SYNTHESIS AND STRUCTURAL STUDIES OF CYCLIC Py-Im POLYAMIDES

Thesis by
David Michael Chenoweth

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

California Institute of Technology
Pasadena, California
2009
(Defended January 15, 2009)
...for Kimberly, my love...
Acknowledgments

This thesis is dedicated to my wife Kim for her constant love, support, and friendship through the years. I can’t thank her enough for her love and support, which has guided me through our time at Caltech. Our scientific collaborations and discussions have been another source of constant enjoyment for me. This thesis would not have been possible without her help and support and I am filled with excitement as I think about the arrival of our first child and our future together.

I would like to thank my Ph.D. committee Peter Dervan, Linda Hsieh-Wilson, Doug Rees, and Brian Stoltz for all the time and effort they have invested in me while at Caltech. They have all been fantastic role models and their dedication and passion for science has been truly inspiring. I would like to express my sincere gratitude to my advisor Peter B. Dervan for his guidance, support, and constant willingness to let me pursue all my crazy ideas and projects while in his group. I have truly valued his guidance and mentoring as I navigated my way through research and I thank him for the privilege to work in his research group. I sincerely thank the chair of my committee Linda Hsieh-Wilson for all of her guidance and support during graduate school. Teaching Chem 145 with her has been one of the highlights of my time at Caltech. I will miss the enjoyment of trying to make the best possible problem sets and test questions together. I think often times I learned more than the students during our endeavors. Doug Rees has been a fantastic mentor and I am grateful for his support throughout my Ph.D, his excitement about science, and his mentorship has profoundly influenced me and will continue to be an inspiration. My interactions with Doug and his research group have been one of the highlights of my graduate school experience at Caltech. I thank Brian Stoltz for serving on my committee. His passion for science, intellectual curiosity, and hard work are truly inspiring. The time he has devoted and his comments, honest criticisms, and mentoring during my proposal exams and thesis defense are greatly appreciated.

Many other professors have made an incredible impression on me while at Caltech and I thank them all. Bill Goddard has been a fantastic collaborator and mentor and his excitement about science is truly infectious. I will miss our discussions about science and collaborations on projects orthogonal to my Ph.D research. His generosity and willingness to give me access to his computers and group resources has been very valuable and is greatly appreciated. I thank Harry Gray for his mentorship and advice and for being another inspirational professor to me. I especially valued our discussions about research (fluorescent polyamides) and his willingness to share his group resources. Dennis Dougherty is thanked for his advice and research discussions (lone pair-\pi

iv
interactions and electrostatic potential map analysis). Dennis has also been an inspirational professor and his Physical Organic Chemistry textbook has served as a constant source of enjoyment to me, often occupying my night stand and keeping me reading late into the night. Jack Roberts is thanked for our many discussions about NMR and his constant willingness to talk to students and help them with problems. Just walking down the hall between Church and Crellin every day and seeing Jack intensely pounding away on his computer has been very motivating. I always knew he was thinking deeply about science and this was very comforting for me.

I have had the privilege to work with many collaborators while at Caltech and I thank them all. Dan Harki has been a great friend and collaborator and I greatly enjoyed our time at Caltech and look forward to future collaborations. I would also like to thank Christian Dose for many great times in the lab and the rest of my collaborators that I have had the privilege of working with (Justin Cohen, John Phillips, Anne Viger, Mike Marques, Mike Brochu, and Julie P.). I also thank the rest of the Dervan group for making the lab a wonderful place.

I sincerely thank Mike Day and Jens Kaiser for putting up with my constant questions and nagging day in and day out during my adventures in crystallography. I thank both of them for their guidance and friendship during my time at Caltech. I will definitely remember the Friday night plasma party with Jens in the Rees group microwave and the many fun synchrotron trips. I thank Bill Scott at UCSB for help with crystallography software setup on my mac and the Chimera team at UCSF for their help with new software features and script writing.

I sincerely thank all of my friends and family outside Caltech for their constant love and support through the years: Mom, Dad, Dannielle, Ed and Sandy Shuler, and Kieth Shuler, Angela Harki, Sam Harki, and Tom Britton. We really miss all of you! Angela and Sam Harki are thanked for being great friends and listening to all the complaints during Grad school. We will definitely miss the two of you (and Dan too)! We will have to find some way to visit each year.

I would especially like to thank Tom Britton for his friendship and mentoring during my time at Eli Lilly. He has made a lasting impression on me and taught me more than I could imagine.
The work presented in this thesis is focused on the molecular recognition of DNA by minor groove binding polyamides. Methods and strategies for the solution-phase synthesis of hairpin and cyclic pyrrole-imidazole polyamides are presented with optimized protocols requiring little to no chromatography. These synthetic strategies have led to the design of cyclic polyamides targeted to the androgen response element and are shown to be biologically active and cell permeable in cell culture experiments in addition their binding affinities rival that of most polyamide architectures. The structural elucidation of an α-amino-turn-linked cyclic polyamide is presented at 1.17 Å resolution providing insight into the detailed molecular recognition process and allosteric modulation responsible for the inhibition of transcription factor-DNA binding. Additionally, structural elucidation of a β-amino-turn-linked cyclic polyamide, highlighting the conformational differences compared to the α-amino-turn linked structure is presented. A structural basis for the inability of polyamides to bind dsRNA is also proposed based on biophysical, structural, and modeling data. In addition to these studies a new class of programmable oligomers targeting the DNA sequence 5′-WGGGGW-3′ were shown to inhibit DNA binding of the Nf-kB transcription factor by EMSA gel shift. Compounds synthesized in this study were found to possess unique fluorescent properties with the ability to modulate their fluorescence by binding their targeted dsDNA, leading to sequence specific fluorescent detection reagents. Efforts toward the templated-assembly of polyamides using higher-order DNA structure (NCP) are also reported and the development of a new profluorescent class of heterocycle, which has the potential to be used as a chemical reporter of ligation events is described.
Table of Contents

Dedication........................................................................................................................................... iii
Acknowledgments................................................................................................................................. iv
Abstract................................................................................................................................................ vi
Table of Contents................................................................................................................................. vii
List of Figures...................................................................................................................................... vii
List of Schemes................................................................................................................................... xii
List of Tables....................................................................................................................................... xii
Nomenclature and Symbology............................................................................................................ xxiii
Chapter 1: Introduction to Molecular Recognition of DNA ............................................................... 1
  1.1 Background and Significance....................................................................................................... 2
  1.2 Nucleic Acid Structure............................................................................................................... 3
  1.3 Molecular Recognition of DNA.................................................................................................. 5
  1.5 Scope of this work....................................................................................................................... 23
  1.6 Notes and Reference................................................................................................................... 25
Chapter 2: Solution-Phase Synthesis of Pyrrole–Imidazole Polyamides ............................................. 30
  Abstract.............................................................................................................................................. 31
  2.1 Introduction................................................................................................................................ 32
  2.2 Results and Discussion............................................................................................................... 33
  2.3 Conclusion................................................................................................................................. 37
  2.4 Experimental Section................................................................................................................ 37
    2.4.1 General................................................................................................................................. 37
    2.4.2 HCl•H₂N-Py-CO₂Me (10)....................................................................................................... 39
    2.4.3 BocHN-PyPy-CO₂Me (11).................................................................................................... 39
    2.4.4 HCl•H₂N-PyPy-CO₂Me (12).................................................................................................. 39
    2.4.5 BocHN-PyPyPy-CO₂Me (13)............................................................................................... 39
    2.4.6 HCl•H₂N-PyPyPy-CO₂Me (14)............................................................................................. 40
    2.4.7 ImPyPyPy-CO₂Me (15)........................................................................................................ 40
    2.4.8 ImPyPyPy-CO₂H (3)............................................................................................................. 40
    2.4.9 BocHN-(R)β-CbzHNγ-Im-CO₂Et (16).................................................................................. 41
    2.4.10 HCl•H₂N-(R)β-CbzHNγ-Im-CO₂Et (5)............................................................................... 42
    2.4.11 ImPyPyPy-(R)β-CbzHNγ-Im-CO₂Et (16)........................................................................... 42
    2.4.12 ImPyPyPy-(R)β-CbzHNγ-Im-CO₂H (17)............................................................................ 43
    2.4.13 ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy-CO₂Me (2)............................................................... 43
    2.4.14 ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy-CO₂H (22)............................................................... 44
    2.4.15 BocHN- (+)-BrOIPA (20)................................................................................................... 44
    2.4.16 ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy- (+)-BrOIPA (23)....................................................... 45
    2.4.17 ImPyPyPy-(R)β-H₂Nγ-ImPyPyPy- (+)-IPA (I)................................................................. 46
    2.4.18 Calculation of Molar Extinction Coefficients.................................................................... 47
  2.5 Notes and References................................................................................................................... 48
  2.6 Spectra and Supplemental Information........................................................................................ 50
Chapter 3: Cyclic Pyrrole–Imidazole Polyamides Targeted to the Androgen Response Element .......... 82
  Abstract.............................................................................................................................................. 83
  3.1 Introduction................................................................................................................................ 84
  3.2 Results and Discussion............................................................................................................... 85
    3.2.1 Solution-Phase Synthesis of Cyclic Polyamides................................................................. 85
    3.2.2 Thermal Stabilization of DNA duplexes by Polyamides..................................................... 87
    3.2.3 Biological Assay for Cell Permeability............................................................................... 89
    3.2.4 ADMET Studies of Polyamides I and 5............................................................................... 89
Chapter 5: Allosteric Modulation of DNA by Small Molecules ...................................................127

5.1 Introduction ......................................................................................................................129

5.2 Synthesis ......................................................................................................................134

5.3 Conclusion ......................................................................................................................134

5.4 Experimental Section .......................................................................................................135

5.4.1 General ....................................................................................................................135

5.4.2 UV Absorption Spectrophotometry ........................................................................135

5.4.3 Measurement of Androgen-Induced PSA mRNA ....................................................136

5.4.4 BocHN-(R)β-CbzHNγ-Im-CO₂H (8) ......................................................................136

5.4.5 BocHN-(R)β-CbzHNγ-ImPyPyPy-CO₂Me (9) .......................................................137

5.4.6 HCl•H₂N-(R)β-CbzHNγ-ImPyPyPy-CO₂Me (10) ..................................................138

5.4.7 BocHN-(R)β-CbzHNγ-ImPyPyPy-CO₂H (11) .........................................................138

5.4.8 BocHN-(R)β-CbzHNγ-ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy-CO₂Me (12) .......139

5.4.9 BocHN-(R)β-CbzHNγ-ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy-CO₂H (13) .........140

5.4.10 BocHN-(R)β-CbzHNγ-ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy-CO₂Pfp (14) ....141

5.4.11 cyclo-(ImPyPyPy-(R)β-H₂Nγ-ImPyPyPy-(R)β-H₂Nγ-) (1) .........................141

5.4.12 cyclo-(ImPyPyPy-(R)β-AcHNγ-ImPyPyPy-(R)β-H₂Nγ-) (3) and cyclo-(ImPyPyPy-(R)β-AcHNγ-ImPyPyPy-(R)β-AcHNγ-) (2) .........142

5.4.13 ImPyPyPy-(R)β-H₂Nγ-ImPyPyPy-(+)IPA (4) .......................................................143

5.4.14 ImPyPyPy-(R)β-AcHNγ-ImPyPyPy-(+)IPA (5) ......................................................143

5.5 Notes and References .......................................................................................................143

5.6 Spectra, Data Statistics, and Supplemental Information ..................................................144

Chapter 6: Structural Elucidation of a β-amino-γ-linked Cyclic Polyamide-DNA Complex and RNA Binding Studies .................................................................153
Appendix A: Next Generation Hairpin Polyamides with (R)-3,4-Diaminobutyric Acid Turn Unit... 226

Abstract .................................................................................................................................. 227
A.1 Introduction .......................................................................................................................... 228
A.2 Results and Discussion ........................................................................................................ 228
A.2.1 Thermal stabilization of DNA duplexes by hairpin polyamides ................................. 228
A.2.2 Sequence specificity at the turn position ....................................................................... 232
A.2.3 Acetylated chiral hairpin polyamides .......................................................................... 234
A.2.4 Biological assay for cell permeability ......................................................................... 235
A.3 Conclusion .......................................................................................................................... 237
A.4 Experimental ...................................................................................................................... 237
A.4.1 General ............................................................................................................................ 237
A.4.2 Synthesis of polyamides ............................................................................................... 238
A.4.3 UV Absorption Spectrophotometry ........................................................................... 239
A.4.4 Molecular Modeling ....................................................................................................... 240
A.4.5 Measurement of Androgen-Induced PSA mRNA ......................................................... 240
A.5 Notes and References ......................................................................................................... 240
A.6 Supplemental Information ................................................................................................. 243

Appendix B: Apredica ADMET Report .................................................................................. 251

Appendix C: Programmable Oligomers for Minor Groove DNA Recognition ................... 287

Abstract .................................................................................................................................. 288
C.1 Introduction .......................................................................................................................... 289
C.2 Experimental ...................................................................................................................... 290
C.2.1 Polyamide Synthesis ....................................................................................................... 290
C.3 Results ................................................................................................................................ 292
C.3.1 DNA Affinity and Sequence Specificity of Dimer Caps .............................................. 292
C.3.2 Design of a Programmable Oligomer for 5’-GTAC-3’ ................................................. 292
C.4 Discussion ............................................................................................................................ 293
C.5 Conclusion .......................................................................................................................... 297
C.6 Notes and References ......................................................................................................... 297
C.7 Supplemental Information ................................................................................................. 299
C.7.1 General ............................................................................................................................ 299
C.7.2 Heterocycle Synthesis ................................................................................................. 300
C.7.3 Polyamide Synthesis ....................................................................................................... 305
C.7.4 Deprotection of the O-Methyl-Protected Polyamides ................................................. 306
C.7.5 Oligomer 9 Synthesis ..................................................................................................... 307
C.7.6 O-Methyl Deprotection ............................................................................................... 309
C.7.7 Cleavage From Resin .................................................................................................... 309

Appendix D: Peptoid Cell Uptake Studies ............................................................................. 310

Abstract .................................................................................................................................. 311
D.1 Summary ............................................................................................................................. 312
D.2 Experimental ...................................................................................................................... 315
D.2.1 Materials ........................................................................................................................ 315
D.2.2 General bromoacetic acid addition procedure ............................................................ 316
D.2.3 General amine addition procedure ............................................................................. 316
D.2.4 Peptoid synthesis procedure (Synthesis of Resin PR-1) ........................................316
D.2.5 Procedure for first mini-PEG coupling (Synthesis of Resin PR-2) .......................317
D.2.6 Preparation of peptoid 5-FAM (1) .........................................................................318
D.2.7 Preparation of peptoid 6-FAM (2) .........................................................................318

Appendix E: Cell Uptake Studies of 4G Targeting Polyamides ...................................................320

E.1 Introduction ..................................................................................................................322
E.2 Results and Discussion ..............................................................................................322
E.3 Conclusion ..................................................................................................................334
E.4 Experimental .............................................................................................................335
E.5 Notes and References ...............................................................................................335
E.6 Spectra and Supplemental Information .....................................................................336
### List of Figures

**CHAPTER 1: Introduction to Molecular Recognition of DNA**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Figure 1.1</strong></td>
<td>Chemical structure of DNA</td>
<td>4</td>
</tr>
<tr>
<td><strong>Figure 1.2</strong></td>
<td>DNA base pairs</td>
<td>5</td>
</tr>
<tr>
<td><strong>Figure 1.3</strong></td>
<td>DNA sugar phosphate backbone</td>
<td>6</td>
</tr>
<tr>
<td><strong>Figure 1.4</strong></td>
<td>DNA polymorphs and A-form RNA</td>
<td>7</td>
</tr>
<tr>
<td><strong>Figure 1.5</strong></td>
<td>Nucleosome core particle structures</td>
<td>9</td>
</tr>
<tr>
<td><strong>Figure 1.6</strong></td>
<td>Anatomy of the DNA base pair edges</td>
<td>10</td>
</tr>
<tr>
<td><strong>Figure 1.7</strong></td>
<td>X-ray structures of DNA binding transcription factors</td>
<td>11</td>
</tr>
<tr>
<td><strong>Figure 1.8</strong></td>
<td>Atomic model of the interferon-β enhancesome</td>
<td>12</td>
</tr>
<tr>
<td><strong>Figure 1.9</strong></td>
<td>DNA-binding natural products</td>
<td>13</td>
</tr>
<tr>
<td><strong>Figure 1.10</strong></td>
<td>DNA recognition by netropsin and distamycin</td>
<td>14</td>
</tr>
<tr>
<td><strong>Figure 1.11</strong></td>
<td>Crystal structure of a 2:1 binding single strand Py/Im polyamide</td>
<td>15</td>
</tr>
<tr>
<td><strong>Figure 1.12</strong></td>
<td>Crystal structure of a 2:1 binding single strand ImHpPyPy-β-Dp polyamide</td>
<td>16</td>
</tr>
<tr>
<td><strong>Figure 1.13</strong></td>
<td>Polyamide pairing rules</td>
<td>17</td>
</tr>
<tr>
<td><strong>Figure 1.14</strong></td>
<td>Consequence of covalent attachment of two polyamide strands</td>
<td>18</td>
</tr>
<tr>
<td><strong>Figure 1.15</strong></td>
<td>Polyamide GABA-based turns</td>
<td>19</td>
</tr>
<tr>
<td><strong>Figure 1.16</strong></td>
<td>NMR models of 1:1 and 6-ring cyclic polyamide DNA complexes</td>
<td>20</td>
</tr>
<tr>
<td><strong>Figure 1.17</strong></td>
<td>Polyamide 2:1 DNA crystal structures colored by B-factor</td>
<td>21</td>
</tr>
<tr>
<td><strong>Figure 1.18</strong></td>
<td>X-ray crystal structures of polyamide-NCP complexes</td>
<td>22</td>
</tr>
<tr>
<td><strong>Figure 1.19</strong></td>
<td>Polyamide clamp bound to the NCP</td>
<td>22</td>
</tr>
<tr>
<td><strong>Figure 1.20</strong></td>
<td>Current state of macromolecular crystallography: A DNA-drug perspective</td>
<td>24</td>
</tr>
</tbody>
</table>

**CHAPTER 2: Solution-Phase Synthesis of Pyrrole-Imidazole Polyamides**

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Figure 2.1</strong></td>
<td>Structure of Py-Im hairpin polyamide 1</td>
<td>33</td>
</tr>
<tr>
<td><strong>Figure 2.2</strong></td>
<td>Retrosynthetic strategy for synthesis of polyamide 1</td>
<td>34</td>
</tr>
</tbody>
</table>
Figure 2.3 Analysis of polyamide 2 purity by analytical HPLC ........................................37
Figure 2.4 UV properties of polyamide 1 .................................................................38
Figure 2.5 $^1$H NMR of HCl•H$_2$N-Py-CO$_2$Me (10) ..................................................50
Figure 2.6 $^{13}$C NMR of HCl•H$_2$N-Py-CO$_2$Me (10) ..................................................51
Figure 2.7 $^1$H NMR of BocHN-PyPy-CO$_2$Me (11) .....................................................52
Figure 2.8 $^{13}$C NMR of BocHN-PyPy-CO$_2$Me (11) ......................................................53
Figure 2.9 $^1$H NMR of HCl•H$_2$N-PyPy-CO$_2$Me (12) ..................................................54
Figure 2.10 $^{13}$C NMR of HCl•H$_2$N-PyPy-CO$_2$Me (12) ..................................................55
Figure 2.11 $^1$H NMR of BocHN-PyPyPy-CO$_2$Me (13) ..................................................56
Figure 2.12 $^{13}$C NMR of BocHN-PyPyPy-CO$_2$Me (13) ..................................................57
Figure 2.13 $^1$H NMR of HCl•H$_2$N-PyPyPy-CO$_2$Me (4) ...............................................58
Figure 2.14 $^{13}$C NMR of HCl•H$_2$N-PyPyPy-CO$_2$Me (4) ...............................................59
Figure 2.15 $^1$H NMR of ImPyPyPy-CO$_2$Me (14) ..........................................................60
Figure 2.16 $^{13}$C NMR of ImPyPyPy-CO$_2$Me (14) ..........................................................61
Figure 2.17 $^1$H NMR of ImPyPyPy-CO$_2$H (3) ..............................................................62
Figure 2.18 $^{13}$C NMR of ImPyPyPy-CO$_2$H (3) ..............................................................63
Figure 2.19 $^1$H NMR of BocHN-($R$)$^β$-CbzHN$γ$-Im-CO$_2$Et (15) ..................................64
Figure 2.20 $^{13}$C NMR of BocHN-($R$)$^β$-CbzHN$γ$-Im-CO$_2$Et (15) ..................................65
Figure 2.21 $^1$H NMR of HCl•H$_2$N-($R$)$^β$-CbzHN$γ$-Im-CO$_2$Et (5) .................................66
Figure 2.22 $^{13}$C NMR of HCl•H$_2$N-($R$)$^β$-CbzHN$γ$-Im-CO$_2$Et (5) .................................67
Figure 2.23 $^1$H NMR of ImPyPyPy-($R$)$^β$-CbzHN$γ$-Im-CO$_2$Et (16) ............................68
Figure 2.24 $^{13}$C NMR of ImPyPyPy-($R$)$^β$-CbzHN$γ$-Im-CO$_2$Et (16) ............................69
Figure 2.25 $^1$H NMR of ImPyPyPy-($R$)$^β$-CbzHN$γ$-Im-CO$_2$H (17) .............................70
Figure 2.26 $^{13}$C NMR of ImPyPyPy-($R$)$^β$-CbzHN$γ$-Im-CO$_2$H (17) .............................71
Figure 2.27 $^1$H NMR of ImPyPyPy-($R$)$^β$-CbzHN$γ$-ImPyPyPy-CO$_2$Me (2) ..................72
Figure 2.28 13C NMR of ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-CO₂Me (2) ......................................73
Figure 2.29 1H NMR of ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-CO₂H (22) ......................................74
Figure 2.30 13C NMR of ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-CO₂H (22) ......................................75
Figure 2.31 1H NMR of BocHN-(+)-BnOIPA (20) .......................................................................76
Figure 2.32 13C NMR of BocHN-(+)-BnOIPA (20) .......................................................................77
Figure 2.33 1H NMR of ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-(+) BnOIPA (23) .........................78
Figure 2.34 13C NMR of ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-(+) BnOIPA (23) .........................79
Figure 2.35 1H NMR of ImPyPyPy-(R)β-HZNγ-ImPyPyPy-(+) IPA (1) .......................................80
Figure 2.36 13C NMR of ImPyPyPy-(R)β-HZNγ-ImPyPyPy-(+) IPA (1) .......................................81

CHAPTER 3: Cyclic Pyrrole–Imidazole Polyamides Targeted to the Androgen Response Element

Figure 3.1 Structures of cyclic and hairpin polyamides 1–5 ...........................................................84
Figure 3.2 Targeting the ARE with DNA-binding polyamides .....................................................88
Figure 3.3 Polyamides 1 and 5 ADMET testing ............................................................................99
Figure 3.4 1H NMR BocHN-(R)β-ChzHNγ-Im-CO₂H (8) ............................................................101
Figure 3.5 13C NMR BocHN-(R)β-ChzHNγ-Im-CO₂H (8) ............................................................102
Figure 3.6 1H NMR BocHN-(R)β-ChzHNγ-ImPyPyPy-CO₂Me (9) .............................................103
Figure 3.7 13C NMR BocHN-(R)β-ChzHNγ-ImPyPyPy-CO₂Me (9) .............................................104
Figure 3.8 1H NMR HCl•H₂N-(R)β-ChzHNγ-ImPyPyPy-CO₂Me (10) ...........................................105
Figure 3.9 13C NMR HCl•H₂N-(R)β-ChzHNγ-ImPyPyPy-CO₂Me (10) ...........................................106
Figure 3.10 1H NMR BocHN-(R)β-ChzHNγ-ImPyPyPy-CO₂H (11) .............................................107
Figure 3.11 13C NMR BocHN-(R)β-ChzHNγ-ImPyPyPy-CO₂H (11) .............................................108
Figure 3.12 1H NMR BocHN-(R)β-ChzHNγ-ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-CO₂Me (12) ....109
Figure 3.13 13C NMR BocHN-(R)β-ChzHNγ-ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-CO₂Me (12) ...110
Figure 3.14 1H NMR BocHN-(R)β-ChzHNγ-ImPyPyPy-(R)β-ChzHNγ-ImPyPyPy-CO₂H (13) ....111
Figure 3.15 $^{13}$C NMR BocHN-$(R)\beta$-ChdHN$\gamma$-ImPyPyPy$-(R)\beta$-ChdHN$\gamma$-ImPyPyPy-CO$_2$H (13) ...

