP (2.9 equiv, Novabiochem) in a solution of DIEA (20 equiv) and DMF at 37°C for 30 min, followed by reaction of the activated solution with the polyamide for 1 h at room temperature. For conjugate **3**, the product was purified by preparatory reverse-phase HPLC, and lyophilization of the appropriate fractions yielded the acetylated IPA conjugate **3**. For conjugates **2, 6, 8, 10, 13**, and **14**, conjugates were deprotected with neat TFA and triethylsilane for 30 min at room temperature before purification.

For conjugates **4, 7, 9, 11**, and **12**, the product was purified by solid-phase extraction using a Waters SepPak C-18 Light Cartridge following conjugation with isophthalic acid. The eluted fractions in acetonitrile and methanol were lyophilized. The Cbz protecting group was removed according to published protocols.<sup>28</sup> A 500 nmol fraction of the Cbz-protected hairpin polyamide conjugate was dissolved in 450  $\mu$ L TFA and 50  $\mu$ L trifluoromethanesulfonic acid (TFMSA). After a reaction time of 5 min, the solution was flash frozen in liquid nitrogen and overlaid with 1 mL DMF. The thawed solution was diluted in 20% aqueous acetonitrile for purification by preparatory reverse-phase HPLC. Lyophilization of the appropriate fractions yielded the IPA conjugates **4, 7, 9, 11**, and **12**.

Conjugate **5** was synthesized by acetylation of conjugate **4**, according to published protocols.<sup>28</sup> Polyamide conjugate **4** was dissolved in 1200  $\mu$ L DMF. A mixture of 135  $\mu$ L pyridine and 15  $\mu$ L acetic anhydride was added, and the reaction was allowed to proceed for 30 min at room temperature. The product was purified by preparatory reverse-phase HPLC, and lyophilization of the appropriate fractions yielded the acetylated IPA conjugate



5. For conjugates **1-14**, extinction coefficients were calculated according to standard protocols.<sup>37</sup> Chemicals not otherwise specified were from Aldrich.

**CtImPyPy-(R)<sup>AcHN</sup> $\gamma$ -PyImPyPy-(+)-FITC (1).** Synthesized with 2,4-diaminobutyric acid turn subunit and conjugated with fluorescein-5-isothiocyanate (FITC). UV-vis  $\lambda_{\max}$  = 307, 444 nm; MALDI-TOF-MS  $m/z$  1288.94 ( $C_{58}H_{71}ClN_{21}O_{10}S^+$  calculated [M+H-FITC]<sup>+</sup> 1288.51)

**CtImPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy-(+)-IPA (2).** Synthesized with 2,4-diaminobutyric acid turn subunit and conjugated with isophthalic acid (IPA). UV-vis  $\lambda_{\max}$  = 318 nm; MALDI-TOF-MS  $m/z$  1394.86 ( $C_{64}H_{73}ClN_{21}O_{12}S^+$  calculated [M+H]<sup>+</sup> 1394.52)

**CtImPyPy-(R)<sup>AcHN</sup> $\gamma$ -PyImPyPy-(+)-IPA (3).** Synthesized with 2,4-diaminobutyric acid turn subunit and conjugated with isophthalic acid (IPA). UV-vis  $\lambda_{\max}$  = 314 nm; MALDI-TOF-MS  $m/z$  1436.74 ( $C_{66}H_{75}ClN_{21}O_{13}S^+$  calculated [M+H]<sup>+</sup> 1436.53)

**CtImPyPy-(R)<sup>H<sub>2</sub>N</sup> $\gamma$ -PyImPyPy-(+)-IPA (4).** Synthesized with 3,4-diaminobutyric acid turn subunit and conjugated with isophthalic acid (IPA). UV-vis  $\lambda_{\max}$  = 318 nm; MALDI-TOF-MS  $m/z$  1394.76 ( $C_{64}H_{73}ClN_{21}O_{12}S^+$  calculated [M+H]<sup>+</sup> 1394.52)

**CtImPyPy-(R)<sup>AcHN</sup> $\gamma$ -PyImPyPy-(+)-IPA (5).** Synthesized with 3,4-diaminobutyric acid turn subunit and conjugated with isophthalic acid (IPA). UV-vis  $\lambda_{\max}$  = 314 nm; MALDI-TOF-MS  $m/z$  1436.78 ( $C_{66}H_{75}ClN_{21}O_{13}S^+$  calculated [M+H]<sup>+</sup> 1436.53)

