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. Ouantitative 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 β -alanine-C₃-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

| | Pag | ze |
|----------------|---|-----------|
| Acknowledge | ments | iv |
| Abstract | | V |
| Table of Cont | ents | vi |
| List of Figure | s and Tables | <i>ii</i> |
| 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 β-Alanine-C ₃ -Linked Hairpin Polyamide Conjugates | 31 |
| Chapter 5 | Disrupting the Oct4 Octamer DNA Element with Pyrrole-Imidazole Polyamide Conjugates | 98 |
| Appendix | Synthetic Experiments toward Fluorinated Heterocycles | 39 |

List of Figures and Tables

vii

Page

Chapter 1

| 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-β-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 |

Chapter 2

| Figure 2.1 | Model for the complex formed between hairpin polyamide 24 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 (13, 14, 16, 17, 19, 21, 24, and 27) | 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 (6, 12, 20, 22, and 23) | 29 |

Chapter 3

| Figure 3.1 | Structures of polyamide-fluorescein conjugates 1 and 2 | 42 |
|------------|--|-----|
| Figure 3.2 | DNase I footprint titration experiments for polyamides 1 and 2 | 43 |
| Figure 3.3 | Overlaid fluorescence emission spectra for polyamides 1 and 2 | .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 1-10 | 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 1 , 2 , and 5 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 1 , 2 , and 5 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 1-13 | 84 |
|------------|---|----|
| Figure 4.6 | DNase I footprint titration experiments for polyamides 2 and 5 | 90 |
| Table 4.4 | Equilibrium association constants for polyamides 2-8 | 91 |
| Figure 4.7 | Nuclear localization of polyamides 2, 3, and 7 in HeLa cells | 92 |
| Figure 4.8 | Relative levels of VEGF mRNA in HeLa cells at 1 μ M polyamide | 93 |
| Figure 4.9 | Relative levels of VEGF mRNA in HeLa cells at 0.2 and 1 μ 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 1-7 |
| Figure 5.4 | Quantitative DNase I footprint titrations for polyamide 1 107 |
| Figure 5.5 | DNA melting temperature analysis for polyamides 1-5 108 |
| Figure 5.6 | Nuclear localization of polyamide-fluorescein conjugate 1 109 |
| Figure 5.7 | Time course for 48-hour incubation experiments with P19 cells |
| Figure 5.8 | Quantitative RT-PCR experiments with polyamides 2 and 6 111 |
| Figure 5.9 | Quantitative RT-PCR experiments with polyamides 2-5112 |
| Figure 5.10 | Quantitative RT-PCR experiments with polyamides 4 and 7 113 |
| Figure 5.11 | Structures of polyamide conjugates 8-14 |
| Figure 5.12 | Time course for 48-hour and 72-hour incubation experiments |
| | with R1 mouse embryonic stem cells |

| Figure 5.13 | Quantitative RT-PCR experiments with polyamides 4 and 7117 |
|-------------|--|
| Figure 5.14 | Quantitative RT-PCR experiments with polyamide 4118 |
| Figure 5.15 | Images of R1 mouse embryonic stem cells after 24-hour treatment120 |
| Figure 5.16 | Images of R1 mouse embryonic stem cells after 48-hour treatment121 |
| Figure 5.17 | Images of R1 mouse embryonic stem cells after 72-hour treatment122 |
| Figure 5.18 | The partial sequence of plasmid phOCT4-EGFP1124 |

Appendix

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