Figure 3.16 $^1$H NMR BocHN-$(R)\beta$-ChdHN$\gamma$-ImPyPyPy$-(R)\beta$-ChdHN$\gamma$-ImPyPyPy-CO$_2$Pfp (14) .

Figure 3.17 $^1$H NMR cyclo-$(R)\beta$-ChdHN$\gamma$-ImPyPyPy$-(R)\beta$-ChdHN$\gamma$-ImPyPyPy-CO$_2$ (1) .................

Figure 3.18 Analytical HPLC characterization of cyclic polyamide 15 ....................

Figure 3.19 Analytical HPLC characterization of cyclic polyamide 1 ..................

Figure 3.20 Analytical HPLC characterization of cyclic polyamide 3 ..................

Figure 3.21 Analytical HPLC characterization of cyclic polyamide 2 ..................

Figure 3.22 Analytical HPLC characterization of cyclic polyamide 5 ..................

CHAPTER 4: Oligomerization Route to Polyamide Macrocycles

Figure 4.1 Structures of macrocyclic polyamides 1z–3z and 1–3 .......................

Figure 4.2 Synthesis of macrocyclic polyamides 1–3 ...............................

Figure 4.3 Reverse phase HPLC analysis of the oligomerization reaction .............

Figure 4.4 $^1$H NMR BocHN-$(R)\beta$-ChdHN$\gamma$-ImPyPyPy-CO$_2$Pfp (4) ................

CHAPTER 5: Allosteric Modulation of DNA by Small Molecules

Figure 5.1 Chemical structure of the cyclic polyamide 1 ...............................

Figure 5.2 Comparison of native DNA to polyamide/DNA complex ..................

Figure 5.3 Analysis of groove width for native DNA and polyamide complexed DNA ....

Figure 5.4 Conformation of the $\alpha$-amino substituted GABA turn .....................

Figure 5.5 Direct and water-mediated non-covalent molecular recognition interactions ... 

Figure 5.6 Solid-phase synthesis of cyclic polyamide 1 ...............................

Figure 5.7 Polyamide analytical data ..............................................................

Figure 5.8 Single crystal of the polyamide-DNA complex and UV-Vis ..................

Figure 5.9 Hydrogen bond map of polyamide-DNA complex ...........................

Figure 5.10 Polyamide/DNA complex map .....................................................
Figure 5.11 DNA structure map ........................................................................................................147
Figure 5.12 Hydrogen bond map of DNA crystal structure ..........................................................148
Figure 5.13 Molecular electrostatic potential maps of compounds 7-10 and complex 11 ..149
Figure 5.14 Comparison of Local base-pair step parameters .....................................................150
Figure 5.15 Comparison of Local base-pair parameters ............................................................151
Figure 5.16 Comparison of Local base-pair helical parameters ................................................152

CHAPTER 6: Structural Elucidation of a β-amino-γ-linked Cyclic Polyamide-DNA Complex and RNA Binding Studies

Figure 6.1 Structure of cyclic polyamide–DNA complex at 0.95 Å resolution .........................156
Figure 6.2 Molecular recognition details from the X-ray structure ........................................158
Figure 6.3 DNA minor and major groove dimensions .............................................................159
Figure 6.4 Allosteric distortion upon polyamide binding ........................................................160
Figure 6.5 Polyamide–RNA binding ......................................................................................162
Figure 6.6 Comparison of Local base-pair step parameters ..................................................168
Figure 6.7 Comparison of Local base-pair parameters ..........................................................169
Figure 6.8 Comparison of Local base-pair helical parameters ................................................170

CHAPTER 7: Programmable Oligomers Targeting 5’-GGGG-3’ in the Minor Groove of DNA and NF-κB Binding Inhibition

Figure 7.1 Structures of pyrrole-benzimidazole dimers ............................................................174
Figure 7.2 Crystal structure of the NF-κB heterodimer bound to the DNA ...............................174
Figure 7.3 Postulated hydrogen-bonding models for the polyamide–DNA complexes .....176
Figure 7.4 Illustration of NF-κB:DNA binding inhibition by oligomer 2 .................................177
Figure 7.5 DNaseI footprinting for polyamides 1 and 2 ............................................................180
Figure 7.6 DNaseI footprinting for polyamides 3 and 4 ............................................................181
Figure 7.7 DNaseI footprinting for polyamides 5 and 6 ............................................................182
Figure 7.8 Gel shift screen for compounds 1–6 .................................................................184
Figure 7.9 EC₅₀ value for compound 2 from gel shift experiment .................................185
Figure 7.10 Antibody supershift on match DNA with anti-p50 or anti-p65 antibody ...185

CHAPTER 8: Fluorescent Sequence-Specific dsDNA Binding Oligomers

Figure 8.1 Structure of oligomers O1 and O2 .................................................................200
Figure 8.2 Design of dsDNA library ..............................................................................201
Figure 8.3 Fluorescence emission spectra of O1 and O2 ..............................................202
Figure 8.4 Plot of dsDNA concentration versus fluorescence .....................................203
Figure 8.5 UV spectra for compounds O1 and O2 .......................................................205
Figure 8.6 Fluorescence emission as a function of DNA concentration .......................206
Figure 8.7 Fluorescence emission spectra compound O1 ..........................................207
Figure 8.8 Fluorescence emission spectra compound O2 ..........................................208

CHAPTER 9: Polyamide/NCP Ligation and Profluorescent Azido-Carbostyrils

Figure 9.1 Crystal structure of an NCP bound polyamide clamp ................................211
Figure 9.2 Illustration of the clamp dimerization strategy ........................................212
Figure 9.3 Sequence of the 146 base pair fragment of α-satellite DNA .......................213
Figure 9.4 Analysis of the linker dependence ...............................................................214
Figure 9.5 Synthesis of alkyl azido linkers .................................................................214
Figure 9.6 Synthesis of azido-polyamides .................................................................215
Figure 9.7 Synthesis of alkynyl-polyamides .............................................................216
Figure 9.8 DNA templated ligation on the NCP .........................................................217
Figure 9.9 Control reactions for the NCP templated ligation ......................................218
Figure 9.10 Linker distances for incorporation of the profluorescent coumarin ..........219
Figure 9.11 Carbostyril linker distance ......................................................................220
Figure 9.12 Synthesis of profluorescent azide 17 .................................................................221
Figure 9.13 Photophysical properties of compound 17 ..........................................................222
Figure 9.14 Representative gels of the NCP reconstitution ....................................................224

APPENDIX A: Next Generation Hairpin Polyamides with (R)-3,4-Diaminobutyric Acid Turn Unit

Figure A.1 Increased DNA-binding affinity caused by different γ-turn units ..................229
Figure A.2 Chemical structures for hairpins 1-16 ..............................................................230
Figure A.3 Normalized UV denaturation profiles ...............................................................231
Figure A.4 Models of different turn conformations ............................................................233
Figure A.5 Chemical structures polyamides 17-19 ............................................................234
Figure A.6 Androgen receptor (AR)-mediated transcription .........................................235
Figure A.7 Chemical structures polyamides 20-23 ..........................................................236
Figure A.8 DNase I footprint titration experiments (polyamides 2–4) ...........................245
Figure A.9 DNase I footprint titration experiments (polyamides 6–8) ............................246
Figure A.10 DNase I footprint titration experiments (polyamides 10–12) ......................247
Figure A.11 DNase I footprint titration experiments (polyamides 14–16) ......................248
Figure A.12 A) Chemical structures of polyamides 20-23 ..............................................250

APPENDIX B: Apredica ADMET Report

Figure B.1 Polyamides 1 and 5 ..........................................................................................252

APPENDIX C: Programmable Oligomers for Minor Groove DNA Recognition

Figure C.1 Structures of dimers ........................................................................................290
Figure C.2 Postulated hydrogen-bonding models for polyamide-DNA complexes ..........291
Figure C.3 Quantitative DNase I footprinting experiments for polyamides 1, 2, and 3 ....294
Figure C.4 Postulated hydrogen-bonding model and structure of oligomer 9 .................296
APPENDIX D: Peptoid Cell Uptake Studies

Figure D.1 Compound 1, 5-FAM, and 2, 6-FAM ...............................................................312
Figure D.2 Synthesis of compound 1 and 2 .................................................................313
Figure D.3 HeLa cell uptake studies for compound 1 ................................................314
Figure D.4 HeLa cell uptake studies for compound 2 ................................................315

APPENDIX E: Cell Uptake Studies of 4G Targeting Polyamides

Figure E.1 Fluorescent polyamide 2 ..............................................................................322
Figure E.2 Tail and turn modifications for 23 compound polyamide library. .........323
Figure E.3 Fluorescent polyamide cell uptake studies (Compounds 1-4) ..........326
Figure E.4 Fluorescent polyamide cell uptake studies (Compounds 5-8) ..........327
Figure E.5 Fluorescent polyamide cell uptake studies (Compounds 9-11) ....327
Figure E.6 Fluorescent polyamide cell uptake studies (Compounds 12-15) ..328
Figure E.7 Fluorescent polyamide cell uptake studies (Compounds 16-19) ..329
Figure E.8 Fluorescent polyamide cell uptake studies (Compounds 20-23) ..330
Figure E.9 Mechanism of mitochondrial staining using MitoTracker ..........331
Figure E.10 Results of HeLa cell uptake study using MitoTracker and compound 4 ...331
Figure E.11 Results of HeLa cell uptake study using Lysotracker and compound 4 ...332
Figure E.12 Cell uptake with (±)-verapamil (Compound 23) ................................333
Figure E.13 Cell uptake study with Lipofectamine 2000 (Compound 23) ..........334
List of Schemes

CHAPTER 2: Solution-Phase Synthesis of Pyrrole–Imidazole Polyamides

**Scheme 2.1** Preparation of 3 and 4 ................................................................. 35

**Scheme 2.2** Preparation of 5 and assembly of core polyamide 2 .................. 36

**Scheme 2.3** Preparation of 21 .......................................................................... 37

**Scheme 2.4** Final steps for the synthesis of Py-Im polyamide 1 ...................... 37

CHAPTER 3: Cyclic Pyrrole–Imidazole Polyamides Targeted to the Androgen Response Element

**Scheme 3.1** Preparation of 10 and 11 .............................................................. 85

**Scheme 3.2** Preparation of 1, 2, and 3 ............................................................ 86

CHAPTER 7: Programmable Oligomers Targeting 5’-GGGG-3’ in the Minor Groove of DNA and NF-κB Binding Inhibition

**Scheme 7.1**. Synthesis of imidazopyridine–imidazole dimers .......................... 178

**Scheme 7.2**. Representative solid-phase synthesis of polyamide 6 ............... 179

APPENDIX C: Programmable Oligomers for Minor Groove DNA Recognition

**Scheme C.1** Representative solid-phase synthesis of polyamide 2 and 3 .... 293

**Scheme C.2** Synthesis of 15 .............................................................................. 301

**Scheme C.3** Synthesis of 7 ................................................................................ 302

**Scheme C.4** Synthesis of 8 ................................................................................ 303

**Scheme C.5** Synthesis of 1 ................................................................................ 304

**Scheme C.6** Compound 2 ................................................................................ 305

**Scheme C.7** Compound 3 ................................................................................ 305

**Scheme C.8** Synthesis of 9-13 ......................................................................... 307

APPENDIX E: Cell Uptake Studies of 4G Targeting Polyamides

**Scheme E.1**. Synthesis of R4 ........................................................................... 324

**Scheme E.2**. Synthesis of compound library .................................................. 325
List of Tables

CHAPTER 1: Introduction to Molecular Recognition of DNA

Table 1.1 Typical nucleic acid structural parameters. ..............................................................8

CHAPTER 3: Cyclic Pyrrole–Imidazole Polyamides Targeted to the Androgen Response Element

Table 3.1 $T_m$ values for polyamides for 1–5 .................................................................87
Table 3.2 Caco-2 permeability summary .................................................................99
Table 3.3 Cytotoxicity summary .............................................................................99
Table 3.4 Fluorescent Cyp IC50 summary ..............................................................100
Table 3.5 hERG FastPatch summary ......................................................................100
Table 3.6 Plasma half-life summary ......................................................................100
Table 3.7 Plasma protein binding summary ..........................................................100

CHAPTER 4: Oligomerization Route to Polyamide Macrocycles

Table 4.1 $T_m$ values for cycles 1–3 in the presence of DNA..................................122

CHAPTER 5: Allosteric Modulation of DNA by Small Molecules

Table 5.1 Data collection and refinement statistics. ........................................143

CHAPTER 6: Structural Elucidation of a β-amino-γ-linked Cyclic Polyamide-DNA Complex and RNA Binding Studies

Table 6.1 Data collection and refinement statistics. ........................................157
Table 6.2 Buckle and opening values ...............................................................161
Table 6.3 Polyamide-DNA and Polyamide-RNA melting temperatures ............162

CHAPTER 7: Programmable Oligomers Targeting 5'-GGGG-3' in the Minor Groove of DNA and NF-κB Binding Inhibition

Table 7.1 Affinities of 5'-GGGG-3' binding oligomers. .........................................183
APPENDIX A: Next Generation Hairpin Polyamides with (R)-3,4-Diaminobutyric Acid Turn Unit

Table A.1 Melting temperatures of DNA/polyamide complexes ........................................232
Table A.2 Melting temperatures of DNA/polyamide complexes .....................................235
Table A.3 Melting temperatures of DNA/polyamide complexes .....................................243
Table A.4 Equilibrium association constants for hairpin polyamides ..............................244
Table A.5 Melting temperatures of polyamides .............................................................249

Appendix C: Programmable Oligomers for Minor Groove DNA Recognition ......................287

Table C.1 Affinities of X/Py ring pairs ........................................................................295
Nomenclature and Symbology

A  adenine
Å  angstrom
A•T adenine Watson-Crick hydrogen bonded to thymine
Ac₂O acetic anhydride
ADMET absorption, distribution, metabolism, excretion, and toxicity
AP activating protein
AR androgen receptor
ARE androgen response element
atm atmosphere
Bi benzimidazole
Boc tert-butyloxycarbonyl
bp base pair
°C degrees Celsius
C cytosine
calc’d calculated
Cbz carbobenzyloxy
CCDC Cambridge Crystallographic Data Centre
Ct 2-carboxy-3-chlorothiophene
Dbu diazabicycloundecane
DCM dichloromethane
DHT dihydrotestosterone
DIEA N,N-diisopropylethylamine
DMF N,N-dimethylformamide
DMSO dimethylsulfoxide
DNA deoxyribonucleic acid
Dp N,N-dimethylaminopropylamine
ds double strand
Em emission
ESI electrospray ionization
Et ethyl
Ex excitation
FAB fast-atom bombardment
Fmoc fluorenylmethyloxycarbonyl
G guanine
g grams
G•C guanine Watson-Crick hydrogen bonded to cytosine
GABA gamma-aminobutyric acid
h hour(s)
HBTU 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HF hartree fock
Hp 3-hydroxypyrrole
HPLC high performance liquid chromatography
HRMS high resolution mass spectrometry
hv  light
Hz  hydroxybenzimidazole
IC$_{50}$  median inhibition concentration (50%)
Im  N-methylimidazole
Ip  imidazopyridine
K$_a$  association constant
K$_d$  dissociation constant
$\lambda$  wavelength
m/z  mass to charge ratio
$\mu$  micro
M  molar
max  maximum
MALDI  Matrix-assisted LASER desorption/ionization
min  minute(s)
mol  mole(s)
mmol  millimole(s)
MS  mass spectrometry
N  normal
N  A, T, G, or C
NCP  nucleosome core particle
No  oxazole
NOESY  nuclear Overhauser enhancement spectroscopy
PCR  polymerase chain reaction
Py-Im  pyrrole-imidazole
PNA  peptide nucleic acid
PSA  prostate specific antigen
Py-Im  pyrrole-imidazole
RT-PCR  reverse transcriptase PCR
Py  N-methylpyrrole
PyBOP  (benzotriazol-1-yloxy)trityrrolidinophosphonium hexafluorophosphate
OBt  hydroxytriazole ester
$R_f$  retention factor
RNA  ribonucleic acid
RP-HPLC  reverse-phase high performance liquid chromatography
sat.  saturated
satd.  saturated
ss  single strand
T  thymine
TFA  trifluoroacetic acid
TMR  tetramethyl rhodamine
TO  thiazole orange
TOF  time-of-flight
U  uracil
UV  ultraviolet
Vis  visible
W  adenine or thymine
Chapter 1: Introduction to Molecular Recognition of DNA
1.1 Background and Significance

The field of molecular recognition has come a long way since the organic solvent based host-guest chemistry of Lehn and Cram (crown ether cation complexes). Understanding in a predictive and mechanistic sense the molecular recognition between synthetic ligands and biological macromolecules in water is fundamental to understanding biochemical processes and cellular composition. The overall free energy of these complexes often includes a superposition of non-covalent forces such as hydrogen bonding interactions, dipole-dipole, induced dipole, cation-π, lone pair-π, and van der Waals interactions in addition to hydrophobic effects. Understanding the intimate interplay of these forces and their contributions to the overall free energy of a host-guest system has remained one of the ultimate challenges in chemistry and biology. The molecular recognition processes involved in nucleic acid-drug and nucleic acid-protein interactions are similar with both being driven by the hydrophobic effect, a phenomenon which is still not well understood. As ligand-receptor recognition proceeds, the optimization of multiple forces ensues including minimization of water exposed hydrophobic surfaces and simultaneous maximization of van der Waals interactions. Additionally, optimization of all buried hydrogen-bond donor and acceptor pairings including solvent-assisted and counterion charge neutralization contribute to the complex recognition event. Intimate structural and biophysical knowledge of these processes is fundamental to the understanding of nature at the molecular level.

The DNA double helix, in addition to being the molecular storage unit of genetic information, represents one of the ultimate challenges in aqueous based molecular recognition. Over billions of years, nature has used selection to evolve protein surfaces that recognize DNA in a cooperative and combinatorial fashion allowing for the stringent regulation of the molecular processes crucial to all living organisms on earth. Prior to the 1960s, histologists and cell biologists realized that certain small molecules could interact specifically with cell nuclei. Dye molecules such as aminoacridines were regularly used for staining tissues and cells and it was recognized that specificity for different nucleic acid structures could be obtained using different dyes. However it was not until the 1960s that a formal DNA drug binding hypothesis would be formulated. The “intercalation hypothesis” formulated by Leonard Lerman (a graduate student of Linus Pauling at Caltech) in 1961, working at the Cambridge MRC laboratory, provided the pivotal turning point in the field of drug-nucleic acid recognition. Since the intercalation hypothesis, a plethora of biophysical, biochemical, and biostructural investigations have unveiled the detailed chemistry and biology of many DNA binding drugs, some of which have had a profound impact on human disease (i.e. actinomycin D). The
intercalating natural product actinomycin D remained one of the most potent chemotherapeutics throughout the 1950s and 1960s along with other nucleic acid binding drugs including cross-linking agents and powerful alkylators, however the first minor groove binding agents would not be discovered until the mid 1960s.\(^5\) Even though Lerman himself relied upon X-ray fiber diffraction data for his intercalation hypothesis in the 1960s, it took another 15-20 years before the first single crystal X-ray structures of drug-nucleic acid complexes (intercalators) would emerge with the seminal work of Sobell, Rich, and Neidle.\(^9,11\) The first X-ray structure of a minor groove binder would not appear until Dickerson’s report on the 1:1 structure of netropsin complexed with DNA in 1985.\(^12\) This was soon followed by the 1:1 structure of the distamycin-DNA complex by Rich in 1987.\(^13\) In a seminal study, structural evidence using NMR for the 2:1 binding motif of distamycin was provided by Wemmer in 1989,\(^14\) however the first single crystal X-ray structure of a 2:1 minor groove binding ligand-DNA complex was not realized until the work of Ramakrishnan in 1994 on distamycin A.\(^15\) Since this work many advances have been made in the field of DNA molecular recognition, with minor groove binders representing one of the most promising classes of DNA-binding molecules for targeted transcriptional therapy.

The modulation of gene expression using small molecules has been one of the ultimate goals of nucleic acid molecular recognition. Complex natural products such as actinomycin D, netropsin, and distamycin A have served as inspiration to chemists for the construction of molecular architectures capable of nucleic acid recognition with specificity and affinities equivalent to and rivaling that of endogenous proteins. Initially inspired by the 2:1 binding natural product distamycin, pyrrole-imidazole polyamides have evolved into a modular programmable molecular recognition system capable of specificities and affinities rivaling that of endogenous transcription factors.\(^16,17\) Modulation of transcription factor-DNA interfaces with small molecules such as pyrrole-imidazole polyamides provides a powerful strategy for controlling regulation of the genetic material and could eventually impact human medicine. The future of molecular recognition is poised to benefit greatly from advances in biochemical, biophysical, computational, and structural (X-ray, NMR, EM, Cryo-EM, etc.) methods along with the new tools of physical biology leading to ever increasing resolution and a quantitative understanding of molecular level processes.\(^4,18\)

1.2 Nucleic Acid Structure

Deoxyribose nucleic acid (DNA) is the fundamental storage material of genetic information and can be characterized chemically as a hetero-polymer consisting of nucleotide monomers linked
through their sugar-phosphate backbones.\textsuperscript{19,20} The 5’ and 3’ hydroxyl groups of the deoxyribose sugar define the directionality of the DNA strand while a set of four nucleobases [adenine (A), guanine (G), cytosine (C), and thymine (T)] distinguish the nucleotide monomers, providing the fundamental building blocks of the genetic code. Figure 1.1 shows the chemical structure of a short DNA strand containing all four bases and Figure 1.2 shows the atom numbering conventions. Early studies by Chargaff demonstrated that A and T occurred in similar molar ratios as did G and C, which in combination with fiber diffraction data from Rosalind Franklin and Maurice Wilkins would eventually lead to Watson and Crick’s base-paired helical model of B-DNA.\textsuperscript{21} The Watson-Crick base paired model of DNA contains a set of rules for which A prefers to bind T through two hydrogen bonds and G prefers to bind C through three hydrogen bonds on opposite strands as shown in Figure 1.2 (U replaces T in RNA).\textsuperscript{22} The strands are oriented in an antiparallel fashion as they base pair and wind around a central axis. These opposite strands form a double helical structure where the Watson-Crick base pairs are stacked and stabilized by a combination of favorable hydrophobic effects and hydrogen bonding between paired bases. Due to the length of the sugar-phosphate backbone, a helical twist is required to minimize the distance between adjacent base pairs and maximize their hydrophobic stacking.\textsuperscript{19,20}

The sugar-phosphate backbone of DNA is highly dynamic allowing for a diverse range of higher order structures depending on environmental conditions. The torsion angles for the sugar-phosphate backbone are defined in Figure 1.3 and typically vary with ionic strength, pH, sequence, and many other factors.\textsuperscript{19,20} In contrast to RNA, where the 2’-hydroxyl of the sugar locks the A-form helix into a fairly rigid structure, the sugar-phosphate backbone of DNA is highly mobile.\textsuperscript{19,20} DNA conformation can often be defined by the sugar puckering modes, which by convention are named after the ring atom and either endo or exo referring to the 5’ side of the furanose ring or the 3’ side, respectively. Figure 1.3 shows typical

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{dna_structure}
\caption{Chemical structure of DNA.}
\end{figure}
sugar pucker conformations along with their preference in nucleic acid structures. In addition to the sugars displaying conformational preferences, the phosphodiester bond exhibits conformational rigidity analogous to a peptide bond. This conformation rigidity, known as the gauche effect, is a result of stereoelectronic effects from lone pair hyperconjugation/donation of the O3' and O5' oxygen atoms into the σ* orbital of the P-O5' and P-O3' bonds, respectively.23 Double helical DNA is a dynamic structure which is capable of forming three primary double strand conformations known as A, B, and Z forms. In contrast to this, double helical RNA is far less flexible with the 2'-OH locking its sugar ring conformation into a C3'-endo pucker resulting in a preference for an A form helix similar to that of A-form DNA. A structural comparison of these ideal DNA polymorphs along with A-form RNA is shown in Figure 1.4 and Table 1.1.19,20,23,24.