The synthesis and characterization of polyamide conjugates **6-14** have been reported previously.<sup>33,35</sup>

#### 5.4.2. Quantitative DNase I footprint titrations

Quantitative DNase I footprint titration experiments were performed on the 5'-<sup>32</sup>P-end-labeled PCR products of plasmid phOCT4-EGFP1 according to standard protocols.<sup>37</sup> The phOCT4-EGFP1 plasmid was a gift from Dr. Wei Cui of the Roslin Institute.<sup>34</sup> Radiolabeled DNA was equilibrated with a solution of polyamide **1** for 14-16 h at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0

prior to DNase I cleavage. Chemical sequencing reactions were performed according to published methods.<sup>38,39</sup> Storage phosphor autoradiography was performed on a Molecular Dynamics Typhoon 8600 phosphorimager. 18 M $\Omega$  water was obtained from an AquaMAX Ultra water purification system, and all buffers were 0.2  $\mu$ m filtered.

#### **5.4.3. UV absorption spectrophotometry**

Melting temperature analysis was performed according to published protocols.<sup>28</sup> UV absorption analysis was performed on a Varian Cary 100 spectrophotometer. Analysis was performed in a degassed aqueous buffer of 10 mM sodium cacodylate, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0. The 12 bp DNA oligonucleotides 5'-GAGATGCATGAC-3' and 5'-GTCATGCATCTC-3' were obtained with HPLC purification from Integrated DNA Technologies. The DNA oligonucleotides and polyamides **1-5** were mixed in 1:1 stoichiometry to a final concentration of 2  $\mu$ M in a volume of 1 mL for each experiment. Prior to analysis, samples were heated to 90°C and cooled to a starting temperature of 25°C with a heating or cooling rate of 5°C/min for each ramp. Denaturation profiles were recorded at  $\lambda = 260$  nm from 25°C to 90°C with a heating rate of 0.5°C/min. The reported melting temperatures were defined as the maximum of the first derivative of the denaturation profile.

#### **5.4.4. Cell cultures**

Cell lines were cultured in a 5% CO<sub>2</sub> atmosphere at 37°C. The human cervical cancer cell line HeLa and the mouse embryonal carcinoma cell line P19 were obtained from ATCC and cultured according to recommended protocols. HeLa cells were cultured in DMEM medium (Invitrogen) supplemented with 10% fetal bovine serum (FBS, Omega Scientific) and 1% penicillin/streptomycin solution (Mediatech).<sup>29,30</sup> P19 cells were cultured in Alpha Minimum Essential Medium with ribonucleosides and deoxyribonucleosides (Invitrogen) supplemented with 7.5% bovine calf serum (Omega Scientific) and 2.5% fetal bovine

serum. R1 mouse embryonic stem cells were cultured according to standard protocols by Caltech Genetically Engineered Mouse Services. The R1 mouse embryonic stem cell line was obtained from the Nagy lab, where they were developed.

#### **5.4.5. Confocal microscopy**

Confocal microscopy experiments were performed according to published protocols.<sup>29,30</sup> Cell lines were trypsinized (Mediatech) for 5 min at 37°C, centrifuged for 10 min at 4°C at 100 × g, and resuspended in fresh medium to a concentration of  $1.33 \times 10^5$  cells/mL. Incubations were performed by adding 150 µL of cells into culture dishes equipped with glass bottoms for direct imaging (MatTek). Cells were grown in the glass-bottom culture dishes for 24 h. The medium was then removed and replaced with 147 µL of fresh medium, followed by addition of 3 µL of the 100 µM solution of polyamide-fluorescein conjugate **1** for a final polyamide concentration of 2 µM. Cells were incubated in a 5% CO<sub>2</sub> atmosphere at 37°C for 12 h. Imaging was performed with a 63× oil-immersion objective lens on a Zeiss LSM 5 Pascal inverted laser scanning microscope. Polyamide-fluorescein conjugate fluorescence and visible light images were obtained using standard filter sets for fluorescein.<sup>29,30</sup>

#### **5.4.6. Quantitative RT-PCR experiments**

For cell culture experiments, P19 mouse embryonal carcinoma cells were plated in a 24-well format with a concentration of  $5 \times 10^4$  cells/mL in a volume of 500 µL. Polyamides were added to the medium at this time as a solution in water. RNA was harvested after 48 h incubation using an RNeasy Mini Kit (Qiagen) and reverse transcribed using SuperScript II Reverse Transcriptase (Invitrogen). Quantitative RT-PCR was performed using SYBR GREEN PCR Master Mix (Applied Biosystems) on an Applied Biosystems 7300 Real Time PCR System. Gene expression levels were measured relative to Gapdh.