In biological systems, especially eukaryotic cells, DNA is assembled around octameric proteins called histones and compacted into macrohelical fibers forming the high-order structure of chromatin. This DNA-histone complex is called the nucleosome core particle (NCP) and represents the fundamental repeating unit of chromatin consisting of 147 base pairs of DNA forming two super helical turns around the histone octomer with 20-80 base pairs of linker DNA separating one NCP from the next. The Richmond group25-27 at ETH Zurich has made seminal contributions to elucidate biologically relevant higher-order DNA structures such as the NCP25,26 and the tetra-NCP27 presented Figure 1.5. In addition, a theoretical model of four tetra-NCPs assembled into a super-helical chromatin fiber is presented in Figure 1.5. Chromatin architecture and accessibility in biological systems represents a higher-order level of regulation and a profoundly important problem for the field of DNA recognition.

1.3 Molecular Recognition of DNA

One of the largest projects in modern science, the human genome project,28-30 is poised to deliver detailed information and make major impacts in biotechnology and medicine through the physical and functional characterization of the approximately 20,000 to 25,000 genes in the human
Figure 1.3 DNA sugar phosphate backbone torsion angle map and sugar puckering conventions. 
a) Chemical structure of DNA sugar phosphate backbone with torsion angles next to the 3-dimensional structure of a nucleoside taken from the native high resolution B-DNA structure solved in Chapter 4 of this thesis. b) Pseudorotation phase angle \((P)\) diagram defining 5-membered ring sugar puckering modes. Equation describing the pseudorotation phase angle and maximum torsion angle for 5-membered rings. c) Chemical structure of the most common sugar puckering modes for B-DNA (C2'-endo) and A-DNA/A-RNA (C3'-endo).

\[ \nu_j = \Phi_{m} \cos[ P + 144^\circ (j - 2)], \; (j = 0, 1, 2, 3, 4) \]
Figure 1.4 A comparison of double helical DNA polymorphs and A-form RNA.

These genes are tightly regulated in higher organisms by transcription factor assemblies that function in a concerted cooperative and combinatorial fashion to modulate eukaryotic gene expression. The molecular recognition processes involved in nucleic acid-protein interactions are completely analogous to those of nucleic acid-drug interactions where initial complexation is often driven by the hydrophobic effect. Optimization of the same forces is also required, involving minimization of water exposed hydrophobic surfaces and maximization of van der Waals interactions in conjunction with the optimization of all buried hydrogen bond donor and acceptor pairings (solvent-assisted or counterion charge neutralization). The recognition of the B-DNA interface by proteins and small molecules can occur at the major groove, minor groove, and phosphate backbone, or any combination, with interactions mediated through electrostatics, hydrogen bonding, and van der Waals interactions along with base pair stacking for the case of intercalators. The DNA base pair edges in the major groove and minor groove provide an array
of functionality for hydrogen bonding, hydrophobic interaction, and steric complementarity with proteins and small molecule binders. The molecule electrostatic potential surfaces for the minor and major groove base pair edges are shown in Figure 1. In addition, primary driving forces such as the hydrophobic effect and shape complementarity are common to both proteins and small molecules.

The regulation of gene transcription is controlled by the sequence specific cooperative assembly of transcription factors, which form regulatory switches and networks in the cell providing stringent control over biochemical processes. The minor groove and major groove of

---

**Table 1.1** Typical nucleic acid structural parameters.

<table>
<thead>
<tr>
<th></th>
<th>A-DNA</th>
<th>B-DNA</th>
<th>Z-DNA</th>
<th>A-RNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helix Sense</td>
<td>Right</td>
<td>Right</td>
<td>Left</td>
<td>Right</td>
</tr>
<tr>
<td>Screw symmetry</td>
<td>11-fold</td>
<td>10-fold</td>
<td>6-fold</td>
<td>12-fold</td>
</tr>
<tr>
<td>bp/repeating unit</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>bp/turn</td>
<td>11</td>
<td>10</td>
<td>12</td>
<td>11-12</td>
</tr>
<tr>
<td>Helix twist (°)</td>
<td>32.7°</td>
<td>36.0°</td>
<td>-10°, -50°</td>
<td>33.1°</td>
</tr>
<tr>
<td>Rise (Å)</td>
<td>2.8</td>
<td>3.4</td>
<td>-3.9°, -3.5°</td>
<td>2.8</td>
</tr>
<tr>
<td>Helix pitch (Å)</td>
<td>28</td>
<td>34</td>
<td>44</td>
<td>36</td>
</tr>
<tr>
<td>Base pair tilt (Å)</td>
<td>20</td>
<td>-6</td>
<td>7</td>
<td>16.7</td>
</tr>
<tr>
<td>Diameter of helix (Å)</td>
<td>23</td>
<td>20</td>
<td>18</td>
<td>21</td>
</tr>
<tr>
<td>Rotation per bp</td>
<td>33</td>
<td>36</td>
<td>-30</td>
<td>32.7</td>
</tr>
<tr>
<td>Glycosidic bond</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dA, dT, dC</td>
<td>anti</td>
<td>anti</td>
<td>anti</td>
<td>anti</td>
</tr>
<tr>
<td>dG</td>
<td>anti</td>
<td>anti</td>
<td>syn</td>
<td>anti</td>
</tr>
<tr>
<td>Sugar pucker</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dA, dT, dC</td>
<td>C3' endo</td>
<td>C2' endo</td>
<td>C2' endo</td>
<td>C3' endo</td>
</tr>
<tr>
<td>dG</td>
<td>C3' endo</td>
<td>C2' endo</td>
<td>C3' endo</td>
<td>C3' endo</td>
</tr>
<tr>
<td>Phosphate-phosphate (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>dA, dT, dC</td>
<td>5.9</td>
<td>7.0</td>
<td>7.0</td>
<td>~5.7</td>
</tr>
<tr>
<td>dG</td>
<td>7.0</td>
<td>7.0</td>
<td>5.9</td>
<td>~5.7</td>
</tr>
<tr>
<td>Major groove</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>width (Å)</td>
<td>2.7</td>
<td>11.7</td>
<td>Convex</td>
<td>11.1</td>
</tr>
<tr>
<td>depth (Å)</td>
<td>13.5</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor groove</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>width (Å)</td>
<td>11.0</td>
<td>5.7</td>
<td>4</td>
<td>11.1</td>
</tr>
<tr>
<td>depth (Å)</td>
<td>2.8</td>
<td>7.5</td>
<td>9</td>
<td>shallow</td>
</tr>
</tbody>
</table>

a CpG step. b GpC step.
DNA provide distinct surfaces for the interaction of transcription factors through specific and nonspecific interactions such as (hydrogen bonding, electrostatics, van der Waals, etc.). Several DNA-binding transcription factors are presented in Figure 1.7 to highlight the diverse architectures used for recognizing DNA, ranging from homodimeric coiled coils interacting with the major groove to monomeric beta-sheet containing proteins interacting with the minor groove. In addition to homodimeric motifs, heterodimeric motifs are utilized along with metal ion coordinated assemblies (i.e. Nf-κB p65-p50 and androgen receptor). Transcriptional co-activating proteins serve to integrate information from transcription factor assemblies and modulate gene expression through communication with RNA polymerase II leading to the transcription of protein-coding regions in the eukaryotic genome.

Transcription factors (TF) can communicate indirectly through allosteric modulation of DNA resulting in cooperative assembly with very little direct protein-protein interaction. Transcription factor binding can cause DNA-sequence dependent structural perturbations which modulate the binding of the next TF. TFs can also interact directly through protein-protein interactions to increase

**Figure 1.5** Nucleosome core particle structures. a) Crystal structure of the nucleosome core particle at 2.80Å determined by Richmond and coworkers. b) X-ray structure of the tetra-NCP determined at 9Å resolution. c) Theoretical model of a chromatin fiber constructed from four tetra-NCPs by rotation and translation about the central axis.
cooperativity. The β-enhancesome (Figure 1.8) is one such example of a cooperative assembly with cooperativity most likely arising at the DNA and coactivator levels. A conserved stretch of 55 bp's

Figure 1.6 Anatomy of the DNA base pair edges and their molecular recognition properties. a) Electrostatic potential maps of the base pair edges presented to the major (top) and minor (bottom) grooves of DNA. Hydrogen bond donors are designated with red arrows and the letter D. Hydrogen bond acceptors are designated with purple arrows and the letter A and functionality for the potential for van der Waals interactions is designated with the appropriate colored line and vdW. b) Top view of the Watson-Crick base pair molecular selectrostatic potential surfaces. Electrostatic potential maps from native DNA crystal structure sovled in Chapter 4 of this Thesis.
(160 Å long) in a nucleosome free region of the IFN-β promoter serves as a regulatory element for the cooperative assembly of 8 proteins into a continuous surface, burying 72% of the DNA solvent accessible area with very little protein-protein interaction.\textsuperscript{35} Transcriptional co-activating proteins serve to integrate information from the assembly to modulate gene expression through communication with RNA polymerase II leading to transcription.\textsuperscript{33,34}

\textbf{Figure 1.7} X-ray structures of DNA binding transcription factors [GCN4 (Harrison, 1993), Nf-κB p65-p50 (Ghosh, 1998), TBP-TATA box (Dickerson, 1996), Zif268 (Pabo, 1996), Androgen receptor (Gewirth, 2004)].
A diverse range of natural products and secondary metabolites have been shown to bind DNA with interaction modes consisting primarily of either intercalation or groove binding. In addition, some ligands rely on a combination of intercalation and groove binding that can also be augment by covalent modifying chemical domains, as in the case of anthramycin and neocarzinostatin. A collection of diverse DNA binding natural products are shown in Figure 1.9 with echinomycin and daunomycin representing intercalators and anthramycin and dystamicin A representing minor groove binders. The natural product distamycin (Figure 1.9) binds to A,T tracks in the minor groove of DNA, four to five base pairs in size, in both a 2:1 and a 1:1 ligand:DNA stoichiometry. The affinity and specificity of distamycin is controlled by a superposition of shape complementarity, hydrophobic effects, and specific hydrogen bonding to the minor groove of B-form DNA. Due to its modular design of repeating pyrrole amino acids and ammenability to rational modification, distamycin has served as the inspiration for the design of several classes sequence specific DNA minor-groove binders, with the ultimate goal of designing highly specific targeted gene regulation agents.
1.4 DNA Recognition by Minor-Groove Binders

Prior to the first structure of a molecule bound to DNA, specific recognition of B-form DNA was predicted to occur in major groove rather than minor groove. An observation that was based on the fact that the hydrogen bond acceptors at N3 of adenine and O2 of thymine A/T base pairs are similarly placed and lack any prominent distinguishing features. With the combination of biophysical and structural data from NMR and X-ray studies, it was verified that the minor groove of B-form DNA was a legitimate target for specific recognition. (For crystal structures of netropsin and distamycin A, see Figure 1.10.) Building upon inspiration from the natural products, netropsin
and dystamycin, minor groove binders have progressed to a modular molecular recognition platform with high affinity and specificity for many different sequences of DNA.\textsuperscript{5,7,8,16,17}

Over the past two decades, the development of minor groove DNA binders has evolved from the initial discovery of the natural product dystamycin to a new class of programmable heterocyclic oligomers demonstrating high affinity and sequence specificity.\textsuperscript{16,17} In addition to the incorporation of alternative heterocycles such as imidazole that have enabled specificity for guanine recognition using the Im-Py pair, much research has gone into linking the two heterocyclic strands in a dimeric motif.\textsuperscript{37-39} Covalent linkage of the two anti-parallel heterocyclic strands by a gamma amino butyric acid (GABA) linkage results in increases in affinity of 100–3600 fold relative to the unlinked homodimeric motif.\textsuperscript{40,41} The incorporation of the turn linkage in the form of a GABA or substituted GABA turn represented a major technological advance allowing for the first time the incorporation of unsymmetrical ring pairs for the targeting of non-palindromic DNA sequences.\textsuperscript{37} In addition, covalent linkage of the two strands has led to sub-nanomolar increases in affinity competing with and often rivaling that of andogenous DNA binding proteins.\textsuperscript{16,17,37,39} This high affinity modular dimeric motif has allowed for the regulation of gene expression by direct interaction with the DNA-

Figure 1.10 X-ray crystal structures illustrating DNA recognition by the natural products netropsin (left) and distamycin (middle, right). The dicationic natural product netropsin binds preferentially to narrow AT tracts in the DNA minor groove as a monomer (PDB: 6BNA). The monocationic natural product distamycin also binds AT tracts of the DNA minor groove in a 1:1 (PDB: 2DND) and 2:1 (PDB: 378D) ligand:DNA stoichiometry.
The four Watson-Crick base pairs can be differentiated by their molecular shape, electrostatic potential, and positions of hydrogen bond donors and acceptors in the DNA minor groove floor. The minor groove edge of a G•C base pair contains a steric hydrogen bond donating “bump” in the form of the exocyclic amine of guanine. The steric properties of the exocyclic amine of guanine form the basis for the A•T selectivity observed for netropsin and dystamycin binding due to steric interaction with the edge of the pyrrole ring. It was discovered in a key study in the early 1990s that imidazole in place of pyrrole in a three ring polyamide analogous to dystamycin could bind the 5 base pair sequence 5’-WGWCW-3’ (where W=A or T) resulting in a 2:1 polyamide-DNA complex where the imidazole ring is stacked against a pyrrole ring allowing differentiation of G•C base pairs from C•G, A•T, or T•A.42 The Im/Py pair has been used extensively in unlinked polyamides and in turn linked polyamides culminating in the recent publication of a polyamide library, that represents

Figure 1.11 Crystal structure (PDB: 365D) of the 2:1 binding single strand Py/Im polyamide targeted to the sequence 5’-CCAGGCTGG-3’ (2.00 Å resolution). Overall complex is shown on the right and the space filling model showing the basis for GC recognition is at the bottom left where the imidazole lone pair forms a hydrogen bond with the exocyclic N-H of guanine.

from C•G, A•T, or T•A.42
the solutions for targeting various 6 base pair sequences with high affinity and specificity.\textsuperscript{16,17,43} Thermodynamic studies have revealed that the Im/Py pair sequence selectivity is primarily driven by favorable enthalpic factors\textsuperscript{44,45} and X-ray crystallographic studies in collaboration with the Rees group provided structural insight based on a specific hydrogen bond between the imidazole lone pair and the exocyclic amine of guanine in unlinked 2:1 homodimeric polyamides (Figure 1.11).\textsuperscript{46}

Discrimination of T•A from A•T base pairs represents a much greater challenge due to the ability of thymine and adenine to both accept a hydrogen bond and the lack of unsymmetrical steric features as in the G•C case.\textsuperscript{16,17} Despite this challenge, a small asymmetric cleft between the C2 of adenine and the O2 of thymine has been exploited for specific targeting by the N-methyl-3-hydroxypyrrole/N-methylpyrrole (Hp/Py) pair, however affinities of these molecules are slightly lower than their Py/Py containing counterparts.\textsuperscript{16,17} In another seminal structural study with the Rees group on Hp containing 2:1 binders, it was revealed that a combination of shape selective recognition of the asymmetric cleft along with a specific hydrogen bond between the Hp hydroxyl and the thymine O2 was responsible for the A•T specificity (Figure 1.12).\textsuperscript{47,48} The combination of

\textbf{Figure 1.12} Crystal structure (PDB: 407D) of the 2:1 binding single strand ImHpPyPy-β-Dp polyamide targeted to the sequence 5'-CCAGTACTGG-3' (2.20 Å resolution). Overall complex is shown on the right and the space filling model of the basis for AT recognition is shown at the bottom left where the hydroxypyrrole-OH is within hydrogen bonding distance of the carbonyl lone pair on T.
Py, Im, and Hp combined as unsymmetrical pairs in opposite strands of a unlinked homodimeric or turn linked polyamide can be used to specifically recognize the four Watson-Crick base pairs.
These interactions can be described as a set of guidelines or pairing rules for the design of sequence specific B-form DNA targeted polyamides where Im/Py specifies G•C and Hp/Py specifies A•T. Some limitations do exist for certain sequences such as homopurine tracts, certain G-rich sequences, and sequences beyond 6 base pairs due to the sequence-dependent DNA microstructure and overcurvature of longer polyamides, however unique solutions to some of these problems have been developed (i.e. incorporation of a flexible β-Ala residue in 1:1 and hairpin polyamide motifs).

In addition to the 2:1 and hairpin polyamide architectures many other strand linkage strategies have been explored such as linking through the N-methyl groups on the central heterocycles (H-pin motif) or in the terminal heterocycles (U-pin motif). However, one of the highest affinity and in some cases most specific polyamide architectures has been the covalent linking of the C- to N-termini at both ends of the polyamide into a macrocycle, eliminating all possibility of extended binding modes. Macrocyclic γ-turn linked polyamides were first explored as 6 ring systems targeting a 5 base pair sequence of DNA in 1995 and were shown to have significantly higher affinity, however their specificity versus mismatch DNA was only 3-fold compared to 40-fold for their hairpin counterparts. Mainly due to limitations in synthetic methodology and initial discouraging thermodynamic results the cyclic polyamide motif was not investigated further until 1999. After improvements in solid-phase synthetic methodology, nanomole to micromole quantities of polyamides could be readily synthesized although cyclic polyamides still remained

![Equilibrium Association Constants (K\textsubscript{a})](image)

**Figure 1.14** Consequence of covalent attachment of two polyamide strands by incorporation of GABA-based turns. (For a structural key to the ball-and-stick nomenclature see the Nomenclature and Symbology section at the beginning of this thesis or Figure 1.13).
Using solid-phase methods cyclic polyamides were reinvestigated with two major architectural changes.\textsuperscript{41,53-55} The first being the use of an 8 rings system as oppose to 6 in the original studies and the second major change was moving the charge from the pyrrole $N$-methyl group to the alpha position of the $\gamma$-turn in the form of $(R)$-2,4-diaminobutyric acid, that had been discovered to increase the affinity, sequence specificity, and orientational preference of hairpin polyamides. This second generation cyclic 8 ring polyamide motif was found to have greatly improved specificity and affinity over its hairpin and unlinked counterparts targeting the sequence 5'-AGTACT-3'.\textsuperscript{41} The results of covalent attachment of the two polyamide strands can be seen in Figure 1.14. In a second study, multiple Hp/Py pairs were introduced into the 8-ring cyclic polyamide motif, which resulted in increased affinity and specificity relative to hairpin polyamides targeted to the sequences 5'-TGAACT-3' and 5'-TGATCT-3'.\textsuperscript{53} Despite these advances the cyclic polyamide motif has received little attention relative to its hairpin counterpart mainly due to synthetic limitations.

The chiral $(R)$-2,4-diaminobutyric acid turn ($\alpha$-turn) was a major advance in polyamide design not only for the cyclic polyamides but primarily for the hairpin motif.\textsuperscript{38} The addition of an amino substituent to the alpha position of the $\gamma$-turn helps to disfavor extended 1:1 binding modes and reverse binding due to a steric clash with the minor groove floor. In addition, the chiral amino turn helps to increase the overall affinity of polyamides while maintaining specificity and improving water solubility.\textsuperscript{38} The chiral $\alpha$-turn was proposed to increase affinity through electrostatic interactions between the protonated cationic amine group and the anionic DNA backbone however this interaction has not been born out in structural studies (see Chapter 5 and 6 of this thesis).\textsuperscript{17} In
addition to substitution of the \( \gamma \)-turn at the alpha position, many other variations have been studied with shorter turns, longer turns, and conformationally constrained variations.\textsuperscript{16,17,56} The most recent and one of the most successful advances in polyamide turn technology is the \( \beta \)-amino \( \gamma \)-linked turn.\textsuperscript{39} Recent studies have demonstrated that this turn can provide substantial increases in affinity for certain polyamide sequences, however the effect is less pronounced as the imidazole content of the polyamide is increased (Figure 1.15).\textsuperscript{39}

NMR structural studies using NOESY-restrained molecular dynamics models have also provided insight into 1:1 and 6-ring cyclic polyamide DNA complexes.\textsuperscript{55,57} Figure 1.16 shows an NMR model of a 1:1 polyamide with 11 out of 40 of the best calculated models overlayed.\textsuperscript{57} This shows very small coordinate deviation towards the center of the DNA helix and bound polyamide with increasing conformational mobility at the ends. The 6-ring cyclic polyamide model represented in Figure 1.16 shows an overlay of 21 of the best calculated models.\textsuperscript{55} This structure shows significant conformational mobility in the 6-ring cyclic complex with a highly flexible DNA sugar-phosphate backbone.

Figure 1.16 NMR NOESY-restrained molecular dynamics models of 1:1 and 6-ring cyclic polyamide DNA complexes.
Figure 1.17 Polyamide 2:1 DNA crystal structures colored by B-factor with red representing the largest B-factors and blue representing the smallest.
X-ray crystallographic studies resulting from collaborations between the Rees and Dervan groups have provided valuable insight into the polyamide-DNA molecular recognition process by elucidating the structure of five 2:1 polyamide-DNA complexes at a resolution ranging from 2.00 to 2.27 Å and a summary comparing specific structural parameters is shown in Figure 1.17. These crystallographic studies revealed that the DNA rise per base pair matches the polyamide rise

**Figure 1.18** X-ray crystal structures of polyamide-NCP complexes. a) Five polyamides in complex with the NCP at 2.30 Å resolution. b) Two polyamides in complex with the NCP at 2.45 Å resolution. c) One polyamide in complex with the NCP at 2.65 Å resolution.

X-ray crystallographic studies resulting from collaborations between the Rees and Dervan groups have provided valuable insight into the polyamide-DNA molecular recognition process by elucidating the structure of five 2:1 polyamide-DNA complexes at a resolution ranging from 2.00 to 2.27 Å and a summary comparing specific structural parameters is shown in Figure 1.17. These crystallographic studies revealed that the DNA rise per base pair matches the polyamide rise

**Figure 1.19** Turn-linked polyamide clamp bound to the NCP with the linker traversing the nucleosome super-groove.
per residue, however the polyamide structure is over-curved with respect to the DNA minor groove and shape complementarity is lost beyond a sequence of 5 contiguous base pairs.\textsuperscript{46} In addition, these crystallographic studies elucidated the basis for GC recognition by the Im/Py pair\textsuperscript{46} and TA recognition by the Hp/Py pair\textsuperscript{47,48} providing fundamental insight into polyamide binding. A variety of space groups were observed including monoclinic, orthorhombic, and trigonal with resolutions ranging from 2.00 to 2.27 Å. Average R-factors were in the mid-20s and polyamide B-factors averaged 47 to 86 Å\textsuperscript{2} whereas DNA B-factors averaged 43 to 67 Å\textsuperscript{2} for all structures presented in Figure 1.17. The 2:1 binding polyamide crystal structures also frequently exhibited disorder in the polyamide tail region and was usually modeled in alternate conformations reflecting the dynamic nature of the β-alanine linked dimethylamino propylamine terminus.

DNA binding polyamides are also able to bind sequence specifically to DNA on the nucleosome core particle.\textsuperscript{58} Hairpin polyamide-NCP crystal structures have been solved at resolutions ranging from 2.05 to 2.65 Å providing structural proof that polyamides can bind biologically relevant higher-order DNA structure however a combination of resolution limits and high B-factors for the polyamide prevented a detailed picture beyond confirmation of the polyamide binding location (Figures 1.18 and 1.19).\textsuperscript{58,59} The current state of macromolecular crystallography, with regard to minor groove binding DNA-drug structures, was assessed prior to beginning the structural work presented in Chapter 5 and 6 of this thesis and is presented in Figure 1.20. This survey demonstrates the lack of high resolution structures of DNA minor groove binders and the notable absence of linked dimeric minor groove binder structures. This survey underscores the pressing need for atomic resolution X-ray crystal structures of DNA minor groove binders to truly understand the molecular basis of recognition.

1.5 Scope of this work

The work presented in this thesis is focused on the molecular recognition of DNA by minor groove binding polyamides. In Chapter 2 of this thesis, a solution-phase synthesis of pyrrole-imidazole polyamides is presented with optimized protocols utilizing little to no chromatography. Chapter 3 builds on synthetic methodology in Chapter 2 allowing the efficient synthesis of cyclic polyamides targeted to the androgen response element. This chapter demonstrates that cyclic polyamides can be synthesized in an efficient manner, are biologically active and cell permeable in cell culture experiments, and rival the binding affinity of most other polyamide architectures. Chapter 4 details an oligomerization route to macrocyclic polyamides and reports on the DNA
Figure 1.20 Current state of macromolecular crystallography: A DNA-drug perspective. Data compiled from the PDB on 11/04/2007. (The number of structures solved is designated in parenthesis.)
binding ability of higher order macrocycles. The structural elucidation of an \( \alpha \)-amino-turn-linked cyclic polyamide is presented in Chapter 5 at 1.18 Å resolution providing insight into the detailed molecular recognition processes. Chapter 6 details the structural elucidation of a \( \beta \)-amino-turn-linked cyclic polyamide highlighting the conformational differences compared to the \( \alpha \)-amino-turns and providing a structural basis for the inability of polyamides to bind dsRNA. In Chapter 7, a new class of programmable oligomers targeting the DNA sequence 5'-WGGGGW-3' were shown to inhibit DNA binding of the Nf-\( \kappa \)B transcription factor by EMSA gel shift. Compounds discovered in Chapter 7 were found to possess unique fluorescent properties with the ability to modulate their fluorescence by binding their targeted dsDNA site and this work is presented in Chapter 8. Chapter 9 describes an ongoing effort in the templated-assembly of polyamides using higher-order DNA structure (NCP). Additionally, this chapter describes the development of a new profluorescent class of heterocycle, that has the potential to be used as a chemical reporter for templated ligation events. Appendix A through F detail results from efforts not covered in the main thesis and a continuation of studies from Chapters 3 and 8.

1.6 Notes and Reference


17. Dervan, P. B., and Edelson, B. S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides *Curr. Opin. Struct. Biol.* 2003, 13, 284-299. (For a cyclic polyamide crystal structure containing the α-amino turn see Chapter 4 of this thesis.)


38. Herman, D. M., Baird, E. E., and Dervan, P. B. Stereochemical Control of the DNA Binding Affinity, Sequence Specificity, and Orientation Preference of Chiral Hairpin Polyamides in the


Chapter 2: Solution-Phase Synthesis of Pyrrole–Imidazole Polyamides

The text of this chapter was taken in part from a manuscript coauthored with Daniel A. Harki, and Peter B. Dervan* (Caltech)

Abstract

Pyrrole–imidazole polyamides are DNA-binding molecules that are programmable for a large repertoire of DNA sequences. Typical syntheses of this class of heterocyclic oligomers rely on solid-phase methods. Solid-phase methodologies offer rapid assembly on a micromole scale sufficient for biophysical characterizations and cell culture studies. In order to produce gram-scale quantities necessary for efficacy studies in animals, polyamides must be readily synthesized in solution. An 8-ring hairpin polyamide 1, which targets the DNA sequence 5'-WGWWCW-3', was chosen for our synthesis studies as this oligomer exhibits androgen receptor antagonism in cell culture models of prostate cancer. A convergent solution-phase synthesis of 1 from a small set of commercially available building blocks is presented which highlights principles for preparing gram quantities of pyrrole–imidazole oligomers with minimal chromatography.
2.1 Introduction

Pyrrole–imidazole polyamides are a class of small molecules that bind the minor groove of DNA sequence-specifically.\textsuperscript{1,2} Encoded by side-by-side arrangements of \textit{N}-methylpyrrole (Py) and \textit{N}-methylimidazole (Im) carboxamide monomers, Im/Py pairs distinguish G•C from C•G base pairs, whereas Py/Py pairs are degenerate for T•A and A•T.\textsuperscript{3} Hairpin Py-Im polyamides have been shown to bind a broad repertoire of DNA sequences,\textsuperscript{4} permeate cell membranes and traffic to the nucleus,\textsuperscript{5} access chromatin,\textsuperscript{6} and disrupt protein–DNA interfaces.\textsuperscript{2} Hairpin polyamide inhibition of transcription factor–DNA binding of HIF-1\textsubscript{α},\textsuperscript{7} androgen receptor (AR),\textsuperscript{8} and AP-1\textsuperscript{9} has been exploited for controlling expression of medically relevant genes such as VEGF, PSA, TGF-β1, and LOX-1 in cell culture experiments.

An underpinning to transition polyamide studies from cell culture to small animal disease models is the ability to synthesize Py-Im polyamides on gram-scale. Over the years advances in polyamide solid-phase synthesis have been reported, including Boc- and Fmoc-based approaches from our laboratories and others.\textsuperscript{10} Solid-phase methodologies offer many advantages for milligram-scale polyamide syntheses, including rapid and reliable amino acid couplings and facile purifications owing to immobilization of the polyamide oligomer on a solid support. However, these techniques intrinsically limit the scale of synthesis. Conversely, efficient gram-scale solution-phase methods for polyamide synthesis that avoid arduous chromatographic purifications and employ commercially available Py-Im amino acid building blocks as reagents are less well developed. Remarkably, solution-phase synthesis of hairpin polyamides was the standard in our laboratory prior to the development of solid-phase methodologies,\textsuperscript{11a} and many variations on this theme have been published.\textsuperscript{11} However, laborious chromatographic purifications and modest reaction yields are commonplace. Therefore, we sought to develop a general solution-phase polyamide synthesis method that would allow access to gram quantities of material in high yield with minimal chromatography.

We report a proof-of-principle study demonstrating that hairpin Py-Im polyamides can be synthesized in solution from a small set of building blocks on large scale with minimal use of chromatography. This method involves Boc-protected dimers, trimers, and tetramers of heterocycles suitable for convergent syntheses. By exploiting differences in the physical solubility properties of starting materials versus products, a solution-phase synthesis of an 8-ring hairpin Py-Im polyamide 1 (Figure 2.1) has been achieved. Notably, our synthesis permits core polyamide 2 (Figure 2.2) to be prepared without a single chromatographic purification, thereby providing large quantities
of 2 for subsequent modification at the C-terminus such as 1. Py-Im polyamide 1, which targets the DNA sequence 5′-WGWWCW-3′, was selected for our studies because it antagonizes AR binding to androgen response elements (ARE) in gene promoters and regulates a subset of AR-driven genes, such as PSA. The regulation of aberrant AR-activated gene expression in prostate cancer is a promising strategy for developing novel therapeutics. This biological activity, coupled with our desire to conduct small animal efficacy experiments with 1, renders this polyamide an ideal candidate for scale-up and optimization studies. In addition, we discuss unifying principles for planning solution-phase polyamide syntheses of different Py-Im arrangements.

### 2.2 Results and Discussion

Our retrosynthetic approach for the preparation of an 8-ring hairpin polyamide, ImPyPyPy-(R)β-ImPyPy+(−)-IPA (1), is shown in Figure 2.2. Sequential couplings of ImPyPyPy tetramer 3 to turn moiety 5, followed by ester saponification and coupling to PyPyPy trimer 4, afford polyamide 2 in a convergent manner. Advanced intermediates 3–5 were prepared from building blocks 6–10, which have been previously synthesized by our laboratory and others. The cornerstone of our synthesis strategy capitalizes on the disparate physical properties of starting materials versus products, which permit purification of each intermediate to be achieved by combinations of precipitation, trituration, and crystallization. Such details are described below.

The synthesis of pyrrole trimer 4 begins with pyrrole amine salt 10 (Scheme 2.1). Amide coupling of 10 with activated pyrrole monomer 7 affords dimer 11 in 93% yield. The utilization of a small excess of 10 relative to 7 drives the reaction to completion, and residual 10 is readily separated from 11 following precipitation in water and aqueous washing of the residual solid 11. Reaction of dimer 11 with anhydrous HCl in diethyl ether removes the carbamate protecting group and facilitates precipitation of 12 as the HCl salt during the course of the reaction. Isolation of solid 12 by filtration, followed by washing of the residual solid with excess Et2O, provides the amine.
HCl salt in 99% yield. By exploiting the aqueous solubility of 10 versus insolubility of 11, PyPy dimer 11 is easily purified from a small excess of 10 by precipitation, whereas deprotected 12 is separable from Boc-protected 11 by virtue of the Et₂O solubility of 11 versus insolubility of 12. This reaction sequence highlights our synthesis strategy: exploiting the different solubility profiles of reactants versus products for chromatography-free purifications. Accordingly, pyrrole trimer 4 was obtained from dimer 12 by coupling with 7, followed by acidic deprotection to yield 4 in 95% yield (two steps) from dimer 12. The ImPyPyPy tetramer 3 was synthesized in two steps from trimer 4. 1-Methyl-2-trichloroacetylimidazole (6), prepared in one step from N-methylimidazole,10c,13 was allowed to react with a small excess of 4 to deliver tetramer 14 in 83% yield following precipitation in H₂O, trituration with Et₂O, and drying in vacuo. Saponification of 14 with aqueous NaOH in 1,4-dioxane, followed by neutralization with aqueous HCl, precipitation, and Et₂O trituration, afforded tetramer 3 in 77% overall yield from trimer 4.

The Im-turn fragment 5 was synthesized in two steps from the Im•HCl salt monomer10a 9 by coupling to PyBOP-activated (R)-3,4-Cbz-Dbu(Boc)-OH (8) yielding protected dimer 15 in 95% yield (Scheme 2.2). The utilization of a small excess of 9 drives the coupling reaction to completion and is easily separated in the aqueous wash step. Removal of the carbamate protecting
group with anhydrous HCl in 1,4-dioxane yielded the final synthon for our studies, imidazole-turn dimer 5, in quantitative yield following filtration and washing of the residual salt. With compound 5 in hand, the assembly of core polyamide 2 was initiated. PyBOP-mediated coupling of tetramer 3 to a small excess of water-soluble Im-turn dimer 5 yielded the advanced intermediate 16 in 97% yield. Saponification of 16 to acid 17, followed by amide coupling with an excess of water-soluble trimer 4, delivered core Py-Im hairpin polyamide 2 in 88% yield for the two steps. Consistent with the previously discussed strategy for synthesizing intermediates 3–5, the differences in solubility of reactants versus products were exploited to isolate pure material by precipitation, washing, and trituration. In most cases a low-boiling-point solvent was employed in the final trituration step to facilitate efficient solvent removal in vacuo.

Core polyamide 2 was synthesized without a single chromatographic purification in high overall purity, as depicted by the analytical HPLC analysis of 2 shown in Figure 2.3. Multigram quantities of 2 have been readily synthesized by this method, providing a stockpile of material for elaboration at the C-terminus into discrete polyamide conjugates, such as 1.

Py-Im polyamide 1 was synthesized in solution from advanced core 2 by coupling the preassembled C-terminal tail moiety 21 with saponified core 22. This convergent approach begins by preparing Boc-protected C-terminus moiety 20 (Scheme 2.3). PyBOP-mediated coupling of mono-Boc-protected triamine linker 18 with monobenzyl-protected isophthalic acid 19 afforded Boc-protected 20 in 98% yield. Deprotection of 20 with anhydrous CF₃CO₂H in dichloromethane (1:1) yielded amine 21, which was used immediately following concentration under high vacuum. Saponification of core polyamide 2 with aqueous NaOH in 1,4-dioxane at 23 °C yielded 8-ring acid 22 in 89% yield (Scheme 2.4). The transformation of 2 to 22 proved somewhat difficult.

Scheme 2.1 Preparation of 3 and 4. Reagents and Conditions: (i) DMF, DIEA, 7, 23 °C, 8 h, 93%; (ii) 2.0 M HCl in Et₂O, 23 °C, 18 h, 99%; (iii) DMF, DIEA, 7, 23 °C, 8 h, 96%; (iv) 4.0 M HCl in 1,4-dioxane, 23 °C, 18 h, 99%; (v) DMF, DIEA, 6, 23 °C, 2 h, 83%; (vi) NaOH (aq), 1,4-dioxane, 42 °C, 2 h, 93%.
in early studies due to formation of an unidentified side product. Avoiding reaction temperatures above 23 °C suppresses most of the byproduct formation, whereas a screen of aqueous bases commonly used for ester saponification (KOH and LiOH) failed to identify a better reagent. Coupling of residual acid 22 with freshly prepared 21 delivered the penultimate oligomer 23 in 87% yield. Unfortunately, crude oligomer 23 could not be satisfactorily purified by our standard method and required chromatography on silica gel to achieve pure material. Hundreds of milligrams of 23 have been prepared by this method in a single reaction sequence. Global deprotection of 23 by hydrogenation (Pd/C, ~1 atm H₂) at 23 °C for 48 h yields Py-Im polyamide 1 in 81% yield. Final product 1 can be separated from residual catalyst by solid-phase extraction and then purified by preparative reverse-phase HPLC.

Scheme 2.2 Preparation of 5 and assembly of core polyamide 2. Reagents and Conditions: (i) DMF, DIEA, PyBOP, (R)-3,4-Cbz-Dbu(Boc)-OH (8), 23 °C, 8 h, 95%; (ii) HCl in 1,4-dioxane, 23 °C, 16 h, 99%; (iii) DMF, DIEA, PyBOP, 3, 23 °C, 2 h, 97%; (iv) KOH (aq), MeOH, 1,4-dioxane, 42 °C, 2 h, 92%; (i) DMF, DIEA, PyBOP, 4, 23 °C, 10 h, 96%.

of laboratory and biologically relevant solvents. As shown in Figure 2.4, a strong solvent influence on the molar extinction coefficient of 1 is observed as the amount of organic cosolvent is increased. For example, an extinction coefficient (ε, M⁻¹cm⁻¹) of 26500 was measured for 1 in distilled and deionized H₂O, whereas this value doubled to 54800 in 50% acetonitrile in aqueous CF₃CO₂H (0.1% v/v CF₃CO₂H), a widely utilized laboratory solvent for purifying and quantifying peptides. A stock solution frequently encountered for preparing biological samples, 10% DMSO in DEPC-treated H₂O yielded an intermediary value of 40900 M⁻¹cm⁻¹. Hence, care must be taken to consider the solvent
2.3 Conclusion

A solution-phase synthesis of Py-Im hairpin polyamide 1 is presented, highlighting unifying principles for the preparation of related polyamides. A convergent synthesis was developed, requiring no chromatographic purifications to arrive at core 8-ring polyamide 2 on multi-gram scale. Final elaboration of the C-terminus affords AR polyamide antagonist 1 in high yield. The synthetic methodology permits gram-scale synthesis of Py-Im polyamides, a minimum next step as we transition these small molecules to animal models for biological efficacy.

2.4 Experimental Section

2.4.1 General

Chemicals were purchased from Sigma-Aldrich and were used without further purification. (R)-3,4-Cbz-Dbu(Boc)-OH was purchased from Senn Chemicals AG (code number 44159). Bulk grade solvents were from Fisher Scientific. Centrifugation system utilized when performing UV spectroscopy to determine polyamide concentrations.
was performed in a Beckman Coulter benchtop centrifuge (Allegra 21R) equipped with a Beckman swing-out rotor (model S4180). Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250 × 4.6 mm, 5 μm) and a diode array detector, and the mobile phase consisted of a gradient of acetonitrile (MeCN) in 0.1% (v/v) aqueous CF$_3$CO$_2$H. Preparative HPLC was performed on an Agilent 1200 system equipped with a solvent degasser, a diode array detector, and a Phenomenex Gemini column (250 × 21.2 mm, 5 μm). A gradient of MeCN in 0.1% (v/v) aqueous CF$_3$CO$_2$H was utilized as the mobile phase. UV–Vis measurements were made on a Hewlett-Packard diode array spectrophotometer (model 8452 A). NMR spectroscopy was performed on a Varian instrument operating at 499.8 MHz (for $^1$H) or 125.7 MHz (for $^{13}$C) at ambient temperature. All NMR analyses were performed in DMSO-$d_6$, and chemical shifts are reported in parts per million relative to the internal solvent peak referenced to 2.49 (for $^1$H) or 39.5 (for $^{13}$C). High-resolution mass spectrometry (HRMS) was recorded in positive-ion mode by fast-atom bombardment (FAB$^+$) on a JEOL JMS-600H instrument or by electrospray ionization (ESI$^+$) on a Waters Acquity UPLC-LCT Premiere XE TOF-MS system.

**Figure 2.4** UV properties of polyamide 1 in solvent systems I–VIII. (A) UV traces of I–VIII from 250–600 nm and (B) tabular form of data. Molar extinction coefficients were calculated from the $\lambda_{max}$ for each individual system, which ranged from 313 to 317 nm. DEPC = diethylpyrocarbonate treated H$_2$O.

<table>
<thead>
<tr>
<th>Solvent System</th>
<th>Molar Extinction Coefficient (M$^{-1}$cm$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>H$_2$O (distilled &amp; deionized) 26,500</td>
</tr>
<tr>
<td>II</td>
<td>H$_2$O (DEPC) 27,600</td>
</tr>
<tr>
<td>III</td>
<td>2% DMSO in H$_2$O (DEPC) 34,300</td>
</tr>
<tr>
<td>IV</td>
<td>5% EtOH in H$_2$O (DEPC) 35,400</td>
</tr>
<tr>
<td>V</td>
<td>10% DMSO in H$_2$O (DEPC) 40,900</td>
</tr>
<tr>
<td>VI</td>
<td>10% MeCN in H$_2$O (0.1% CF$_3$CO$_2$H) 46,800</td>
</tr>
<tr>
<td>VII</td>
<td>10% DMF in H$_2$O (0.1% CF$_3$CO$_2$H) 53,800</td>
</tr>
<tr>
<td>VIII</td>
<td>50% MeCN in H$_2$O (0.1% CF$_3$CO$_2$H) 54,800</td>
</tr>
</tbody>
</table>
2.4.2 HCl•H₂N-Py-CO₂Me (10)
Prepared as previously described.¹⁰¹H NMR: δ 10.09 (br s, 3H), 7.25 (d, J = 2.0 Hz, 1H), 6.80 (d, J = 2.2 Hz, 1H), 3.85 (s, 3H), 3.74 (s, 3H); ¹³C NMR: δ 160.2, 123.7, 120.8, 113.9, 111.4, 51.3, 36.6; HRMS (FAB+) calc’d for C₇H₁₁N₂O₂ [M+H]⁺ 155.0821, found 155.0847.

2.4.3 BocHN-PyPy-CO₂Me (11)
A solution of BocHN-Py-OBt ⁷ (6.16 g, 17.2 mmol) and HCl•H₂N-Py-CO₂Me ¹⁰ (3.61 g, 19.0 mmol) in DMF (39 mL) and DIEA (6 mL, 34.4 mmol) was stirred at 23 ºC for 8 h. The solution was then added to distilled H₂O (500 mL) preacidified with aqueous HCl (1 N, 300 mL, 300 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H₂O (80 mL) and collected by centrifugation (repeated 2X). The resultant solid, which contained a small amount of residual H₂O, was frozen and lyophilized to dryness. Drying of the light-brown solid in vacuo yielded dimer ¹¹ (6.03 g, 93%). ¹H NMR: δ 9.84 (s, 1H), 9.10 (s, 1H), 7.44 (d, J = 1.7 Hz, 1H), 6.88 (m, 2H), 6.82 (s, 1H), 3.82 (s, 3H), 3.79 (s, 3H), 3.72 (s, 3H), 1.44 (s, 9H); ¹³C NMR: δ 160.8, 158.4, 152.8, 123.0, 122.6, 122.4, 120.7, 118.4, 117.1, 108.3, 103.8, 78.3, 50.9, 36.1, 36.0, 28.2; HRMS (FAB+) calc’d for C₁₈H₂₅N₄O₅ [M+H]⁺ 377.1825, found 377.1835.

2.4.4 HCl•H₂N-PyPy-CO₂Me (12)
Dimer ¹¹ (4.0 g, 10.6 mmol) in a solution of anhydrous HCl in Et₂O (2.0 M, 400 mL) was stirred at 23 ºC for 18 h. The mixture was then diluted with 400 mL of anhydrous Et₂O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et₂O and dried in vacuo to yield dimer ¹² as a tan solid (3.3 g, 99%). ¹H NMR: δ 10.12 (s, 1H), 10.07 (br s, 3H), 7.46 (d, J = 2.0 Hz, 1H), 7.10 (d, J = 2.0 Hz, 1H), 7.00 (d, J = 2.0 Hz, 1H), 6.91 (d, J = 2.0 Hz, 1H), 3.87 (s, 3H), 3.83 (s, 3H), 3.72 (s, 3H); ¹³C NMR: δ 160.8, 157.7, 124.6, 122.6, 121.7, 120.8, 118.6, 113.1, 108.4, 107.2, 51.0, 36.6, 36.2; HRMS (FAB+) calc’d for C₁₃H₁₇N₄O₃ [M+H]⁺ 277.1301, found 277.1292.

2.4.5 BocHN-PyPy-CO₂Me (13)
A solution of BocHN-Py-OBt ⁷ (3.1 g, 8.7 mmol) and dimer ¹² (3.0 g, 9.59 mmol) in DMF (20 mL) and DIEA (3 mL, 17.4 mmol) was stirred at 23 ºC for 8 h. The solution was then added to distilled H₂O (250 mL) preacidified with aqueous HCl (1 N, 150 mL, 150 mmol), yielding a precipitate that
was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H$_2$O (40 mL) and collected by centrifugation (repeated 2x). The resultant solid, which contained a small amount of residual H$_2$O, was frozen and lyophilized to dryness. Drying of the light-brown solid \textit{in vacuo} yielded trimer 13 (4.17 g, 96%). $^1$H NMR: $\delta$ 9.91 (s, 1H), 9.86 (s, 1H), 9.09 (s, 1H), 7.46 (d, $J$ = 2.0 Hz, 1H), 7.21 (d, $J$ = 1.7 Hz, 1H), 7.05 (d, $J$ = 1.5 Hz, 1H), 6.89 (m, 2H), 6.83 (s, 1H), 3.80 (s, 3H), 3.73 (s, 3H), 1.45 (s, 9H); $^{13}$C NMR: $\delta$ 160.8, 158.5, 158.4, 152.8, 123.0, 122.8, 122.4, 122.30, 122.29, 120.7, 118.48, 118.47, 117.0, 108.3, 104.8, 103.8, 78.3, 50.9, 36.2, 36.05, 36.04, 28.2; HRMS (FAB$^+$) calc’d for C$_{24}$H$_{30}$N$_6$O$_6$ [M$^+$] $^+$ 498.2227, found 498.2233.

2.4.6 HCl$\cdot$H$_2$N-PyPyPy-CO$_2$Me (4)

Trimer 13 (4.0 g, 8.02 mmol) in a solution of anhydrous HCl in 1,4-dioxane (4.0 M, 300 mL) was stirred at 23 ºC for 18 h. The mixture was then diluted with 600 mL of anhydrous Et$_2$O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et$_2$O and dried \textit{in vacuo} to yield trimer 4 as a brown-orange solid (3.45 g, 99%). $^1$H NMR: $\delta$ 10.16 (s, 3H), 10.13 (s, 1H), 9.97 (s, 1H), 7.46 (d, $J$ = 2.0 Hz, 1H), 7.25 (d, $J$ = 1.7 Hz, 1H), 7.11 (d, $J$ = 2.0 Hz, 1H), 7.08 (d, $J$ = 2.0 Hz, 1H), 7.02 (d, $J$ = 2.0 Hz, 1H), 6.91 (d, $J$ = 2.0 Hz, 1H), 3.88 (s, 3H), 3.84 (s, 3H), 3.82 (s, 3H), 3.72 (s, 3H); $^{13}$C NMR: $\delta$ 160.8, 158.4, 157.7, 124.8, 123.0, 122.6, 121.9, 121.6, 120.8, 118.7, 118.5, 113.0, 108.4, 107.2, 104.8, 51.0, 36.6, 36.2, 36.1; HRMS (FAB$^+$) calc’d for C$_{19}$H$_{22}$N$_6$O$_4$ [M$^+$] $^+$ 398.1702, found 398.1685.