R1 mouse embryonic stem cells were cultured according to standard protocols by Caltech Genetically Engineered Mouse Services. The R1 mouse embryonic stem cell line was obtained from the Nagy lab, where they were developed. R1 mouse embryonic stem cells were plated in a 24-well format on a feeder layer of mouse primary fibroblasts. Polyamides were added to the medium as a solution in water. Cocktail 1 contains 5  $\mu$ M each of polyamides **6**, **8**, **9**, and **10**. Cocktail 2 contains 5  $\mu$ M each of polyamides **11-14**. Retinoic acid (Fluka) was added to the medium as a solution in ethanol. As a control, an equivalent volume of ethanol without retinoic acid was added to the medium. RNA was harvested after 48 h or 72 h using the same procedure described above.

Stealth RNAi transfection experiments were performed according to recommended protocols for P19 cells from Invitrogen. For Oct4 (Pou5f1) Stealth RNAi experiments, P19 cells were plated in a 24-well format with a concentration of  $1 \times 10^5$  cells/mL in a volume of 400  $\mu$ L growth medium without antibiotics one day before transfection. For each transfection sample, 50 pmol of Stealth RNAi was diluted in 50  $\mu$ L of Opti-MEM I Reduced Serum Medium. Separately, 1  $\mu$ L of Lipofectamine 2000 was diluted in 50  $\mu$ L of Opti-MEM I Reduced Serum Medium and incubated for 15 min at room temperature. The diluted Stealth RNAi and Lipofectamine 2000 solutions were combined and incubated for 15 min at room temperature before addition to each well.

Stealth RNAi for the Pou5f1 gene was obtained from Invitrogen. Duplex #1 contained the 25 bp RNA oligonucleotides 5'-CCAAUGCCGUGAAGUUGGAGAAGGU-3' and 5'-ACCUUCUCCAACUUCACGGCAUUGG-3'. Duplex #2 contained the 25 bp RNA oligonucleotides 5'-CCCGGAAGAGAAAGCGAACUAGCAU-3' and 5'-AUGCUAGUUCGCUUUCUCUUCGCGG-3'. Duplex #3 contained the 25 bp RNA oligonucleotides 5'-CCAAUCAGCUUGGGCUAGAGAAGGA-3' and 5'-UCCUUCUCUAGCCCAAGCUGAUUGG-3'. While all three Stealth RNAi duplexes

achieved knockdown of the gene of interest, duplex #2 was the most effective. A negative Stealth RNAi containing medium GC content (Invitrogen, Lot #434330) was used as a control.

#### **5.4.7. Primers for quantitative RT-PCR**

The following sets of primers were used for quantitative RT-PCR experiments. Oligonucleotides were obtained from Integrated DNA Technologies in LabReady format normalized to 100  $\mu$ M in IDTE pH 8.0. PCR amplicons were analyzed by gel electrophoresis on a 2% agarose gel (Invitrogen) stained with SYBR Gold (Invitrogen) and imaged on a Molecular Dynamics Typhoon 8600. PCR amplification efficiency and dissociation curve analysis were performed as recommended by Applied Biosystems.