2.4.7 ImPyPyPy-CO$_2$Me (14)

A solution of trimer 4 (1.019 g, 2.34 mmol) and 1-methyl-2-trichloroacetylimidazole (6)$^{10c,13}$ (478 mg, 2.10 mmol) in DMF (4.5 mL) and DIEA (910 μL, 5.22 mmol) was stirred at 23 ºC for 2 h. The solution was then added to distilled H$_2$O (40 mL) pre-acidiﬁed with aqueous HCl (1N, 910 μL, 0.91 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H$_2$O (40 mL) and collected by centrifugation (repeated 1X). The resultant solid, which contained a small amount of residual H$_2$O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et$_2$O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et$_2$O and dried \textit{in vacuo} to yield tetramer 14 as a light brown solid (886 mg, 83%). $^1$H NMR: $\delta$ 10.68 (s, 1H), 10.00 (s, 1H), 9.94 (s, 1H), 7.48 (s, 1H), 7.47 (d, $J$ = 2.0 Hz, 1H), 7.31 (d, $J$ = 1.7 Hz, 1H), 7.24 (d, $J$ = 1.7 Hz, 1H), 7.18 (s, 1H), 7.17 (d, $J$ = 2.0 Hz, 1H), 7.08 (d, $J$ = 2.0 Hz, 1H), 6.91 (d, $J$ = 1.7 Hz, 1H), 4.01 (s, 3H), 3.86 (s,
3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.73 (s, 3H); 13C NMR: δ 160.8, 158.5, 158.4, 155.0, 138.2, 126.4, 125.6, 123.1, 123.0, 122.5, 122.2, 121.2, 120.7, 118.8, 118.6, 118.5, 108.3, 104.9, 104.8, 50.9, 36.18, 36.17, 36.1, 35.4; HRMS (FAB+) calc'd for C24H27N8O5 [M+H]+ 507.2104, found 507.2116.

2.4.8 ImPyPyPy-CO2H (3)
A solution of tetramer 14 (804 mg, 1.59 mmol) in 1,4-dioxane (8 mL) and aqueous NaOH (1N, 8.0 mL, 8.00 mmol) was stirred at 42 ºC for 2 h. The solution was then added to distilled H2O (40 mL) pre-acidified with aqueous HCl (1N, 8 mL, 8.00 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H2O (40 mL) and collected by centrifugation (repeated 1X). The resultant solid, which contained a small amount of residual H2O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et2O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et2O and dried in vacuo to yield tetramer 3 as a brown solid (728 mg, 93%). 1H NMR: δ 10.45 (s, 1H), 9.97 (s, 1H), 9.91 (s, 1H), 7.43 (s, 1H), 7.38 (s, 1H), 7.29 (s, 1H), 7.24 (s, 1H), 7.17 (s, 1H), 7.07 (s, 1H), 7.03 (s, 1H), 6.85 (s, 1H), 3.99 (s, 3H), 3.85 (s, 3H), 3.84 (s, 3H), 3.82 (s, 3H); 13C NMR: δ 162.0, 158.48, 158.46, 156.1, 138.8, 127.0, 126.4, 123.0, 122.7, 122.6, 122.2, 121.5, 120.3, 119.5, 118.7, 118.5, 108.4, 105.0, 104.8, 36.14, 36.07, 35.1; HRMS (FAB+) calc’d for C23H25N8O5 [M+H]+ 493.1948, found 493.1952.

2.4.9 BocHN-(R)-CbzHN-γ-Im-CO2Et (15)
A solution of (R)-3,4-Cbz-Dbu(Boc)-OH 8 (1.03 g, 2.93 mmol) and PyBOP (1.83 g, 3.51 mmol) in DMF (12 mL) and DIEA (1.5 mL, 8.8 mmol) was stirred at 23 ºC for 22 min. The solution was then added to solid (powdered) HCl•H2N-Im-CO2Et 9 (850 mg, 3.15 mmol) and stirred at 23 ºC for 8 h. The solution was then added to distilled H2O (30 mL) pre-acidified with aqueous HCl (1N, 9 mL, 9 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H2O (40 mL) and collected by centrifugation (repeated 2X). The resultant solid, which contained a small amount of residual H2O, was frozen and lyophilized to dryness. Drying of the brown solid in vacuo yielded dimer 15 (1.4 g, 95%). 1H NMR: δ 10.57 (s, 1H), 7.51 (s, 1H), 7.31-7.27 (m, 5H), 7.02 (d, J = 8.5 Hz, 1H), 6.79 (m, 1H), 4.97 (s, 2H), 4.25 (q, J = 7.2 Hz, 2H), 3.93 (m, 1H), 3.89 (s, 3H), 3.01 (m, 2H), 2.44-2.35 (m, 2H), 1.35 (s, 9H), 1.27 (t, J = 7.2 Hz, 3H); 13C NMR: δ 167.6, 158.4, 155.8, 155.4, 137.4, 137.1, 130.7, 128.2, 127.6, 127.5, 114.8, 77.7, 65.1, 60.5, 48.6, 43.5, 38.1, 35.4, 28.2, 14.0; HRMS (FAB+) calc’d for C24H34N5O7...
[M+H]+ 504.2458, found 504.2462.

2.4.10 HCl•H2N-(R)β-ChInγ-Im-CO2Et (5)

Dimer 15 (500 mg, 0.993 mmol) in a solution of anhydrous HCl in 1,4-dioxane (4.0 M, 10 mL) and anhydrous Et₂O (4 mL) was stirred at 23 °C for 16 h. The mixture was then diluted with 20 mL of anhydrous Et₂O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et₂O and dried in vacuo to yield dimer 5 as a white solid (432 mg, 99%). 1H NMR: δ 10.75 (s, 1H), 8.05 (m, 3H), 7.52 (s, 1H), 7.38 (d, J = 8.5 Hz, 1H), 7.34-7.28 (m, 5H), 5.01 (m, 2H), 4.25 (q, J = 7.1 Hz, 2H), 4.13 (m, 1H), 3.90 (s, 3H), 2.96-2.84 (m, 2H), 2.60-2.51 (m, 2H), 1.27 (t, J = 7.1 Hz, 3H); 13C NMR: δ 166.8, 158.4, 155.7, 137.2, 136.8, 130.9, 128.3, 127.8, 127.7, 114.9, 65.5, 60.6, 46.6, 42.1, 38.2, 35.4, 14.0; HRMS (FAB+) calc’d for C₁₉H₂₆N₅O₅ [M+H]+ 404.1934, found 404.1928.

2.4.11 ImPyPyPy-(R)β-ChInγ-Im-CO2Et (16)

A solution of tetramer 3 (1.28 g, 2.59 mmol) and PyBOP (1.49 g, 2.86 mmol) in DMF (8 mL) and DIEA (1.8 mL, 10.4 mmol) was stirred at 23 °C for 20 min. The solution was then treated with solid (powdered) HCl•H₂N-(R)β-ChInγ-Im-CO₂Et 5 (1.2 g, 2.73 mmol) and stirred at 23 °C for 2 h. The solution was then added to distilled H₂O (30 mL) pre-acidified with aqueous HCl (1 N, 20 mL, 20 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H₂O (40 mL) and collected by centrifugation (repeated 2x). The resultant solid, which contained a small amount of residual H₂O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et₂O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et₂O and dried in vacuo to yield ImPyPyPy-(R)β-ChInγ-Im-CO₂Et 16 as a tan solid (2.2 g, 97%). 1H NMR: δ 10.60 (s, 1H), 10.46 (s, 1H), 9.96 (s, 1H), 9.91 (s, 1H), 7.97 (t, J = 5.6 Hz, 1H), 7.52 (s, 1H), 7.39 (s, 1H), 7.29-7.26 (m, 6H), 7.24 (m, 1H), 7.18-7.14 (m, 3H), 7.04 (m, 2H), 6.90 (m, 1H), 4.98 (m, 2H), 4.24 (q, J = 7.1 Hz, 2H), 4.10 (m, 1H), 3.99 (s, 3H), 3.89 (s, 3H), 3.845 (s, 3H), 3.839 (s, 3H), 3.835 (s, 3H), 3.77 (s, 3H), 3.29 (m, 2H), 2.49 (m, 2H, obstructed by NMR solvent), 1.27 (t, J = 7.2 Hz, 3H); 13C NMR: δ 167.7, 161.5, 158.49, 158.48, 158.45, 156.0, 155.6, 138.7, 137.4, 137.1, 130.8, 128.2, 127.7, 127.6, 126.8, 126.4, 123.0, 122.8, 122.7, 122.24, 122.17, 121.4, 118.7, 118.5, 118.0, 114.9, 105.0, 104.7, 104.4, 65.2, 60.5, 48.6, 42.1, 38.2, 36.13, 36.11, 36.0, 35.4, 35.2, 14.0; HRMS (FAB+) calc’d for C₄₂H₄₈N₁₇O₉ [M+H]+ 878.3698, found 878.3668.
2.4.12 ImPyPyPy-(R)\(\beta\)-CbzHN\(\gamma\)-Im-CO\(_2\)H (17)

A solution of ImPyPyPy-(R)\(\beta\)-CbzHN\(\gamma\)-Im-CO\(_2\)Et 16 (2.0 g, 2.28 mmol) dissolved in 1,4-dioxane (2 mL), MeOH (6 mL), and aqueous KOH (1 N, 9.1 mL, 9.1 mmol) was stirred at 42 °C for 2 h. The solution was then acidified with aqueous HCl (1 N, ~9.1 mL, ~9.1 mmol) to a pH = 4.5, yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H\(_2\)O (10 mL) and collected by centrifugation (repeated 1x). The resultant solid, which contained a small amount of residual H\(_2\)O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et\(_2\)O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et\(_2\)O and dried in vacuo to yield ImPyPyPy-(R)\(\beta\)-CbzHN\(\gamma\)-Im-CO\(_2\)H 17 as a tan solid (1.78 g, 92%). \(^1\)H NMR: \(\delta\) 10.50 (s, 1H), 10.47 (s, 1H), 9.96 (s, 1H), 9.92 (s, 1H), 7.98 (m, 1H), 7.48 (s, 1H), 7.40 (s, 1H), 7.30-7.27 (m, 6H), 7.24 (m, 1H), 7.19-7.14 (m, 3H), 7.05 (m, 2H), 6.90 (m, 1H), 4.99 (m, 2H), 4.10 (s, 1H), 3.99 (s, 3H), 3.88 (s, 3H), 3.845 (s, 3H), 3.838 (s, 3H), 3.77 (s, 3H), 3.30 (m, 2H), 2.49 (m, 2H, obstructed by NMR solvent); \(^1\)C NMR: \(\delta\) 167.7, 161.5, 160.0, 158.5, 155.8, 155.6, 138.6, 137.1, 131.6, 128.3, 127.7, 127.6, 126.7, 126.4, 123.0, 122.8, 122.7, 122.23, 121.4, 118.7, 118.5, 118.0, 114.6, 105.0, 104.7, 104.4, 65.2, 48.6, 42.2, 38.2, 36.13, 36.10, 36.0, 35.4, 35.2; HRMS (FAB\(^+\)) calc’d for C\(_{40}\)H\(_{44}\)N\(_{13}\)O\(_9\) [M+H]\(^+\) 850.3385, found 850.3383.

2.4.13 ImPyPyPy-(R)\(\beta\)-CbzHN\(\gamma\)-ImPyPyPy-CO\(_2\)Me (2)

A solution of ImPyPyPy-(R)\(\beta\)-CbzHN\(\gamma\)-Im-CO\(_2\)H 17 (1.5 g, 1.77 mmol) and PyBOP (546 mg, 1.85 mmol) in DMF (8.8 mL) and DIEA (922 μL, 5.3 mmol) was stirred at 23 ºC for 10 min. The solution was then treated with solid (powdered) HCl•H\(_2\)N-PyPyPy-CO\(_2\)Me 4 (806 mg, 1.85 mmol) and stirred at 23 ºC for 10 h. The solution was then added to distilled H\(_2\)O (35 mL) preacidified with aqueous HCl (1 N, 5 mL, 5 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H\(_2\)O (40 mL) and collected by centrifugation (repeated 3x). The resultant solid, which contained a small amount of residual H\(_2\)O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et\(_2\)O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et\(_2\)O and dried in vacuo to yield ImPyPyPy-(R)\(\beta\)-CbzHN\(\gamma\)-ImPyPyPy-CO\(_2\)Me 2 as a tan solid (2.09 g, 96%). \(^1\)H NMR: \(\delta\) 10.66 (s, 1H), 10.21 (s, 1H), 10.00 (s, 1H), 9.98 (s, 1H), 9.96 (s, 1H), 9.93 (s, 1H), 9.92 (s, 1H), 8.00 (m, 1H), 7.48 (s, 1H), 7.46 (d, \(J = 2.0\) Hz, 1H), 7.45 (s, 1H), 7.31-7.27
(m, 7H), 7.23 (m, 2H), 7.20-7.14 (m, 5H), 7.06 (m, 2H), 6.92 (m, 1H), 6.90 (d, J = 2.0 Hz, 1H), 5.00 (m, 2H), 4.11 (m, 1H), 4.00 (s, 3H), 3.95 (s, 3H), 3.854 (s, 3H), 3.850 (s, 3H), 3.842 (s, 3H), 3.837 (s, 3H), 3.828 (s, 3H), 3.78 (s, 3H), 3.73 (s, 3H), 3.32 (m, 2H), 2.53 (m, 2H); 13C NMR: δ 167.9, 161.6, 160.8, 158.5, 158.42, 158.38, 155.8, 155.6, 137.6, 137.1, 136.0, 134.0, 128.3, 127.7, 127.6, 126.4, 123.3, 123.1, 123.0, 122.80, 122.77, 122.5, 122.26, 122.24, 122.1, 121.2, 120.9, 120.8, 118.9, 118.70, 118.64, 118.5, 118.0, 114.1, 108.4, 104.9, 104.79, 104.76, 104.5, 65.2, 51.0, 48.8, 42.2, 38.4, 36.25, 36.21, 36.19, 36.13, 36.10, 36.0, 35.8, 35.0; HRMS (FAB+) calc’d for C_{59}H_{64}N_{19}O_{12} [M+H]+ 1230.498, found 1230.504.

2.4.14 ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy-CO₂H (22)
A solution of polyamide 2 (500 mg, 0.406 mmol) dissolved in 1,4-dioxane (8 mL) and aqueous NaOH (1 N, 8.0 mL, 8.0 mmol) was stirred at 23 ºC for 11 h. The solution was then cooled to 0 ºC in an ice bath and the pH adjusted to pH = 4.0 with aqueous HCl (1 N, ~8 mL, 8 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H₂O (40 mL) and collected by centrifugation (repeated 2x). The resultant solid, which contained a small amount of residual H₂O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et₂O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et₂O and dried in vacuo to yield ImPyPyPy-(R)β-CbzHNγ-ImPyPyPy-CO₂H 22 as a tan solid (442 mg, 89%). 1H NMR: δ 10.52 (s, 1H), 10.22 (s, 1H), 10.00 (s, 1H), 9.95 (s, 1H), 9.93 (s, 1H), 9.90 (s, 1H), 8.00 (m, 1H), 7.45 (s, 1H), 7.42 (m, 2H), 7.30-7.27 (m, 7H), 7.24 (m, 2H), 7.19-7.14 (m, 4H), 7.09 (s, 1H), 7.06 (m, 2H), 6.92 (m, 1H), 6.84 (d, J = 2.0 Hz, 1H), 5.01 (m, 2H), 4.11 (m, 1H), 3.99 (s, 3H), 3.95 (s, 3H), 3.851 (s, 3H), 3.848 (s, 3H), 3.843 (s, 3H), 3.838 (s, 3H), 3.81 (s, 3H), 3.79 (s, 3H), 3.32 (m, 2H), 2.53 (m, 2H); 13C NMR: δ 167.9, 162.0, 161.6, 158.5, 158.42, 158.38, 155.8, 155.6, 138.6, 137.1, 136.0, 134.0, 128.3, 127.7, 127.6, 126.4, 123.06, 123.05, 122.8, 122.74, 122.70, 122.6, 122.24, 122.19, 122.15, 121.4, 121.2, 120.3, 119.5, 118.7, 118.55, 118.47, 118.0, 114.1, 108.4, 104.9, 104.86, 104.80, 104.75, 104.5, 65.2, 48.7, 42.2, 38.3, 36.2, 36.13, 36.10, 36.07, 35.97, 35.2, 34.9; HRMS (FAB+) calc’d for C_{58}H_{62}N_{19}O_{12} [M+H]+ 1216.483, found 1216.487.

2.4.15 BocHN-(+)-BnOIP A (20)
A solution of acid 19 (211 mg, 0.824 mmol) and PyBOP (624 mg, 1.2 mmol) in DMF (3 mL) and DIEA (211 μL, 1.2 mmol) was stirred at 23 ºC for 10 min. Protected triamine 18 (373 mg, 1.52
mmol) was then added to the solution and stirring was continued at 23 °C for 3 h. The solution was then added to distilled H₂O (15 mL) and a viscous oil was isolated following centrifugation (~4500 rpm) and decanting of the aqueous layer. The residual oil was again washed with distilled H₂O (10 mL) and collected by centrifugation (repeated 3x). Drying of the residual oil in vacuo yielded BocHN-(+) -BnOIPA 20 as a viscous amber oil (391 mg, 98%). ¹H NMR: δ 8.77 (t, J = 5.5 Hz, 1H), 8.43 (app t, J = 1.5 Hz, 1H), 8.14-8.09 (m, 2H), 7.63 (app t, J = 7.8 Hz, 1H), 7.48-7.34 (m, 5H), 6.85 (t, J = 4.9 Hz, 1H), 5.38 (s, 2H), 3.30 (m, 2H), 3.00 (m, 1H), 2.94 (m, 2H), 2.72-2.51 (m, 4H), 2.38 (br s, 3H), 1.75 (m, 2H), 1.59 (m, 2H), 1.34 (s, 9H); ¹³C NMR: δ 165.3, 165.2, 155.6, 136.0, 135.0, 131.9, 131.7, 129.8, 129.0, 128.6, 128.2, 128.1, 127.9, 77.5, 66.4, 54.3, 54.0, 40.6, 37.7, 37.3, 28.2, 25.9, 25.5; HRMS (FAB+) calc’d for C₂₇H₃₈N₃O₅ [M+H]+ 484.2811, found 484.2793.

2.4.16 ImPyPyPy-(R)-β-CbzHN- γ-ImPyPyPy-(+)-BnOIPA (23)

A solution of ImPyPyPy-(R)-β-CbzHN-γ-ImPyPyPy-CO₂H 22 (192 mg, 0.158 mmol) and PyBOP (99 mg, 0.189 mmol) in DMF (2 mL) and DIEA (220 μL, 1.3 mmol) was stirred at 23 °C for 30 min. In a separate vial, BocHN-(+)-BnOIPA 20 (122 mg, 0.252 mmol) was deprotected by treating with a solution of CH₂Cl₂:CF₃CO₂H (1:1, 2 mL) for 20 minutes at 23 °C followed by concentration to dryness in vacuo. This residual material was then treated with the preactivated solution of ImPyPyPy-(R)-β-CbzHN-γ-ImPyPyPy-CO₂H 22 and allowed to stir at 23 °C for 12 h. The solution was then added to distilled H₂O (30 mL) preacidified with aqueous HCl (1 N, 2 mL, 2 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H₂O (40 mL) and collected by centrifugation (repeated 3x). The resultant solid, which contained a small amount of residual H₂O, was frozen and lyophilized to dryness. The residual brown solid was purified by SiO₂ chromatography with the mobile phase consisting of a step gradient of 49:1 CH₂Cl₂:MeOH to 44:6:1 CH₂Cl₂:MeOH:NH₃ to provide ImPyPyPy-(R)-β-CbzHN-γ-ImPyPyPy-(+)-BnOIPA 23 as a tan solid (217 mg, 87%) after drying under high vacuum. ¹H NMR: δ 10.43 (s, 1H), 10.20 (s, 1H), 9.98 (s, 1H), 9.95 (s, 1H), 9.93 (s, 1H), 9.92 (s, 1H), 9.87 (s, 1H), 8.73 (t, J = 5.5 Hz, 1H), 8.41 (app t, J = 1.6 Hz, 1H), 8.11-8.07 (m, 2H), 8.03-7.98 (m, 2H), 7.60 (app t, J = 7.7 Hz, 1H), 7.48-7.32 (m, 7H), 7.30-7.23 (m, 9H), 7.19-7.14 (m, 5H), 7.06 (d, J = 1.7 Hz, 1H), 7.05 (d, J = 1.7 Hz, 1H), 7.03 (d, J = 1.0 Hz, 1H), 6.92 (d, J = 1.2 Hz, 1H), 6.84 (d, J = 1.7 Hz, 1H), 5.37 (s, 2H), 5.00 (m, 2H), 4.11 (m, 1H), 3.99 (s, 3H), 3.95 (s, 3H), 3.85 (s, 3H), 3.84 (m, 6H), 3.83 (s, 3H), 3.79 (s, 3H), 3.77 (s, 3H), 3.30 (m, 4H), 3.18 (m, 2H), 2.53 (m, 2H), 2.35 (m, 4H), 2.16 (br s, 3H), 1.71-1.60 (m, 4H); ¹³C NMR: δ 167.9, 165.20, 165.16, 161.6, 161.2, 158.50,
2.4.17 ImPyPyPy-(R)β-H2Nγ-ImPyPyPy-(+)IPA (1)

A solution of ImPyPyPy-(R)β-CbzNγ-ImPyPyPy-(+)IPA 23 (25 mg, 0.016 mmol) and Pd/C (10 wt % dry, 10 mg) in DMF (2 mL) was stirred at 23 ºC for 48 h under H2 (~1 atm, balloon). [Note: The protecting groups are cleaved at different rates, and early reaction aliquots reveal a monoprotected compound.] The reaction was then filtered through a Sep-Pak cartridge (5 g of C-18 sorbent), and the Sep-Pak was washed with DMF (4 mL), aqueous MeCN (50%, 20 mL), MeCN (250 mL), and MeOH (250 mL). The filtrate was then concentrated in vacuo, purified by reverse-phase HPLC, and lyophilized to dryness. The solid was triturated with anhydrous Et2O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et2O and dried in vacuo to yield ImPyPyPy-(R)β-H2Nγ-ImPyPyPy-(+)IPA 1 as a light tan solid (18.9 mg, 81%).

1H NMR: δ 10.61 (s, 1H), 10.46 (s, 1H), 9.95 (s, 2H), 9.94 (s, 1H), 9.92 (s, 1H), 9.89 (s, 1H), 9.46 (br s, 1H), 8.82 (t, \( J = 5.8 \) Hz, 1H), 8.42 (app t, \( J = 1.4 \) Hz, 1H), 8.20 (m, 1H), 8.15 (m, 1H), 8.08 (d, \( J = 1.7 \) Hz, 1H), 8.07 (d, \( J = 1.7 \) Hz, 1H), 7.97 (m, 3H), 7.60 (app t, \( J = 7.6 \) Hz, 1H), 7.45 (s, 1H), 7.40 (d, \( J = 0.7 \) Hz, 1H), 7.28 (d, \( J = 2.0 \) Hz, 1H), 7.26 (d, \( J = 1.7 \) Hz, 1H), 7.22 (d, \( J = 2.0 \) Hz, 1H), 7.21 (d, \( J = 1.7 \) Hz, 1H), 7.18 (d, \( J = 1.7 \) Hz, 1H), 7.16 (m, 2H), 7.15 (d, \( J = 1.7 \) Hz, 1H), 7.08 (m, 2H), 7.05 (d, \( J = 1.0 \) Hz, 1H), 7.03 (d, \( J = 1.7 \) Hz, 1H), 6.94 (d, \( J = 1.8 \) Hz, 1H), 3.99 (s, 3H), 3.96 (s, 3H), 3.85 (s, 3H), 3.844 (s, 3H), 3.842 (s, 3H), 3.83 (s, 3H), 3.82 (s, 3H), 3.79 (s, 3H), 3.69 (m, 1H), 3.44 (m, 2H), 3.35 (m, 2H), 3.24 (m, 2H), 3.18-3.12 (m, 2H), 3.11-3.04 (m, 2H), 2.76 (m, 3H), 2.50 (m, 2H, obstructed by NMR solvent), 1.95-1.91 (m, 2H), 1.90-1.83 (m, 2H); 13C NMR: δ 166.84, 166.82, 165.7, 162.1, 161.6, 158.52, 158.49, 158.40, 158.1, 157.8, 156.0, 155.7, 138.7, 135.6, 134.6, 134.2, 131.9, 131.5, 131.0, 128.7, 128.0, 126.9, 126.3, 123.1, 123.0, 122.8, 122.7, 122.5, 122.4, 122.24, 122.16, 122.10, 121.4, 121.1, 118.7, 118.5, 118.3, 118.0, 115.7, 105.0, 104.9, 104.8, 104.54, 104.52, 59.2, 53.3, 48.2, 36.5, 36.2, 36.08, 36.07, 35.99, 35.6, 35.38, 35.35, 35.1, 35.0, 24.3, 24.0; HRMS (ESI+) calc’d for C60H77N22O12 [M+H]+ 1357.6086, found 1357.6091.
2.4.18 Calculation of Molar Extinction Coefficients

Solid 1 (0.51 mg, 0.35 μmoles) was weighed into a tared vial on a microbalance (Sartorius Micro M4). A molecular weight of 1471.46 was utilized for 1, which assumes 1 exists as the CF₃CO₂H salt. The material was dissolved in distilled and deionized H₂O (1000 μL), yielding a 0.35 μM stock solution. Fifty-fold dilutions of this stock solution were made into the appropriate solvent systems (Figure 2.4). Data were collected at 23 °C in a 1 cm quartz cuvette, and molar extinction coefficients are based on the λ_max for each solvent system (range: 313–317 nm). The instrument was blanked on each discrete solvent system prior to data collection. Duplicate analysis of each sample yielded similar results (data not shown). A second polyamide stock solution was generated by dissolving solid 1 in 50% MeCN in aqueous CF₃CO₂H (0.1% v/v), followed by 50-fold dilution into the individual solvent systems, and data collection yielded similar results to those shown in Figure 2.4 (data not shown). DEPC-treated H₂O (RNase- and DNase-free) was purchased from USB Corp. EtOH was absolute grade from Pharmaco-AAPER, and DMSO was molecular biology grade from Sigma-Aldrich. DMF was anhydrous synthesis grade from Sigma-Aldrich, and acetonitrile was HPLC grade from Fisher.
2.5 Notes and References


12. Sources for commercially available building blocks 7–10 shown in Figure 2.2; 7 (Fluorochem, Cat # 019570), 8 (Fluka, Cat # 17974), 9 (Bachem, Cat # F-3480), 10 (Bachem, Cat # F-3485).


2.6 Spectra and Supplemental Information

$^1$H and $^{13}$C NMR spectra for synthesized compounds.

**Figure 2.5** $^1$H NMR of HCl•H$_2$N-Py-CO$_2$Me (10)
Figure 2.6: 13C NMR of HCl•H2N-Py-CO2Me (10)

[Diagram of 13C NMR spectrum with chemical structures]

EXP. Carbon

Data: Oct 17 2014
Sample: HCl•H2N-Py-CO2Me
File: /home/sdtuser/sdt.car

Acquisition: 31 p
SW: 51165.5
AT: 2.348
FP: 7642.5
TB: 84
Display: 50.0
Transmitter: 25.111.1
Decoupler: 4.203

[Detailed parameters for NMR experiment]
Figure 2.7: $^1$H NMR of BocHN-PyPy-CO$_2$Me (11)
Figure 2.8 $^{13}$C NMR of BocHN-PyPy-CO$_2$Me (11)
Figure 2.9 $^1$H NMR of HCl-H$_2$N-PyPy-CO$_2$Me (12)
Figure 2.10 13C NMR of HCl•H2N-PyPy-CO2Me (12)
Figure 2.11: $^1$H NMR of BocHN-PyPyPy-CO$_2$Me (13)
Figure 2.12 $^{13}$C NMR of BocHN-PyPyPy-CO$_2$Me (13)
Figure 2.13 $^1$H NMR of HCl$\cdot$H$_2$N-PyPyPy-CO$_2$Me (4)
Figure 2.14 $^{13}$C NMR of HCl•H$_2$N-PyPyPy-CO$_2$Me (4)
Figure 2.15 $^1$H NMR of ImPyPyPy-CO$_2$Me (14)
Figure 2.16 $^{13}$C NMR of ImPyPyPy-CO$_2$Me (14)
Figure 2.17 1H NMR of ImPyPyPy-CO2H (3)
Figure 2.18 13C NMR of ImPyPyPy-CO₂H (3)
Figure 2.19 $^1$H NMR of BocHN-($R$)β-ChnHN$\gamma$-Im-CO$_2$Et (15)
Figure 2.20 13C NMR of BocHN-(R)β-Cbzγ-Im-CO2Et (15)
Figure 2.21 $^1$H NMR of HCl$\cdot$H$_2$N-(R)$^{\beta}$-Cbz$\cdot$H$\cdot$Y-Im-CO$_2$Et (5)
Figure 2.22 $^{13}$C NMR of HCl•H$_2$N-(R)$^{\beta}$-CholN$_{\gamma}$-Im-CO$_2$Et (5)
Figure 2.23 $^1$H NMR of ImPyPyPy-($R$)$^{\beta}$ChaN$_\gamma$-Im$\cdot$CO$_2$Et (16)
Figure 2.24 $^{13}$C NMR of ImPyPyPy-$\beta$-Chol$\Delta$Py-Im-CO$_2$Et (16)
Figure 2.25: $^1$H NMR of ImPyPyPy-(R)-CbzHN $\beta$-Im-CO$_2$H (17)
Figure 2.26 $^{13}$C NMR of ImPyPyPy-$(R)\text{ChlIN}_{\gamma}$-Im-CO$_2$H (17)
Figure 2.27 $^1H$ NMR of ImPyPyPy-(R)$^\beta$-Cbz-$\gamma$-ImPyPyPy-CO$_2$Me (2)
Figure 2.28 $^{13}$C NMR of ImPyPyPy-$(R)^\beta$-ChuIm-ImPyPyPy-CO$_2$Me (2)
Figure 2.29 1H NMR of ImPyPyPy-(R)-CbzHNγ-ImPyPyPy-CO2H (22)
Figure 2.30 $^{13}$C NMR of ImPyPyPy-$(R)\beta$-ChltIN$_2$-ImPyPyPy-CO$_2$H (22)
Figure 2.31 $^1$H NMR of BocHN-(+)-Bn$^\text{Ot}$IPA (20)
Figure 2.32: 13C NMR of BocHN-(+)-BnOIPA (20)
Figure 2.33 1H NMR of ImPyPyPy-(R)-CbzHN-γ-ImPyPyPy-(+)-BnOIPA (23)
Figure 2.34 $^{13}$C NMR of ImPyPyPy-$\beta$-HCl$\text{H\textsubscript{2}N}$-ImPyPyPy-$\beta$-BnOIPA (23)
Figure 2.35 $^1$H NMR of ImPyPyPy-$\beta$H2N-$\gamma$-ImPyPyPy-(+)IPA (1)
Figure 2.36 $^{13}$C NMR of ImPyPyPy-$(R)^6{^{12}}$N-ImPyPyPy-(+)-IPA (1)
Chapter 3: Cyclic Pyrrole–Imidazole Polyamides Targeted to the

Androgen Response Element

The text of this chapter was taken in part from a manuscript coauthored with Daniel A. Harki, John W. Phillips, Christian Dose, and Peter B. Dervan* (Caltech)

Abstract

Hairpin pyrrole–imidazole (Py-Im) polyamides are a class of cell-permeable DNA-binding small molecules that can disrupt transcription factor–DNA binding and regulate endogenous gene expression. The covalent linkage of antiparallel Py-Im ring pairs with an γ-amino acid turn unit affords the classical hairpin Py-Im polyamide structure. Closing the hairpin with a second turn unit yields a cyclic polyamide, a lesser-studied architecture mainly attributable to synthetic inaccessibility. We have applied our methodology for solution-phase polyamide synthesis to cyclic polyamides with an improved high-yield cyclization step. Cyclic 8-ring Py-Im polyamides 1–3 targets the DNA sequence 5’-WGWWCW-3’ which corresponds to the androgen response element (ARE) bound by the androgen receptor transcription factor to modulate gene expression. We find that cyclic Py-Im polyamides 1–3 bind DNA with exceptionally high affinities and regulate the expression of AR target genes in cell culture studies, from which we infer that the cycle is cell permeable.
3.1 Introduction

Modulating the expression of eukaryotic gene networks by small molecules is a challenge at the frontier of chemical biology. Pyrrole–imidazole polyamides are a class of cell-permeable small molecules that bind to the minor groove of DNA in a sequence-specific manner.\textsuperscript{1,2} Side-by-side stacked \textit{N}-methylpyrrole (Py) and \textit{N}-methylimidazole (Im) carboxamides (Im/Py pairs) distinguish G•C from C•G base pairs, whereas Py/Py pairs specify for both T•A and A•T.\textsuperscript{3} Py-Im hairpin polyamides have been programmed for a broad repertoire of DNA sequences with affinities similar to endogenous transcription factors.\textsuperscript{4} They are cell permeable and influence gene transcription by disrupting protein–DNA interfaces.\textsuperscript{2,5,6} Hairpin polyamide interference of DNA binding by transcription factors such as HIF-1\texttext{	extalpha},\textsuperscript{7} androgen receptor (AR),\textsuperscript{8} and AP-1\textsuperscript{9} has been described in recent years, yielding a new approach toward gene control by small molecules.

In parallel with our gene regulation studies, a significant effort has been devoted to maximizing the biological potency of hairpin Py-Im polyamides through structural modifications. In particular, we have recently demonstrated that hairpin polyamides bearing the (\textit{R})-\textit{\beta}-amino-\textit{\gamma}-turn, such as polyamide 4, possess favorable binding affinities to DNA and are useful in gene regulation studies (Figure 3.1).\textsuperscript{5g} A significant effort exists in our laboratory to regulate aberrant AR-activated gene expression in prostate cancer.\textsuperscript{8} To further optimize lead oligomer 4, it would seem reasonable that closing the hairpin with an identical linker, yielding a cyclic structure 1, would further enhance DNA affinity (Figure 3.1). Previous syntheses of cyclic polyamides using

![Figure 3.1 Structures of cyclic and hairpin polyamides 1–5 targeted to the DNA sequence 5’-WGWWCW-3’ and their ball-and-stick models. Ball-and-stick representation legend: black and white circles represent \textit{N}-methylimidazole and \textit{N}-methylpyrrole units, respectively, half-circle with - sign represents the terminal isophthalic acid substituent, and white half-diamond with + sign represents the triamine linker unit.](image-url)
solid-phase protocols are characterized by low reaction yields due to inefficient macrocyclization.\textsuperscript{10}

We report here the solution-phase synthesis of cyclic polyamides 1–3 with an improved high–yield cyclization step. In addition, we examined the DNA binding properties of these compounds by thermal duplex DNA melting and performed preliminary studies of their \textit{in vitro} ADMET properties. Cyclic Py-Im polyamides 1–3 were shown to regulate endogenous gene expression in cell culture experiments.

### 3.2 Results and Discussion

#### 3.2.1 Solution-Phase Synthesis of Cyclic Polyamides

Due to the symmetrical nature of cyclic polyamides 1–3 and their sequence similarity to previously described hairpin polyamide 4,\textsuperscript{11} PyPyPy trimer 6 and Im-turn dimer 7 provide all the necessary atoms to synthesize 1–3. The preparation of advanced intermediates 6 and 7 has been detailed in the Chapter 2 (this thesis)\textsuperscript{11} from readily available building blocks.\textsuperscript{12} The cornerstone of our synthesis strategy capitalizes on the disparate physical properties of starting materials versus products, which permit purification of most intermediates to be achieved by combinations of precipitation, trituration, and crystallization. In addition, \textit{in situ} deprotection of advanced pentafluorophenyl ester polyamide 14 at high dilution leads to macrocyclization in high yield, affording cyclic polyamide 15.

The synthesis of tetramer-turn 9 begins with Im-turn dimer 7 (Scheme 3.1). Saponification of 7 with aqueous KOH in methanol at 37°C, followed by neutralization, precipitation, and Et\textsubscript{2}O trituration, yields Im-turn acid 8 in 95% yield. Amide coupling of 8 with pyrrole trimer 6 provides pentamer 9 in 96% yield. The utilization of a small excess of 6 relative to 8 drives the reaction to completion, and residual 6 is readily separated from 9 following precipitation in water and aqueous washing of residual solid

![Scheme 3.1](image)

**Scheme 3.1** Preparation of 10 and 11. Reagents and Conditions: (i) KOH (aq), MeOH, 37 °C, 2 h, 95%; (ii) 8, PyBOP, DMF, DIEA, 6, 23 ºC, 4 h, 96%; (iii) HCl in 1,4-dioxane, 23 ºC, 2 h, 99% (iv) NaOH (aq), 1,4-dioxane, 42 ºC, 3 h, 95%.
9. With all atoms in place for the target cyclic polyamide 1, compound 9 was elaborated to amine salt 10 (99% yield) by reaction with HCl in 1,4-dioxane. Carboxylic acid 11 was generated by saponification of 9 with NaOH in 1,4-dioxane in 95% yield.

Assembly of the acyclic advanced intermediate 12 was achieved by PyBOP-mediated coupling of intermediates 10 and 11 in 94% yield (Scheme 3.2). A small excess of amine salt 10 was utilized to drive the reaction to completion. Saponification of ester 12 proceeded smoothly with aqueous NaOH in 1,4-dioxane, yielding 13 in 93% yield. Activation of acid 13 as the pentafluorophenol ester 14 provided the necessary functionality to afford macrocyclization following removal of the terminal tert-butyl carbamate (Boc) protecting group. In our hands, we found that the pentafluorophenol ester sufficiently activated the terminal acid for amide coupling while avoiding undesired oligomerization and/or decomposition processes that are conceivable with more reactive functionalities, such as acid chlorides. Premature initiation of the macrocyclization reaction was tempered by keeping the terminal amine protonated until it was transferred into a dilute solution of acetonitrile. Addition of an amine base (DIEA) then generated the free terminal amine, which could then undergo macrocyclization in dilute solvent conditions to deliver 15, which was directly deprotected following purification. The benzyl carbamate protecting

Scheme 3.2 Preparation of 1, 2, and 3. Reagents and Conditions: (i) PyBOP, DMF, DIEA, 23 °C, 2 h, 94%; (ii) NaOH (aq), 1,4-dioxane, 40 °C, 4 h, 93%; (iii) CH₂Cl₂, DCC, pentafluorophenol, DMAP, 23 °C, 12 h, 80%; (iv) a) CF₃CO₂H, CH₂Cl₂, 23 °C, concentrate; b) DMF, acetonitrile, DIEA, 0–23 °C, 3 days; (v) CF₃SO₃H, CF₃CO₂H, 23 °C, 5 min, 68% over 3 steps; (vi) NMP, DIEA, Ac₂O, 23 °C, 18% of 1 (recovered), 22% of 2, 40% of 3.
groups were cleaved by treatment with superacid conditions (trifluoromethylsulfonic acid–trifluoroacetic acid) to provide 1 in 68% yield over three steps. Controlled acetylation of 1 by reaction with substoichiometric quantities of Ac₂O in NMP/DIEA provided a statistical population of 1 (18%), 2 (22%), and 3 (40%) that were easily separable by preparative HPLC. Acetylated hairpin 5 was prepared using excess Ac₂O/pyridine in 95% yield from previously reported amine hairpin 4. ¹¹

### 3.2.2 Thermal Stabilization of DNA duplexes by Polyamides

Quantitative DNase I footprint titrations have historically been utilized to measure polyamide–DNA binding affinities and specificities. ¹³ However, this method is limited to measuring $K_a$ values $\leq 2 \times 10^{10}$ M⁻¹, which invalidates this technique for quantifying the exceptionally high DNA-binding affinities of cycles 1–3. ¹⁴ The magnitude of DNA thermal stabilization ($\Delta T_m$) of DNA–polyamide complexes has been utilized to rank order polyamides with high DNA binding affinities. ⁵g,¹⁵ Accordingly, we have employed melting temperature analysis ($\Delta T_m$) for dissecting differences in DNA-binding affinities of hairpin versus cyclic polyamides. Spectroscopic analyses were performed on a 14-mer duplex DNA mimicking the androgen response element (ARE) DNA sequence, 5′-TTGCTGTTCTGCAA-3′ DNA duplex, which contains one polyamide binding site. As shown in Table 3.1 polyamides 1–5 provided an increase in the duplex DNA melting temperature relative to the individual DNA duplex, thereby confirming polyamide-DNA binding. Chiral hairpin 4 led to an increased melting temperature $\Delta T_m = 18.4$ °C whereas cyclic polyamide 1 yielded a higher $\Delta T_m$-value of 23.6 °C. Cyclic polyamides 1–3 reveal stronger stabilizations than parent hairpin analogs 4 and 5. Acylation of the $\beta$-amino turns was shown to decrease the thermal stabilization values in both hairpin and cyclic motifs, presumably due to the loss of beneficial electrostatics from the protonated-cationic amine on the turn unit.

<table>
<thead>
<tr>
<th>Table 3.1 $T_m$ values for polyamides for 1–5. ¹⁺⁺</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARE dsDNA sequence = 5′-TTGCTGTTCTGCAA-3′</td>
</tr>
<tr>
<td>3′-AAGC ACAAAGA CGTT-5′</td>
</tr>
<tr>
<td>Polyamides</td>
</tr>
<tr>
<td>---------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

¹⁺⁺All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses. $\Delta T_m$ values are given as $T_m$(DNA/polyamide) – $T_m$(DNA). The propagated error in $\Delta T_m$ measurements is the square root of the sum of the square of the standard deviations for the $T_m$ values.

---

Figure 3.2 Targeting the ARE with DNA-binding polyamides. (a) X-ray crystal structure of androgen receptor homodimer DNA-binding domain bound to the sequence 5'-CTGTTCTTGATGT-TCTGG-3' (PDB 1r4i). (b) Map of the PSA-ARE site (top) and schematic representation of a cyclic polyamide targeting the PSA-ARE site 5'-AGAACA-3'. (c) Inhibition of induced PSA mRNA expression in LNCaP cells by cyclic PI polyamides 1–3 and hairpin polyamides 4 and 5 (dosed at 0.3, 3, and 30 μM) by real-time quantitative PCR. The results were normalized to a DHT-induced, untreated control (control=1), and the error bars represent the standard error of the mean of a single experiment performed in biological triplicate. The entire experiment was reproduced four times, with similar results. NI = noninduced, I = induced, DHT = dihydrotestosterone.
3.2.3 Biological Assay for Cell Permeability

Hairpin polyamides have been shown to modulate endogenous gene expression in living cells by disrupting transcription factor–DNA binding in gene promoters.\textsuperscript{2,7,9} Recently, hairpin polyamide 4 was shown to inhibit androgen receptor-mediated expression of prostate-specific antigen (PSA) in LNCaP cells by targeting the DNA sequence 5’-AGAACA-3’ found in the ARE.\textsuperscript{5g} We utilized this cell culture transcription assay to investigate the biological activity of cyclic polyamides 1–3 in comparison to hairpin polyamides 4 and 5. Since small structural changes to polyamides have been shown to correlate with differences in cellular uptake properties,\textsuperscript{5} it was not obvious whether cyclic polyamides 1–3 would permeate cell membranes and exhibit biological activity comparable to that of hairpin polyamides 4 and 5. Quantitative real-time RT-PCR analysis of DHT-induced PSA expression revealed that cyclic polyamides 1–3 all decreased PSA mRNA levels in LNCaP cells, with cycle 1 exhibiting activity comparable to that of acetylated hairpin polyamide 5 (Figure 3.2). On the basis of these results, we can infer that this class of cyclic Py-Im polyamides are cell permeable and can regulate endogenous gene expression in cell culture.

3.2.4 ADMET Studies of Polyamides 1 and 5

Due to the promising cell culture results obtained with cyclic polyamide 1 and hairpin polyamide 5 against PSA gene expression, we contracted preclinical in vitro absorption, distribution, metabolism, excretion, and toxicity (ADMET)\textsuperscript{17} studies for both compounds.\textsuperscript{18} Results from this study are summarized below and additional detail can be found in section 3.6 (Spectra and Supporting Information) and Appendix B of this thesis. Polyamides 1 and 5 were both found to exhibit low Caco-2 permeability, suggesting that neither compound may be orally available. Both 1 and 5 were found to be almost exclusively protein bound in plasma, with half-lives greater than 2 h. A recent positron emission tomography (PET)-based biodistribution study of a related hairpin polyamide in mice revealed high levels of liver occupancy following tail vein dosage.\textsuperscript{19} On the basis of this result, we investigated the liver stability of candidate polyamides 1 and 5. Microsomal intrinsic clearance studies found half-lives greater than 3 h for 1 and 5 in both human and rat liver microsomes, and no significant inhibition was measured against any cytochrome P450 isoform examined (Cyp1A2/CEC, Cyp2C8/DBF, Cyp2C9/DBF, Cyp2C19/DBF, Cyp2D6/AMMC, Cyp3A4/BFC, Cyp3A4/DBF). Furthermore, no obvious toxicity (IC\textsubscript{50} > 100 μM) was observed in the human hepatocellular carcinoma cell line HepG2. In addition, standard hERG FastPatch assays of cardiac toxicity found that both polyamides (1 and 5) were devoid of unwanted inhibition (IC\textsubscript{50}...
3.3 Conclusion
We describe a solution-phase synthesis methodology for preparing cyclic Py-Im polyamides, highlighted by an efficient macrocyclization between the alkyl linker amine and a pentafluorophenol ester-activated amino acid. The three cyclic Py-Im polyamides possessed high DNA-binding affinities and were capable of accessing the nucleus in cell culture, as judged by their ability to downregulate AR-activated PSA expression in cell culture. Preclinical ADMET analysis of cyclic polyamide 1 and hairpin polyamide 5 revealed favorable drug-like properties such as high liver stability and low toxicity. Ongoing work is focused on characterizing the precise molecular interactions between cyclic polyamides and their cognate DNA sequences by high-resolution crystallographic studies.

3.4 Experimental Section
3.4.1 General
Chemicals and solvents were purchased from Sigma-Aldrich and were used without further purification. (R)-3,4-Cbz-Dbu(Boc)-OH was purchased from Senn Chemicals AG (code number 44159). All DNA oligomers were purchased HPLC purified from Integrated DNA Technologies. Water (18 MΩ) was purified using a Millipore Milli-Q purification system. Centrifugation was performed in a Beckman Coulter benchtop centrifuge (Allegra 21R) equipped with a Beckman swing-out rotor (model S4180). Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250 × 4.6 mm, 5 μm) and a diode array detector, and the mobile phase consisted of a gradient of acetonitrile (MeCN) in 0.1% (v/v) aqueous CF$_3$CO$_2$H. Preparative HPLC was performed on an Agilent 1200 system equipped with a solvent degasser, a diode array detector, and a Phenomenex Gemini column (250 × 21.2 mm, 5 μm). A gradient of MeCN in 0.1% (v/v) aqueous CF$_3$CO$_2$H was utilized as the mobile phase. UV–Vis measurements were made on a Hewlett-Packard diode array spectrophotometer (model 8452 A), and polyamide concentrations were measured in 0.1% (v/v) aqueous CF$_3$CO$_2$H using an extinction coefficient of 69200 M$^{-1}$·cm$^{-1}$ at $\lambda$$_{max}$ near 310 nm. NMR spectroscopy was performed on a Varian instrument operating at 499.8 (for $^1$H) or 125.7 MHz (for $^{13}$C) at ambient temperature. All NMR analyses were performed in DMSO-$d_6$, and chemical shifts are reported in parts per million relative to the internal solvent peak referenced to 2.49 (for $^1$H) or 39.5 (for $^{13}$C). High-resolution
mass spectrometry (HRMS) was recorded in positive-ion mode by fast-atom bombardment (FAB+) on a JEOL JMS-600H instrument or by electrospray ionization (ESI+) on a Waters Acquity UPLC-LCT Premiere XE TOF-MS system.

3.4.2 UV Absorption Spectrophotometry
Melting temperature analysis was performed on a Varian Cary 100 spectrophotometer equipped with a thermo-controlled cell holder possessing a cell path length of 1 cm. A degassed aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂ at pH 7.0 was used as analysis buffer. DNA duplexes and polyamides were mixed in 1:1 stoichiometry to a final concentration of 2 μM for each experiment. Prior to analysis, samples were heated to 90 °C and cooled to a starting temperature of 23 °C with a heating rate of 5 °C/min for each ramp. Denaturation profiles were recorded at λ = 260 nm from 23 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.

3.4.3 Measurement of Androgen-Induced PSA mRNA
Experiments were performed as described previously with the following modifications: (1) all compounds and controls were prepared in neat DMSO and then diluted with media to a final concentration of 0.1% DMSO, and (2) mRNA was isolated with the RNEasy 96 kit (Qiagen, Valencia, CA).