Gapdh\_03\_L: 5'-CGTCCCGTAGACAAAATGGT-3' (20 bp, 36-55)

Gapdh\_03\_R: 5'-TTGATGGCAACAATCTCCAC-3' (20 bp, 126-145)

110 bp amplicon

Pou5f1\_02\_L: 5'-GTTGGAGAAGGTGGAACCAA-3' (20 bp, 415-434)

Pou5f1\_02\_R: 5'-TCTTCTGCTTCAGCAGCTTG-3' (20 bp, 490-509)

95 bp amplicon

Sox2\_01\_L: 5'-GCACATGAACGGCTGGAGCAACG-3' (23 bp, 900-922)

Sox2\_01\_R: 5'-TGCTGCGAGTAGGACATGCTGTAGG-3' (25 bp, 1082-1106)

207 bp amplicon

Nanog\_01\_L: 5'-TCTTCCTGGTCCCCACAGTTT-3' (21 bp, 227-247)

Nanog\_01\_R: 5'-GCAAGAATAGTTCTCGGGATGAA-3' (23 bp, 304-326)

100 bp amplicon

Utf1\_01\_L: 5'-GTCCCTCTCCGCGTTAGC-3' (18 bp, 531-548)

Utf1\_01\_R: 5'-GGAAGAAGTGAATCTGAGCG-3' (20 bp, 620-639)

109 bp amplicon

Utf1\_02\_L: 5'-TGTCCCGGTGACTACGTCT-3' (19 bp, 102-120)

Utf1\_02\_R: 5'-CCCAGAAGTAGCTCCGTCTCT-3' (21 bp, 216-236)

135 bp amplicon

Fbx15\_01\_L: 5'-TGCCAATTGTTGGGAGTACA-3' (20 bp, 111-130)

Fbx15\_01\_R: 5'-CATGCTGCTTCGTGACAGAT-3' (20 bp, 184-203)

93 bp amplicon

Fbx15\_04\_L: 5'-TGTTAGAGGCTCATCTGTCACG-3' (22 bp, 172-193)

Fbx15\_04\_R: 5'-GATGGCATTCTGTCCAGGGAT-3' (21 bp, 256-276)

105 bp amplicon

Fgf4\_02\_L: 5'-GTGTGCCTTTCTTTACCGACG-3' (21 bp, 547-567)

Fgf4\_02\_R: 5'-CTGAGGGCCATGAACATACCG-3' (21 bp, 632-652)

106 bp amplicon

### **Acknowledgments.**

We thank Shirley Pease and the Caltech Genetically Engineered Mouse Services for assistance with R1 mouse embryonic stem cell experiments. We are grateful to the National Institutes of Health for research support.

John Phillips performed quantitative DNase I footprint titration experiments. Daniel Harki synthesized polyamides **6**, **7**, and **9-14**. James Puckett synthesized polyamides **7** and **8**.

## References

1. Wobus, A.M. and Boheler, K.R. (2005) Embryonic stem cells: Prospects for developmental biology and cell therapy. *Physiol. Rev.*, **85**, 635-678.
2. Chambers, I. and Smith, A. (2004) Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene*, **23**, 7150-7160.
3. Rodda, D.J., Chew, J.L., Lim, L.H., Loh, Y.H., Wang, B., Ng, H.H. and Robson, P. (2005) Transcriptional regulation of Nanog by Oct4 and Sox2. *J. Biol. Chem.*, **280**, 24731-24737.
4. Pesce, M., Gross, M.K. and Scholer, H.R. (1998) In line with our ancestors: Oct-4 and the mammalian germ. *Bioessays*, **20**, 722-732.
5. Pesce, M. and Scholer, H.R. (2001) Oct-4: Gatekeeper in the beginnings of mammalian development. *Stem Cells*, **19**, 271-278.
6. Pan, G.J., Chang, Z.Y., Scholer, H.R. and Pei, D.Q. (2002) Stem cell pluripotency and transcription factor Oct4. *Cell Res.*, **12**, 321-329.
7. Niwa, H., Miyazaki, J. and Smith, A.G. (2000) Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nature Genet.*, **24**, 372-376.
8. Takahashi, K. and Yamanaka, S. (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*, **126**, 663-

676.

9. Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K. and Yamanaka, S. (2007) Induction of pluripotent stem cells from adult human fibroblasts by defined factors. *Cell*, **131**, 861-872.
10. Yu, J.Y., Vodyanik, M.A., Smuga-Otto, K., Antosiewicz-Bourget, J., Frane, J.L., Tian, S., Nie, J., Jonsdottir, G.A., Ruotti, V., Stewart, R., Slukvin, II and Thomson, J.A. (2007) Induced pluripotent stem cell lines derived from human somatic cells. *Science*, **318**, 1917-1920.
11. Lewitzky, M. and Yamanaka, S. (2007) Reprogramming somatic cells towards pluripotency by defined factors. *Curr. Opin. Biotechnol.*, **18**, 467-473.
12. Niwa, H. (2007) How is pluripotency determined and maintained? *Development*, **134**, 635-646.
13. Herr, W. and Cleary, M.A. (1995) The Pou Domain - Versatility in Transcriptional Regulation by a Flexible 2-in-One DNA-Binding Domain. *Genes Dev.*, **9**, 1679-1693.
14. Chew, J.L., Loh, Y.H., Zhang, W.S., Chen, X., Tam, W.L., Yeap, L.S., Li, P., Ang, Y.S., Lim, B., Robson, P. and Ng, H.H. (2005) Reciprocal transcriptional regulation of Pou5f1 and Sox2 via the Oct4/Sox2 complex in embryonic stem cells. *Mol. Cell. Biol.*, **25**, 6031-6046.
15. Remenyi, A., Lins, K., Nissen, L.J., Reinbold, R., Scholer, H.R. and Wilmanns, M.



- (2003) Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev.*, **17**, 2048-2059.
16. Tomioka, M., Nishimoto, M., Miyagi, S., Katayanagi, T., Fukui, N., Niwa, H., Muramatsu, M. and Okuda, A. (2002) Identification of Sox-2 regulatory region which is under the control of Oct-3/4-Sox-2 complex. *Nucleic Acids Res.*, **30**, 3202-3213.
  17. Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., Ko, M.S.H. and Niwa, H. (2007) Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat. Cell Biol.*, **9**, 625-U26.
  18. Boer, B., Kopp, J., Mallanna, S., Desler, M., Chakravarthy, H., Wilder, P.J., Bernadt, C. and Rizzino, A. (2007) Elevating the levels of Sox2 in embryonal carcinoma cells and embryonic stem cells inhibits the expression of Sox2 : Oct-3/4 target genes. *Nucleic Acids Res.*, **35**, 1773-1786.
  19. Pan, G.J. and Thomson, J.A. (2007) Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Res.*, **17**, 42-49.
  20. Loh, Y.H., Wu, Q., Chew, J.L., Vega, V.B., Zhang, W.W., Chen, X., Bourque, G., George, J., Leong, B., Liu, J., Wong, K.Y., Sung, K.W., Lee, C.W.H., Zhao, X.D., Chiu, K.P., Lipovich, L., Kuznetsov, V.A., Robson, P., Stanton, L.W., Wei, C.L., Ruan, Y.J., Lim, B. and Ng, H.H. (2006) The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nature Genet.*, **38**, 431-

440.

21. Nishimoto, M., Fukushima, A., Okuda, A. and Muramatsu, M. (1999) The gene for the embryonic stem cell coactivator UTF1 carries a regulatory element which selectively interacts with a complex composed of Oct-3/4 and Sox-2. *Mol. Cell. Biol.*, **19**, 5453-5465.
22. Tokuzawa, Y., Kaiho, E., Maruyama, M., Takahashi, K., Mitsui, K., Maeda, M., Niwa, H. and Yamanaka, S. (2003) Fbx15 is a novel target of Oct3/4 but is dispensable for embryonic stem cell self-renewal and mouse development. *Mol. Cell. Biol.*, **23**, 2699-2708.
23. Dailey, L., Yuan, H.B. and Basilico, C. (1994) Interaction between a Novel F9-Specific Factor and Octamer-Binding Proteins Is Required for Cell-Type-Restricted Activity of the Fibroblast Growth-Factor-4 Enhancer. *Mol. Cell. Biol.*, **14**, 7758-7769.
24. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.*, **9**, 2215-2235.
25. 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.
26. 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.
27. Herman, D.M., Baird, E.E. and Dervan, P.B. (1998) Stereochemical control of the

DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.*, **120**, 1382-1391.

28. Dose, C., Farkas, M.E., Chenoweth, D.M. and Dervan, P.B. (2008) Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J. Am. Chem. Soc.*, **130**, 6859-6866.
29. 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.
30. 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.
31. 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.
32. 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.
33. 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.