3.4.4 BocHN-(R)β-ChlHNγ-Im·CO₂H (8)
A solution of BocHN-(R)β-ChlHNγ-Im·CO₂Et 7 (450 mg, 0.894 mmol) dissolved in MeOH (1.0 mL) and aqueous KOH (1 N, 2.0 mL, 2.0 mmol) was stirred at 37 ºC for 2 h. The reaction mixture was added to a cooled (ice bath) solution of distilled H₂O (10 mL) preacidified with aqueous HCl (1 N, 2.0 mL, 2.0 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The resultant solid was again suspended in distilled H₂O (10 mL) and collected by centrifugation. The resultant solid, which contained a small amount of residual H₂O, was frozen and lyophilized to dryness and then suspended in excess anhydrous Et₂O and filtered, and the filter cake washed with copious amounts of anhydrous Et₂O. Drying of the brown solid in vacuo yielded saponified dimer 8 (404 mg, 95%). ¹H NMR: δ 10.46 (s, 1H), 7.47 (s, 1H), 7.31-7.26 (m, 5H), 7.02 (d, J = 8.3 Hz, 1H), 6.79 (t, J = 5.4 Hz, 1H), 4.97 (s, 2H), 3.92 (m, 1H), 3.88 (s, 3H), 3.60 (s, 1H), 3.01 (m, 2H),
2.41 (m, 2H), 1.35 (s, 9H); \(^{13}\)C NMR: δ 167.6, 160.0, 155.8, 155.4, 137.11, 137.09, 131.6, 128.2, 127.66, 127.55, 114.6, 77.7, 65.1, 48.7, 43.5, 38.0, 35.4, 28.2; HRMS (FAB\(^+\)) calc’d for C\(_{22}\)H\(_{30}\)N\(_5\)O\(_7\) [M+H]\(^+\) 476.2145, found 476.2130.

3.4.5 BocHN-(R)\(^\beta\)-CbzHN\(^\gamma\)-ImPyPyPy-CO\(_2\)Me (9)
A solution of BocHN-(R)\(^\beta\)-CbzHN\(^\gamma\)-Im-CO\(_2\)H \(\mathbf{8}\) (300 mg, 0.631 mmol) and PyBOP (345 mg, 0.663 mmol) in DMF (3.2 mL) and DIEA (330 \(\mu\)L, 1.9 mmol) was stirred at 23 ºC for 10 min. The solution was then treated with solid (powdered) HCl•H\(_2\)N-PyPyPy-CO\(_2\)Me \(\mathbf{6}\) (288 mg, 0.663 mmol) and stirred at 23 ºC for 4 h. The solution was then added to distilled H\(_2\)O (10 mL) preacidified with aqueous HCl (1 N, 2 mL, 2 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H\(_2\)O (10 mL) and collected by centrifugation (repeated 3x). The resultant solid, which contained a small amount of residual H\(_2\)O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et\(_2\)O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et\(_2\)O and dried in vacuo to yield BocHN-(R)\(^\beta\)-CbzHN\(^\gamma\)-ImPyPyPy-CO\(_2\)Me \(\mathbf{9}\) as a tan solid (518 mg, 96%). \(^1\)H NMR: δ 10.17 (s, 1H), 10.00 (s, 1H), 9.95 (s, 1H), 9.93 (s, 1H), 7.46 (d, \(J = 1.7\) Hz, 1H), 7.44 (s, 1H), 7.31-7.29 (m, 5H), 7.27 (d, \(J = 1.7\) Hz, 1H), 7.23 (d, \(J = 1.7\) Hz, 1H), 7.14 (d, \(J = 1.7\) Hz, 1H), 7.07 (d, \(J = 1.7\) Hz, 1H), 7.04 (d, \(J = 8.3\) Hz, 1H), 6.90 (d, \(J = 2.0\) Hz, 1H), 6.81 (t, \(J = 5.9\) Hz, 1H), 4.98 (s, 2H), 3.96 (m, 1H), 3.95 (s, 3H), 3.85 (s, 3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.73 (s, 3H), 3.03 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H); \(^{13}\)C NMR: δ 167.8, 160.8, 158.5, 158.4, 155.84, 155.81, 155.5, 137.1, 136.0, 134.0, 128.3, 127.7, 127.6, 123.06, 123.00, 122.5, 122.2, 121.2, 120.7, 118.7, 118.6, 118.5, 114.0, 104.9, 77.8, 65.2, 50.9, 48.8, 43.6, 38.2, 36.20, 36.18, 36.09, 34.9, 28.2; HRMS (FAB\(^+\)) calc’d for C\(_{41}\)H\(_{49}\)N\(_{11}\)O\(_{10}\) [M\(^+\)]\(^+\) 855.3663, found 855.3688.

3.4.6 HCl•H\(_2\)N-(R)\(^\beta\)-CbzHN\(^\gamma\)-ImPyPyPy-CO\(_2\)Me (10)
A solution of BocHN-(R)\(^\beta\)-CbzHN\(^\gamma\)-ImPyPyPy-CO\(_2\)Me \(\mathbf{9}\) (125 g, 0.146 mmol) in anhydrous HCl in 1,4-dioxane (4.0 M, 10 mL) was stirred at 23 ºC for 2 h. The mixture was then diluted with 100 mL of anhydrous Et\(_2\)O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et\(_2\)O and dried in vacuo to yield HCl•H\(_2\)N-(R)\(^\beta\)-CbzHN\(^\gamma\)-ImPyPyPy-CO\(_2\)Me \(\mathbf{10}\) as a brown solid (114 mg, 99%). \(^1\)H NMR: δ 10.38 (s, 1H), 9.98 (s, 1H), 9.96 (s, 1H), 9.94 (s, 1H), 8.10 (m, 3H), 7.46 (d, \(J = 1.7\) Hz, 1H), 7.45 (s, 1H), 7.42 (d, \(J = 8.3\) Hz, 1H), 7.34-7.28 (m, 5H), 7.28 (d, \(J = 1.7\) Hz, 1H), 7.24 (d, \(J = 1.7\) Hz, 1H), 7.16 (d, \(J = 1.7\) Hz, 1H), 7.08 (d, \(J = 1.7\) Hz, 1H).
93 Hz, 1H), 6.90 (d, J = 1.7 Hz, 1H), 5.02 (m, 2H), 4.14 (m, 1H), 3.95 (s, 3H), 3.85 (s, 3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.73 (s, 3H), 3.02 (m, 2H), 2.63 (m, 2H); \( ^{13} \)C NMR: \( \delta \) 167.0, 160.8, 158.5, 158.4, 155.7, 136.8, 135.7, 134.0, 128.3, 127.8, 127.7, 123.05, 122.97, 122.5, 122.2, 121.1, 120.7, 118.64, 118.60, 118.5, 104.9, 65.6, 50.9, 46.7, 42.0, 38.2, 36.2, 36.1, 36.0, 34.9; HRMS (FAB\(^{+}\)) calc’d for C\(_{36}\)H\(_{42}\)N\(_{11}\)O\(_8\) [M+H]\(^{+}\) 756.3218, found 756.3211.

3.4.7 BocHN-(R)\(^{\beta}\)-CbzHN\(^{\gamma}\)-ImPyPyPy-CO\(_2\)H (II)

A solution of BocHN-(R)\(^{\beta}\)-CbzHN\(^{\gamma}\)-ImPyPyPy-CO\(_2\)Me 9 (200 mg, 0.234 mmol) dissolved in 1,4-dioxane (2.3 mL) and aqueous NaOH (1 N, 2.3 mL, 2.3 mmol) was stirred at 42 ºC for 3 h. The solution was then added to distilled H\(_2\)O (5 mL) preacidified with aqueous HCl (1 N, 2.3 mL, 2.3 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H\(_2\)O (10 mL) and collected by centrifugation (repeated 2x). The resultant solid, which contained a small amount of residual H\(_2\)O, was frozen and lyophilized to dryness and then suspended in excess anhydrous Et\(_2\)O and filtered, and the filter cake washed with copious amounts of anhydrous Et\(_2\)O. Drying of the tan solid \textit{in vacuo} yielded BocHN-(R)\(^{\beta}\)-CbzHN\(^{\gamma}\)-ImPyPyPy-CO\(_2\)H 11 (187 mg, 95%). \(^1\)H NMR: \( \delta \) 12.15 (s, 1H), 10.21 (s, 1H), 10.00 (s, 1H), 9.96 (s, 1H), 9.92 (s, 1H), 7.44 (s, 1H), 7.42 (d, \( J = 1.7 \) Hz, 1H), 7.31-7.29 (m, 5H), 7.28 (d, \( J = 1.5 \) Hz, 1H), 7.24 (d, \( J = 1.5 \) Hz, 1H), 7.16 (d, \( J = 1.5 \) Hz, 1H), 7.08 (m, 2H), 6.85 (d, \( J = 1.7 \) Hz, 1H), 6.82 (t, \( J = 5.7 \) Hz, 1H), 4.98 (s, 2H), 3.95 (m, 4H), 3.85 (s, 3H), 3.84 (s, 3H), 3.81 (s, 3H), 3.03 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H); \(^{13}\)C NMR: \( \delta \) 167.8, 162.0, 158.44, 158.38, 155.80, 155.77, 155.4, 137.1, 136.0, 133.9, 128.3, 127.7, 127.6, 123.0, 122.7, 122.6, 122.2, 121.2, 120.2, 119.5, 118.6, 118.5, 114.0, 108.4, 104.87, 104.83, 77.7, 65.1, 48.8, 43.5, 38.2, 36.2, 36.13, 36.06, 34.9, 28.2; HRMS (FAB\(^{+}\)) calc’d for C\(_{40}\)H\(_{47}\)N\(_{11}\)O\(_{10}\) [M+H]\(^{+}\) 841.3507, found 841.3498.

3.4.8 BocHN-(R)\(^{\beta}\)-CbzHN\(^{\gamma}\)-ImPyPyPy-Py-(R)\(^{\beta}\)-CbzHN\(^{\gamma}\)-ImPyPyPy-CO\(_2\)Me (12)

A solution of BocHN-(R)\(^{\beta}\)-CbzHN\(^{\gamma}\)-ImPyPyPy-CO\(_2\)H 11 (25 mg, 0.029 mmol) and PyBOP (17 mg, 0.031 mmol) in DMF (150 \( \mu \)L) and DIEA (16 \( \mu \)L, 0.089 mmol) was stirred at 23 ºC for 20 min. The solution was then treated with solid (powdered) HCl•H\(_2\)N-(R)\(^{\beta}\)-CbzHN\(^{\gamma}\)-ImPyPyPy-CO\(_2\)Me 10 (25 mg, 0.031 mmol) and stirred at 23 ºC for 2 h. The solution was then added to distilled H\(_2\)O (10 mL) preacidified with aqueous HCl (1 N, 1 mL, 1 mmol), yielding a precipitate that was isolated by centrifugation (~4500 rpm). The residual solid was again suspended in distilled H\(_2\)O (10 mL) and collected by centrifugation (repeated 3x). The resultant solid, which contained a small amount of
residual H$_2$O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et$_2$O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et$_2$O and dried in vacuo to yield BocHN-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-CO$_2$Me 12 as a tan solid (44 mg, 94%). $^1$H NMR: $\delta$ 10.20 (s, 1H), 10.16 (s, 1H), 9.98 (s, 2H), 9.94-9.91 (m, 4H), 7.99 (m, 1H), 7.46 (d, $J = 1.7$ Hz, 1H), 7.45 (s, 1H), 7.44 (s, 1H), 7.32-7.14 (m, 18H), 7.07 (m, 2H), 7.03 (d, $J = 8.3$ Hz, 1H), 6.92 (s, 1H), 6.90 (d, $J = 2.0$ Hz, 1H), 6.80 (t, $J = 5.6$ Hz, 1H), 4.99 (m, 4H), 4.10 (m, 1H), 3.95 (m, 7H), 3.85-3.83 (m, 15H), 3.79 (s, 3H), 3.73 (s, 3H), $\sim$3.30 (m, 2H, obstructed by H$_2$O peak), 3.04 (m, 2H), 2.53-2.44 (m, 4H, partially obstructed by NMR solvent), 1.36 (s, 9H); $^{13}$C NMR: $\delta$ 167.9, 167.8, 161.6, 160.8, 158.5, 158.44, 155.8, 155.5, 155.5, 137.1, 136.0, 134.00, 133.98, 128.3, 127.7, 127.63, 127.60, 123.11, 123.07, 123.00, 122.80, 122.77, 122.5, 122.3, 122.2, 122.1, 121.3, 120.8, 118.69, 118.66, 118.62, 118.52, 118.0, 114.1, 108.4, 104.9, 104.8, 104.5, 77.8, 65.21, 65.16, 50.9, 48.83, 48.78, 43.6, 42.2, 38.4, 38.2, 36.2, 36.10, 36.07, 36.0, 34.9, 28.2; HRMS (TOF-ESI$^+$) calc’d for C$_{76}$H$_{87}$N$_{22}$O$_{17}$ [M+H]$^+$ 1579.6620, found 1579.6580.

3.4.9 BocHN-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-Py-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-CO$_2$H (13)

A solution of BocHN-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-Py-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-CO$_2$Me 12 (25 mg, 0.0158 mmol) dissolved in 1,4-dioxane (376 μL) and aqueous NaOH (1 N, 253 μL, 0.253 mmol) was stirred at 40 ºC for 4 h. The solution was then added to distilled H$_2$O (5 mL) preacidified with aqueous HCl (1 N, 253 μL, 0.253 mmol), yielding a precipitate that was diluted with another 15 mL of H$_2$O and was then isolated by centrifugation (~4500 rpm). The resultant solid, which contained a small amount of residual H$_2$O, was frozen and lyophilized to dryness and then suspended in excess anhydrous Et$_2$O, triturated, and filtered, and the filter cake washed with copious amounts of anhydrous Et$_2$O. Drying of the tan solid in vacuo yielded BocHN-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-(R)$^{\beta}$-CbzHN$^{\gamma}$-ImPyPyPy-CO$_2$H 13 (23 mg, 93%). $^1$H NMR: $\delta$ 12.13 (br s, 1H), 10.23 (s, 1H), 10.20 (s, 1H), 9.98-9.90 (m, 6H), 8.01 (m, 1H), 7.443 (s, 1H), 7.439 (s, 1H), 7.42 (d, $J = 1.7$ Hz, 1H), 7.30-7.15 (m, 18H), 7.07 (m, 3H), 6.92 (m, 1H), 6.84 (d, $J = 2.0$ Hz, 1H), 6.81 (t, $J = 5.6$ Hz, 1H), 5.00 (m, 2H), 4.98 (s, 2H), 4.10 (m, 1H), 3.95 (m, 7H), 3.85-3.83 (m, 12H), 3.81 (s, 3H), 3.78 (s, 3H), $\sim$3.30 (m, 2H, obstructed by H$_2$O peak), 3.03 (m, 2H), 2.53 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H); $^{13}$C NMR: $\delta$ 167.9, 167.87, 162.4, 162.0, 161.6, 158.54, 158.50, 158.43, 155.8, 155.75, 155.74, 155.6, 155.5, 137.1, 136.0, 133.91, 133.90, 128.3, 127.7, 127.60, 127.57, 123.10, 123.07, 122.8, 122.7, 122.6, 122.3, 122.24, 122.17, 121.16, 121.15, 120.3, 119.5, 118.64, 118.61, 118.5, 118.0, 114.10,
A solution of BocHN-\((\text{R})^\beta\text{-CbzHN}^\gamma\text{-ImPyPyPy-(-R)}^\beta\text{-CbzHN}^\gamma\text{-ImPyPyPy-}\text{CO}_2\text{Pfp}\) (14) in anhydrous CF₃CO₂H:CH₂Cl₂ (1:1, 4 mL) was stirred at 23 ºC for 10 min prior to being concentrated to dryness in a 500 mL round-bottom flask. The residue was then dissolved in cold (0 ºC) DMF (10 mL), followed by immediate dilution with MeCN (300 mL) and DIEA (1.6 mL). The reaction mixture was left at 23 ºC for 3 days without stirring. [Note: The solution turns cloudy as the macrocyclization proceeds.] The reaction mixture was concentrated to a volume of 11 mL and added to a solution of H₂O (30 mL) and aqueous HCl (1 N, 9.2 mL) at 0 ºC. The protected intermediate cyclo-(\(-\text{ImPyPyPy-(-R)}^\beta\text{-CbzHN}^\gamma\text{-ImPyPyPy-(-R)}^\beta\text{-CbzHN}^\gamma\text{-}\)) (15) was isolated by centrifugation (~4500 rpm), lyophilized to dryness, and then suspended in excess anhydrous Et₂O, triturated, and filtered, and the filter cake washed with copious amounts of anhydrous Et₂O. Drying of the tan solid in vacuo yielded the protected intermediate cyclo-(\(-\text{ImPyPyPy-(-R)}^\beta\text{-CbzHN}^\gamma\text{-ImPyPyPy-(-R)}^\beta\text{-CbzHN}^\gamma\text{-}\)) (15), HRMS (TOF-ESI⁺) calc’d for C₇₀H₇₆N₂₂O₁₄ [M+2H]²⁺/2 866.3192, found 866.3236.
was immediately deprotected by dissolving in CF$_3$CO$_2$H (2 mL) followed by addition of CF$_3$SO$_3$H (100 μL) at 23 ºC for 5 min. The solution was then frozen, and DMF (2 mL) was layered over the frozen solution. The thawed solution was diluted with H$_2$O (6 mL) and purified by reverse-phase HPLC to give a white solid after lyophilization. The solid was suspended in excess anhydrous Et$_2$O, triturated, filtered, and the filter cake washed with copious amounts of anhydrous Et$_2$O. Drying of the white solid in vacuo yielded cyclo-(ImPyPyPy-(R)$^\beta$H$_2$N$_\gamma$-ImPyPyPy-(R)$^\beta$H$_2$N$_\gamma$-) 1 (46 mg, 68%). $^1$H NMR: δ 10.56 (s, 2H), 9.91 (s, 4H), 9.88 (s, 2H), 8.17 (t, $J$ = 5.6 Hz, 2H), 7.96 (m, 6H), 7.40 (s, 2H), 7.31 (d, $J$ = 1.6 Hz, 2H), 7.27 (d, $J$ = 1.6 Hz, 2H), 7.19 (d, $J$ = 1.6 Hz, 2H), 7.00 (d, $J$ = 1.7 Hz, 2H), 6.96 (d, $J$ = 1.6 Hz, 2H), 6.94 (d, $J$ = 1.7 Hz, 2H), 3.94 (s, 6H), 3.83 (s, 6H), 3.80 (s, 6H), 3.71 – 3.66 (m, 2H), 3.49 – 3.27 (m, 4H, partially obstructed by H$_2$O peak), 2.79 (dd, $J$ = 16.1 Hz, 6.0 Hz, 2H), 2.60 (dd, $J$ = 15.2 Hz, 5.2 Hz, 2H). HRMS (TOF-ESI+) calc’d for C$_{54}$H$_{63}$N$_{22}$O$_{10}$ [M+H]$^+$ 1179.5098, found 1179.5087.

3.4.12 cyclo-(ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-ImPyPyPy-(R)$^{\beta}$H$^2$N$_\gamma$-) (3) and cyclo-(ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-) (2)

A solution of cyclo-(ImPyPyPy-(R)$^{\beta}$H$_2$N$_\gamma$-ImPyPyPy-(R)$^{\beta}$H$_2$N$_\gamma$-) 1 (2.81 mg, 2.0 μmol) in anhydrous NMP (200 μL) and DIEA (20 μL) at 23 ºC was treated with a solution of Ac$_2$O in NMP (0.122 M, 6.8 μL). After 10 min, the reaction mixture was treated with another 6.8 μL of Ac$_2$O in NMP (0.122 M) and allowed to stand for 5 h. The reaction mixture was then diluted to a volume of 10 mL by addition of a 4:1 solution of aqueous CF$_3$CO$_2$H (0.1% v/v):MeCN (5 mL), followed by additional aqueous CF$_3$CO$_2$H (0.1% v/v, 5 mL), and then purified by reverse-phase HPLC to yield cyclo-(ImPyPyPy-(R)$^{\beta}$H$_2$N$_\gamma$-ImPyPyPy-(R)$^{\beta}$H$_2$N$_\gamma$-) 1 (363 nmol, 18%), cyclo-(ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-ImPyPyPy-(R)$^{\beta}$H$_2$N$_\gamma$-) 3 (800 nmol, 40%), and cyclo-(ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-) 2 (432 nmol, 22%). Cyclo-(ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-ImPyPyPy-(R)$^{\beta}$H$_2$N$_\gamma$-) 3 HRMS (TOF-ESI+) calc’d for C$_{56}$H$_{65}$N$_{22}$O$_{11}$ [M+H]$^+$ 1221.5203, found 1221.5204. Cyclo-(ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-) 2 HRMS (TOF-ESI+) calc’d for C$_{58}$H$_{68}$N$_{22}$O$_{12}$ [M+2H]$^{2+}$/2 633.2646, found 633.2631.

3.4.13 ImPyPyPy-(R)$^{\beta}$H$_2$N$_\gamma$-ImPyPyPy-(+) IPA (4)

Prepared as described in Chapter 2 of this thesis.

3.4.14 ImPyPyPy-(R)$^{\beta}$AcHN$_\gamma$-ImPyPyPy-(+) IPA (5)
A solution of polyamide 4\(^1\) (7.4 mg, 5.03 \(\mu\)moles, assumes 4 as the mono-CF\(_3\)CO\(_2\)H salt) in DMF (1.76 mL) was treated with a solution of Ac\(_2\)O in pyridine (10% v/v, 240 \(\mu\)L, 0.254 mmoles Ac\(_2\)O). The solution was allowed to stand at 23 °C for 30 min, and then acidified with aqueous CF\(_3\)CO\(_2\)H (15% v/v, 2 mL). After 5 min the solution was further diluted with distilled H\(_2\)O (5 mL), purified by preparative RP-HPLC, and lyophilized to dryness. Suspension of the residual solid in anhydrous Et\(_2\)O, following by filtration and drying under high vacuum yielded 5 (6.7 mg, 95%). HRMS (FAB\(^+\)) calc’d for C\(_{67}H_{79}N_{22}O_{13}\) [M+H]\(^+\) 1399.6191, found 1399.6181.

### 3.5 Notes and References


14. For examples of polyamides with $K_a$ values $> 2 \times 10^{10} \text{M}^{-1}$ and a discussion of the limitations of quantitative DNase I footprint titrations, please refer to reference 5g. An analogous polyamide to 4, ImPyPyPy-(R)$^{\beta \alpha}$-$\text{ImPyPyPy-}$,$\gamma$-$\text{ImPyPyPy-Dp}$, was found to have $K_a > 2 \times 10^{10} \text{M}^{-1}$. Additionally, previous studies with cyclic polyamide cyclo-(ImPyPyPy-(R)$^{\alpha \gamma}$-$\text{ImPyPyPy-}$,$\gamma$-$\text{ImPyPyPy}$)$ found $K_a$ values far exceeding $2 \times 10^{10} \text{M}^{-1}$ by DNase I footprint titrations. $^{10c,d}$ Cyclic polyamide 1 possesses dual $\beta$-amino functionalities; a modification that yields even greater DNA binding affinities compared with $\alpha$-amino and unsubstituted $\gamma$-turns for hairpin polyamides of sequence ImPyPyPy-$\gamma$-$\text{ImPyPyPy}$. $^{5e}$ The DNA binding affinity of 1 is most likely superseded that of predecessor cyclo-(ImPyPyPy-(R)$^{\alpha \gamma}$-$\text{ImPyPyPy-}$,$\alpha$-$\text{ImPyPyPy}$-$\gamma$-$\text{ImPyPyPy}$).


18. Aprendica, 313 Pleasant St., Watertown, MA 02472 (http://www.apredica.com/).

3.6 Spectra and Supplemental Information

![Cyclic polyamide 1](image1)

*Denoted as **DMC2-239** in ADMET data tables (Table 3.2-3.7) and in full ADMET report located in Appendix B*

![Hairpin polyamide 5](image2)

*Denoted as **DH-V-88** in ADMET data tables (Table 3.2-3.7) and in full ADMET report located in Appendix B*

**Figure 3.3** Polyamides 1 and 5 were subjected to preclinical ADMET testing by contract service at Aprelica (Watertown, MA). Shown below (Table 3.2-3.7) are summaries of the ADMET results taken directly from the final report provided by Aprelica. The full ADMET report, which includes experimental conditions, can be found in Appendix B of this thesis.