34. Gerrard, L., Zhao, D.B., Clark, A.J. and Cui, W. (2005) Stably transfected human embryonic stem cell clones express OCT4-specific green fluorescent protein and maintain self-renewal and pluripotency. *Stem Cells*, **23**, 124-133.
35. 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.
36. Kopp, J.L., Ormsbee, B.D., Desler, M. and Rizzino, A. (2008) Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells*, **26**, 903-911.
37. Trauger, J.W. and Dervan, P.B. (2001) Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. *Methods in Enzymology*, **340**, 450-466.
38. Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods in Enzymology*, **65**, 499-560.
39. Iverson, B.L. and Dervan, P.B. (1987) Adenine Specific DNA Chemical Sequencing Reaction. *Nucleic Acids Res.*, **15**, 7823-7830.

## **Appendix**

### **Synthetic Experiments toward Fluorinated Heterocycles**

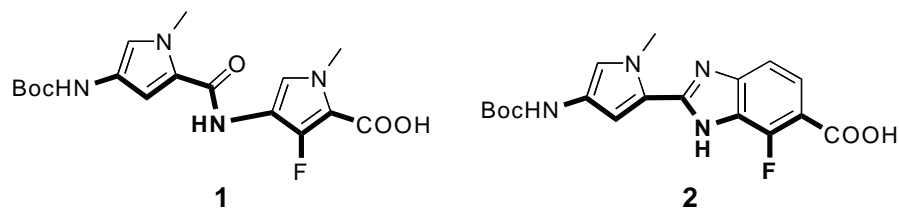
## Appendix

### Synthetic Experiments towards Fluorinated Heterocycles

#### A.1. Background

The incorporation of fluorine into pyrrole-imidazole polyamides as a recognition element in the DNA minor groove has been demonstrated at an internal position as well as in the N-terminal cap position. The pairing of  $\alpha$ -fluoro- $\beta$ -alanine side-by-side with  $\beta$ -alanine in the 2- $\beta$ -2 hairpin motif produces little (1.4-fold) specificity for T•A over A•T.<sup>1</sup> In addition, there was a 30-fold decrease in binding affinity relative to the  $\beta/\beta$  pairing.<sup>1</sup> At the N-terminal cap position, 3-fluorothiophene has been paired across from Py in an eight-ring hairpin polyamide to produce 3-fold specificity for T•A over A•T.<sup>2</sup> Synthetic efforts towards an internal *N*-methyl-3-fluoropyrrole monomer have been reported.<sup>3</sup>

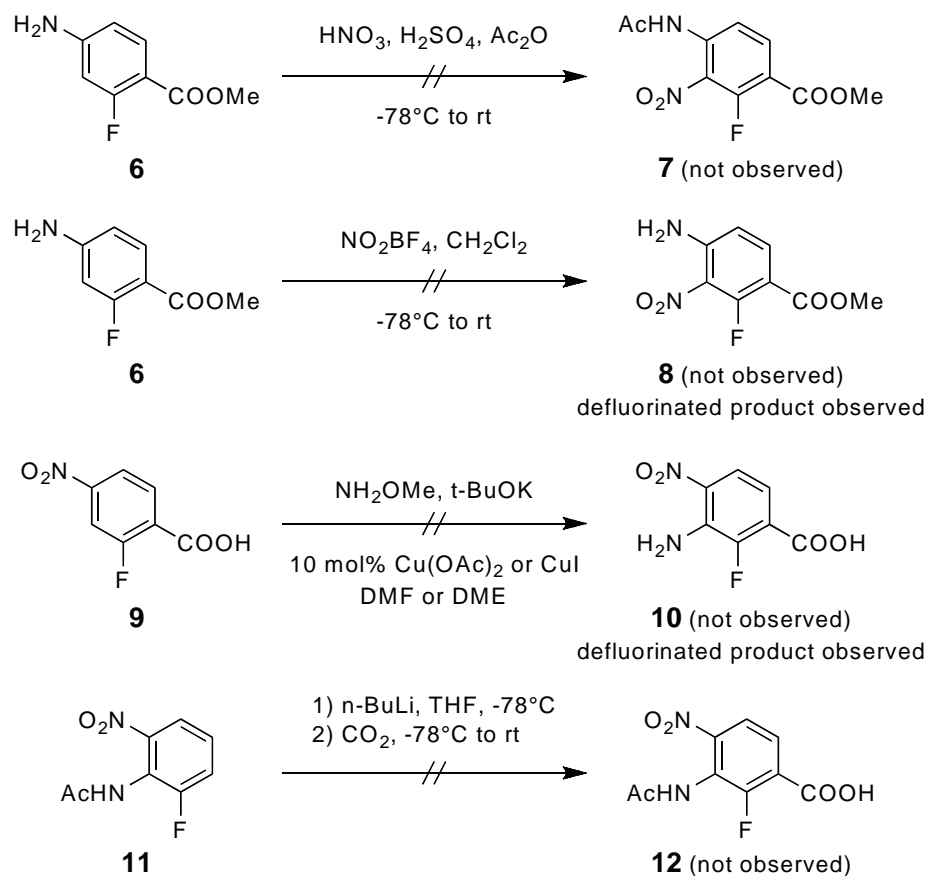
The second-generation fused heterocycles provide an additional scaffold for the incorporation of fluorine into polyamides. The fused heterocycle hydroxybenzimidazole paired with Py in an eight-ring hairpin polyamide recognizes T•A in the DNA minor groove.<sup>4</sup> The pyrrole-fluoropyrrole dimer **1** and the pyrrole-fluorobenzimidazole dimer **2** present the same atomic connectivity along the recognition surface of the polyamide (Figure A.1).<sup>4</sup> Therefore, the fluorobenzimidazole heterocycle was proposed as a synthetic target.



**Figure A.1.** The pyrrole-fluoropyrrole dimer **1** and the pyrrole-fluorobenzimidazole dimer **2** present the same atomic connectivity along the recognition edge.<sup>4</sup>



**Figure A.2.** Coupling of **3** and **4**, followed by condensation, was envisioned for the synthesis of the intermediate **5**.<sup>4</sup> Fluorodiamine **4** was the key synthetic target.



**Figure A.3.** Reactions attempted to achieve the desired tetrasubstitution of **4**.

## A.2. Synthetic efforts

It was envisioned that pyrrole-fluorobenzimidazole dimer **2** could be synthesized by an HBTU coupling, followed by condensation, of NO<sub>2</sub>-Py-COOH monomer **3** and fluorodiamine **4** to give compound **5** (Figure A.2).<sup>4</sup> The intermediate **5** could be reduced, Boc-protected, and saponified to give the Boc-protected amino acid **2** suitable for solid-phase coupling.<sup>4</sup> For this synthetic route, the fluorodiamine methyl ester **4** was the key synthetic target.

Figure A.3 illustrates reactions that were attempted to achieve the desired tetrasubstitution of synthetic target **4**. Two sets of nitration conditions were attempted on starting material **6**, which was available in three steps from commercially available 2-fluoro-4-nitrotoluene (Aldrich).<sup>5-7</sup> Nitric acid treatment did not yield nitration at the sterically hindered 3-position as in structure **7**. The use of nitronium tetrafluoroborate in an attempt to form structure **8** resulted in a defluorinated product, as no peaks were observed in the <sup>19</sup>F NMR spectrum. A copper-catalyzed direct amination with methoxylamine of starting material **9**<sup>5</sup> also resulted in a defluorinated product.<sup>8</sup> An attempt to use fluorine as a directing group in *ortho*-lithiation did not result in the synthesis of structure **12**.<sup>9</sup>



## References

1. Floreancig, P. E.; Swalley, S. E.; Trauger, J. W.; Dervan, P. B., *J. Am. Chem. Soc.* **2000**, 122, 6342-6350.
2. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., *Bioorg. Med. Chem.* **2003**, 11, 4333-4340.
3. Foister, S. Ph.D. Thesis, California Institute of Technology, Pasadena, CA, 2004.
4. Renneberg, D.; Dervan, P. B., *J. Am. Chem. Soc.* **2003**, 125, 5707-5716.
5. Schmelkes, F. C.; Rubin, M., *J. Am. Chem. Soc.* **1944**, 66, 1631-1632.
6. Cary, J. M.; Moore, J. S., *Org. Lett.* **2002**, 4, 4663-4666.
7. Soundararajan, N.; Platz, M. S., *J. Org. Chem.* **1990**, 55, 2034-2044.
8. Seko, S.; Miyake, K.; Kawamura, N., *J. Chem. Soc.-Perkin Trans. I* **1999**, 1437-1444.
9. Snieckus, V., *Chem. Rev.* **1990**, 90, 879-933.