**Table 3.2** Caco-2 permeability summary from Aprelica report (Appendix B).

<table>
<thead>
<tr>
<th>Client ID</th>
<th>test conc (µM)</th>
<th>Assay duration (h)</th>
<th>mean A-&gt;B P&lt;sub&gt;app&lt;/sub&gt; (10&lt;sup&gt;-6&lt;/sup&gt; cm&lt;sup&gt;2&lt;/sup&gt; s&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>mean A-&gt;B P&lt;sub&gt;app&lt;/sub&gt; (10&lt;sup&gt;-6&lt;/sup&gt; cm&lt;sup&gt;2&lt;/sup&gt; s&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>Asymmetry ratio&lt;sup&gt;a&lt;/sup&gt;</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warfarin</td>
<td>50</td>
<td>2</td>
<td>35.4</td>
<td>7.9</td>
<td>0.2</td>
<td>high permeability control</td>
</tr>
<tr>
<td>Ranitidine</td>
<td>50</td>
<td>2</td>
<td>ND</td>
<td>ND</td>
<td>1.7</td>
<td>low permeability control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10</td>
<td>2</td>
<td>ND</td>
<td>ND</td>
<td>0.11</td>
<td>UD</td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10</td>
<td>2</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

<sup>a</sup>Apparent permeability  
<sup>b</sup>P<sub>app(B->A) / P<sub>app(A->B)</sub>  
ND = no compound detected in receiver solution

**Table 3.3** Cytotoxicity summary from Aprelica report (Appendix B).

<table>
<thead>
<tr>
<th>Client ID</th>
<th>Cell line</th>
<th>IC50 (µM)</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorpromazine</td>
<td>HepG2</td>
<td>13</td>
<td>Highly cytotoxic control</td>
</tr>
<tr>
<td>Propranolol</td>
<td>HepG2</td>
<td>80</td>
<td>Low cytotoxic control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>HepG2</td>
<td>&gt;100</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>HepG2</td>
<td>&gt;100</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.4 Fluorescent Cyp IC50 summary from Apredica report (Appendix B).

<table>
<thead>
<tr>
<th>Client ID</th>
<th>IC50 (µM)</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-4031</td>
<td>99% at 0.5</td>
<td></td>
</tr>
<tr>
<td>DH-V-88</td>
<td>&gt;100</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>&gt;100</td>
<td></td>
</tr>
</tbody>
</table>

*The solubility limit for this experiment, as determined by vehicle controls, was 17.3 x 103 LSU (horizontal black line). Based on the data obtained, there may be solubility issues for both test articles at 30 and 100 µM in our physiological saline solution (HB-PS, 0.3%DMSO). Precipitation of DH-V-88 at 100 µM was visible to the naked eye.

Table 3.5 hERG FastPatch summary from Apredica report (Appendix B).

<table>
<thead>
<tr>
<th>Compound</th>
<th>test conc (µM)</th>
<th>medium</th>
<th>T1/2 (min)</th>
<th>Fraction remaining, max time (%)</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propantheline</td>
<td>10.0</td>
<td>Human Plasma</td>
<td>35.5</td>
<td>5.8%</td>
<td>control</td>
</tr>
<tr>
<td>Propantheline</td>
<td>10.0</td>
<td>Rat Plasma</td>
<td>149.0</td>
<td>51.6%</td>
<td>control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10.0</td>
<td>Human Plasma</td>
<td>&gt;120</td>
<td>95.6%</td>
<td></td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10.0</td>
<td>Plasma Human</td>
<td>&gt;120</td>
<td>94.0%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10.0</td>
<td>Plasma Human</td>
<td>&gt;120</td>
<td>124.5%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10.0</td>
<td>Plasma Rat</td>
<td>&gt;120</td>
<td>120.3%</td>
<td></td>
</tr>
</tbody>
</table>

*Half-life

Table 3.7 Plasma protein binding summary from Apredica report (Appendix B).

<table>
<thead>
<tr>
<th>Client ID</th>
<th>test conc (µM)</th>
<th>Assay duration</th>
<th>Species</th>
<th>Mean free fraction (%)</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warfarin</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>0.73%</td>
<td>high binding control</td>
</tr>
<tr>
<td>Warfarin</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>5.47%</td>
<td>high binding control</td>
</tr>
<tr>
<td>Atenolol</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>76.2%</td>
<td>low binding control</td>
</tr>
<tr>
<td>Atenolol</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>84.7%</td>
<td>low binding control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>0.0015%</td>
<td></td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>0.0016%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>0.0000%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>0.0040%</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.4 $^1$H NMR BocHN-$(R)^{\beta}$H CHN$\gamma$-Im-CO$_2$H (8)
Figure 3.5 $^{13}$C NMR BocHN-(R)$^{\beta}$-Chl$^{\gamma}$Im-CO$_2$H (8)
Figure 3.6 $^1$H NMR BocHN-$(R)^\beta$-ChaH$_2$N$_2$-ImPyPyPy-CO$_2$Me (9)
Figure 3.7 $^{13}$C NMR BocHN-(R)$^{6}$-Ch$_2$N$_2$γ-ImPyPyPy$_2$Me$_2$ (9)
Figure 3.8 $^1$H NMR HCl•H$_2$N-(R)$_{\beta}$ChlH$_N$$\gamma$-ImPyPyPy-CO$_2$Me (10)
Figure 3.9 $^{13}$C NMR HCl•H$_2$N-$(R)^\text{ChdIN}_{\gamma}$-ImPyPyPy-CO$_2$Me (10)
Figure 3.10: H NMR BocHN-(R)-<sup>-</sup>CbzHN-<sup>-</sup>ImPyPyPy-CO<sub>2</sub>H (11)
Figure 3.11 $^{13}$C NMR BocHN-(R)$^{\beta}$ChHN$_{\gamma}$-ImpPyPyPy-CO$_2$H (11)
Figure 3.12 $^1$H NMR BocHN-(R)$^{\beta}$-CbzHN,$^{\gamma}$-ImPyPyPy-(R)$^{\beta}$-CbzHN,$^{\gamma}$-ImPyPyPy-CO$_2$Me (12)
Figure 3.13 $^{13}$C NMR BocHN-$(R)^{\beta}$-CbzHN$_{\gamma}$-ImPyPyPy-$(R)^{\beta}$-CbzHN$_{\gamma}$-ImPyPyPy-CO$_2$Me (12)
Figure 3.14 $^1$H NMR BocHN-(R)$^\beta$-CholN$\gamma$-ImPyPyPy-(R)$^\beta$-CholN$\gamma$-ImPyPyPy-CO$_2$H (13)
Figure 3.15 $^{13}$C NMR BocHN-(R)$^{\beta}$-CbzHN$_{\gamma}$-ImpyPyPyPy-(R)$^{\beta}$-CbzHN$_{\gamma}$-ImpyPyPyPy-CO$_2$H (13)
Figure 3.16 $^1$H NMR BocHN-(R)$^\beta$-Chol$^\gamma$-ImPyPyPy-(R)$^\beta$-Chol$^\gamma$-ImPyPyPy·CO$_2$Pfp (14)
Figure 3.17 $^1$H NMR cyclo-(ImPyPyPy-$(R)^{\beta,H2N}_{\gamma}$-ImPyPyPy-$(R)^{\beta,H2N}_{\gamma}$) (1)
Analytical HPLC characterization of cyclic polyamide 15.

Figure 3.18 Analytical HPLC characterization of cyclic polyamide 15.

Analytical HPLC characterization of cyclic polyamide 1.

Figure 3.19 Analytical HPLC characterization of cyclic polyamide 1.

Analytical HPLC characterization of cyclic polyamide 3.

Figure 3.20 Analytical HPLC characterization of cyclic polyamide 3.
**Figure 3.21** Analytical HPLC characterization of cyclic polyamide 2.

**Figure 3.22** Analytical HPLC characterization of cyclic polyamide 5.
Chapter 4: Oligomerization Route to Polyamide Macroycles

The text of this chapter was taken in part from a manuscript coauthored with Daniel A. Harki and Peter B. Dervan* (Caltech)

(Chenoweth, D.M., Harki, D.A., Dervan, P. B. In Preparation.)
Abstract

Cyclic pyrrole–imidazole polyamides are sequence-specific DNA-binding small molecules that are cell permeable and can regulate endogenous gene expression. Syntheses of cyclic polyamides have been achieved by solid-phase and more recently, solution-phase methods. We report a rapid solution-phase oligomerization approach to cyclic polyamides that yields 8, 12, and 16 membered macrocycles. A remarkable preference for DNA binding by the 8 and 16 membered oligomers was observed over the 12-ring macrocycle, which we attributed to a conformational constraint not present in the smaller and larger systems.
4.1 Introduction

Pyrrole–imidazole polyamides are a class of cell-permeable small molecules that bind to the minor groove of DNA in a sequence specific manner.\textsuperscript{1,2} Antiparallel arrangements of \textit{N}-methylpyrrole (Py) and \textit{N}-methylimidazole (Im) carboxamides (Im/Py) recognize G•C from C•G base pairs, whereas Py/Py specifies for both T•A and A•T.\textsuperscript{3} Hairpin Py-Im polyamides have been programmed for a broad repertoire of DNA sequences with high affinities.\textsuperscript{1} These cell permeable\textsuperscript{5} ligands can influence gene transcription by disrupting protein-DNA interfaces,\textsuperscript{2} and have been shown to control transcription of genes important in human disease.\textsuperscript{6} Py-Im polyamides have also been used for a variety of applications ranging from DNA detection\textsuperscript{7} and transcriptional activation\textsuperscript{8} to the self-assembly of higher order structures.\textsuperscript{9} Conjugation of polyamides to functional domains have yielded artificial transcription factor mimics,\textsuperscript{8} fluorescent sequence-specific DNA probes,\textsuperscript{7} and DNA nano-architectures.\textsuperscript{9}

We recently reported solution-phase methods for the synthesis of hairpin\textsuperscript{10} and cyclic polyamides.\textsuperscript{11} Key to the cyclic polyamide synthesis was a highly efficient macrocyclization that yielded polyamide 1z (Figure 4.1) from an acyclic precursor. Activation of the C-terminal Py amino acid of 1z as a pentafluorophenyl ester allowed efficient macrocyclization by the γ-NH\textsubscript{2} on the turn moiety under dilute reaction conditions. Our studies of polyamide 1, the deprotected analogue of 1z, revealed it possessed extremely high DNA binding affinities, was cell permeable, and could disrupt androgen receptor-activated gene expression in cell culture. Additionally, preliminary studies of the in vitro ADMET properties of 1 revealed excellent metabolic stability.\textsuperscript{11,12}

An orthogonal polymerization/oligomerization strategy for the synthesis of 1 and related polyamides is reported here. This method delivers symmetrical Py-
Im polyamide macrocycles from simple Py-Im building blocks in a rapid and convergent manner (Figure 4.2). Separable higher-order oligomers, such as the 12-membered (2) and 16-membered (3) cyclic polyamides are also produced by this method. In addition to describing the synthetic chemistry to prepare 1–3, we report for the first time the DNA binding properties of such expanded polyamide macrocycles.

4.2 Results and Discussion

Our strategy for this oligomerization route relied on the palindromic nature of polyamide 1. Disconnection of 1 at both γ-amino turns affords two identical halves of the molecule. Bimolecular coupling between two molecules, followed by intramolecular ring closure delivers cyclic Py-Im polyamides. Bifunctional monomer 4 contains every atom needed to construct cyclic polyamides 1–3 by this process (Figure 4.2).

The pentafluorophenyl ester 4 was prepared in one step from the previously reported carboxylic acid of 4.11 Acidic deprotection of the γ-amino functionality of 4 followed by drying in vacuo yields intermediate 5 which is the substrate for the homodimerization/oligomerization reaction. To initiate this sequence, the protected trifluoroacetate salt 5 was diluted with DMSO, then treated with an organic base (DIEA) to unmask the highly nucleophilic primary γ-amine. The ensuing oligomerization/macrocyclization process provides benzylcarbamate protected cyclic polyamides 1z, 2z, 3z, and trace amounts of unisolated higher-order oligomers. A distribution of uncyclized intermediates corresponding to the dimer (8-ring cycle, 1z), trimer (12-ring cycle, 2z), tetramer (16-ring cycle, 3z), and higher order adducts can be observed at early time points, as evidenced by HPLC analysis at 2 hr (Figure 4.3). Extended
reaction times (20 hours) reveals cyclized polyamides $1z$, $2z$, and $3z$ in a ratio of 6.6:2.6:1 almost exclusively (Figure 4.3). Isolation of $1z$ (13.9%), $2z$ (5.5%), and $3z$ (2.1%) by preparative HPLC, followed by Cbz-deprotection under acidic conditions ($\text{CF}_3\text{CO}_2\text{H}/\text{CF}_3\text{SO}_3\text{H}$) provides polyamide macrocycles 1–3.

With polyamide macrocycles 2 and 3 in hand we evaluated their ability to bind duplex DNA relative to previously reported cycle 1. Polyamide 1 increases the dsDNA melting temperature by 23.6 °C. To our surprise, trimeric macrocycle 2 failed to bind its target double stranded DNA sequence as evidenced by the complete lack of ligand-promoted thermal stabilization of duplex DNA melting (Table 4.1). This result is presumably due to inherent geometrical constraints of 2, preventing the side-by-side antiparallel alignment of the PyPyPyIm strands, a motif well accommodated by the DNA minor groove. In the case of tetrameric macrocycle 3, dsDNA binding and thermal stabilization was completely restored to a comparable value to dimer 1. We hypothesize that an even number of PyPyPyIm strands allows 3 to possess a collapsed or folded tetramer geometry, with two adjacent, antiparallel PyPyPyIm strands followed by an identical repeat of this motif linked through two intervening turn units.

### 4.3 Conclusion

In summary, we have demonstrated that macrocyclic polyamides can by synthesized by
oligomerization of a bifunctional polyamide to yield a distribution of cyclic polyamide oligomers. Additionally, we show that certain cyclic polyamide geometries are completely devoid of the ability to bind dsDNA, a result which could be utilized in the design of highly specific molecules for targeting non-B-form DNA structures or other higher-order nucleic acid motifs.

4.4 Experimental Section

4.4.1 General

Chemicals were purchased from Sigma-Aldrich and were used without further purification. (R)-3,4-Cbz-Dbu(Boc)-OH was purchased from Senn Chemicals AG (code number 44159). Bulk grade solvents were from Fisher Scientific. Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250 × 4.6 mm, 5 μm), a diode array detector, and the mobile phase consisted of a gradient of acetonitrile (MeCN) in 0.1% (v/v) aqueous CF₃CO₂H. Preparative HPLC was performed on an Agilent 1200 system equipped with a solvent degasser, diode array detector, and a Phenomenex Gemini column (250 × 21.2 mm, 5 μm). A gradient of MeCN in 0.1% (v/v) aqueous CF₃CO₂H was utilized as the mobile phase. NMR spectroscopy was performed on a Varian instrument operating at 499.8 MHz (for ¹H) at ambient temperature. All NMR analyses were performed in DMSO-d₆, and chemical shifts are reported in parts per million relative to the internal solvent peak referenced to 2.49 (for ¹H). High-resolution mass spectrometry (HRMS) was recorded in positive-ion mode by fast-atom bombardment (FAB+) on a JEOL JMS-600H instrument or by matrix-assisted, LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) on an Applied Biosystems Voyager DE-Pro spectrometer.
using α-cyano-4-hydroxycinnamic acid as matrix.

4.4.2 BocHN-(R)β-ChdHNγ-ImPyPyPy-CO₂H
Synthesized as previously described in Chapter 3 of this thesis.

4.4.3 BocHN-(R)β-ChdHNγ-ImPyPyPy-CO₂Pfp (4)
A solution of BocHN-(R)β-ChdHNγ-ImPyPyPy-CO₂H (100 mg, 0.119 mmol) and DCC (49 mg, 0.238 mmol) in CH₂Cl₂ (5.2 mL) was stirred at 23 °C for 45 min. The solution was then treated with DMAP (1.4 mg, 0.012 mmol) followed by pentafluorophenol (131.2 mg, 0.713 mmol) and stirred at 23 °C for 12 h. The reaction mixture was then loaded onto a silica gel column with CH₂Cl₂ and eluted with step gradients of 100% CH₂Cl₂ to 100% acetone with incremental steps of 5% acetone. The product was concentrated in vacuo to yield BocHN-(R)β-ChdHNγ-ImPyPyPy-CO₂Pfp (4) as an off-white solid (84 mg, 71%). ¹H NMR (500 MHz, DMSO-d₆): δ 10.16 (s, 1H), 10.08 (s, 1H), 9.99 (s, 1H), 9.97 (s, 1H), 7.74 (d, J = 1.8 Hz, 1H), 7.44 (s, 1H), 7.33 – 7.30 (m, 5H), 7.29 (d, J = 1.9 Hz, 1H), 7.27 (d, J = 1.8 Hz, 1H), 7.25 (d, J = 1.8 Hz, 1H), 7.14 (d, J = 1.8 Hz, 1H), 7.13 (d, J = 1.8 Hz, 1H), 7.03 (d, J = 8.4 Hz, 1H), 6.80 (t, J = 5.8 Hz, 1H), 4.98 (s, 2H), 3.95 (m, 4H), 3.89 (s, 3H), 3.860 (s, 3H), 3.856 (s, 3H), 3.03 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H). HRMS (FAB+) calc’d for C₄₆H₄₇N₁₁O₁₀F₅ [M+H]+ 1008.343, found 1008.342.

4.4.4 Oligomerization procedure
A glass vial (1 dram) was charged with 4 (5.0 mg, 4.96 μmole) and treated with a solution of CF₃CO₂H in CH₂Cl₂ (1:1 CF₃CO₂H:CH₂Cl₂, 1 mL) and stirred at 23 °C for 10 min. The solvent was removed in vacuo and the residual solid was dried under high vacuum for 20 min. The solid was diluted with DMSO (500 μL) followed by DIEA (80 μL) and the solution was stirred at 23 °C for 20 h. After 20 h the reaction was complete by analytical HPLC analysis. The reaction was diluted to a final volume of 10 mL by addition of a solution of DMF in aqueous CF₃CO₂H (2:3 DMF:0.1% aqueous CF₃CO₂H). NOTE: A small amount of yellow insoluble material was observed and discarded. Purification by RP-HPLC yielded 1z (13.9% yield), 2z (5.5% yield), and 3z (2.1% yield). The yield of 1z is calculated from the mass of the purified and isolated material (0.5 mg). Yields for 2z and 3z were calculated based on 1z using the relative product distribution as measured by integration of the preparative HPLC chromatogram at 310 nm (product distribution: 6.6:2.6:1.0 ratio of 1z:2z:3z; UV integral values were normalized to the number of ImPyPyPy strands contained
in each cyclic oligomer). The benzyl carbamate (Cbz) protecting groups of 1z-3z were removed as previously described. Characterization data for dimer 1 has been reported previously in Chapter 3 of this thesis. Trimer 2 HRMS (MALDI-TOF) calc’d for C_{81}H_{94}N_{33}O_{15} [M+H]⁺ 1768.7607, found 1768.7566. Tetramer 3 HRMS (MALDI-TOF) calc’d for C_{108}H_{125}N_{44}O_{20} [M+H]⁺ 2358.0112, found 2358.0143.

4.4.5 UV Absorption Spectrophotometry
Melting temperature analysis was performed on a Varian Cary 100 spectrophotometer equipped with a thermo-controlled cell holder possessing a cell path length of 1 cm. A degassed aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂ at pH 7.0 was used as analysis buffer. DNA duplexes and hairpin polyamides were mixed in 1:1 stoichiometry to a final concentration of 2 μM for each experiment. Prior to analysis, samples were heated to 90 °C and cooled to a starting temperature of 23 °C with a heating rate of 5 °C/min for each ramp. Denaturation profiles were recorded at λ = 260 nm from 23 °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.

4.5 References and Notes


12. Compound 1 targets the DNA sequence 5’-WGWWCW-3’, where W =A/T or T/A.
Figure 4.4 $^1$H NMR BocHN-($R$)-$\beta$-Chl$\text{N}$_$\gamma$-ImPyPyPy-CO$_2$Pfp (4)
Chapter 5: Allosteric Modulation of DNA by Small Molecules
Abstract

Many human diseases are caused by dysregulated gene expression. The oversupply of one or more transcription factors may be required for the growth and metastatic behavior of human cancers. Cell permeable small molecules which can be programmed to disrupt transcription factor-DNA interfaces could silence aberrant gene expression pathways. Pyrrole-imidazole polyamides are DNA minor-groove binding molecules that are programmable for a large repertoire of DNA motifs. A high resolution X-ray crystal structure of an 8-ring cyclic Py/Im polyamide bound to the central six base pairs of the sequence d(5'-CCAGGC_CTGG-3'), reveals a 4 Å widening of the minor groove and compression of the major groove along with a >18° bend in the helix axis toward the major groove. This allosteric perturbation of the DNA helix provides a molecular basis for disruption of transcription factor-DNA interfaces by small molecules, a minimum step in chemical control of gene networks.
5.1 Introduction

Py/Im polyamides bind the minor groove of DNA sequence specifically,1,2 encoded by side-by-side arrangements of N-methylpyrrole (Py) and N-methylimidazole (Im) carboxamide monomers. Im/Py pairs distinguish G•C from C•G base pairs, whereas Py/Py pairs are degenerate for T•A and A•T.3,6 Antiparallel Py/Im strands are connected by γ-aminobutyric acid linker (GABA) to create a hairpin-shaped oligomer. Hairpin Py/Im polyamides have been programmed to bind a broad library of different DNA sequences. They have been shown to permeate cell membranes,8-10 access chromatin,11,12 and disrupt protein-DNA interactions.2 Disruption of transcription factor-DNA interfaces six bp in size such as HIF-1α,13-15 androgen receptor (AR),16 and AP-117,18 have been exploited for controlling expression of medically relevant genes such as VEGF, PSA, TGF-β1 and LOX-1 in cell culture experiments.13-18 X-ray crystallography of antiparallel 2:1 binding polyamides in complex with DNA reveal a 1 Å widening of the minor groove.5,6 This modest structural perturbation to the DNA helix by the side-by-side stacked arrangement of aromatic rings does not explain the large number of transcription factor-DNA interfaces disrupted by minor-groove binding hairpin Py/Im polyamides.2,5,6,19 It must be that the turn unit in the hairpin oligomer connecting the two antiparallel strands plays a structural role.

Here we report the atomic resolution structure (1.18 Å resolution) of an 8-ring cyclic polyamide in complex with double helical DNA. The cyclic polyamide 1 is comprised of two antiparallel ImImPyPy strands capped by (R)-α-amino-γ turn units. Polyamide 1, which codes for the sequence 5’-WGGCCW-3’ was co-crystallized with the palindromic DNA oligonucleotide sequence 5’-CCAGGCCTGG-3’ 10 base pairs in length (Figure 5.1). We observe significant structural allosteric perturbations of the DNA helix induced upon binding of GABA (γ-aminobutyric acid) turn-linked polyamides in the minor groove. A detailed view of the α-amino-γ-turn conformation and hydration

![Diagram of cyclic polyamide 1 and DNA sequence](image-url)
reveal a network of well-ordered water-mediated interactions between the polyamide and the minor groove floor of DNA.

5.2 Results and Discussion

The structure was solved by direct methods to 1.18 Å resolution with synchrotron radiation. One cyclic polyamide bound to a single DNA duplex is present in the asymmetric unit of the crystal in the P1 space group. In the DNA complex the aromatic amino acids are bound with an N- to C- orientation of each ImImPyPy strand of the cycle adjacent to the 5’ to 3’ direction of the DNA. To assess DNA structural perturbations imposed by polyamide binding we compared our polyamide-DNA complex to the free DNA. We solved the X-ray structure of d(CCAGGC\textsubscript{2}CTGG\textsubscript{2}) to a resolution of 0.98 Å for comparison.\textsuperscript{20} Interestingly, this structure shows several discrete alternate conformations in 7 of the 10 nucleotides in each strand of the DNA duplex, illustrating the dynamic and conformationally mobile nature of the B-DNA sugar-phosphate backbone. The comparison of polyamide-DNA complex to free DNA is shown in Figure 5.2. Polyamide binding locks out the alternate DNA conformations, rigidifying the sugar-phosphate backbone, and strongly perturbing the overall helix structure. Binding of the polyamide widens the minor groove up to 4 Å while simultaneously compressing the major

Figure 5.2 Comparison of native DNA to polyamide/DNA complex. a) Native DNA crystal structure at 0.98 Å resolution. b) Comparison to DNA/polyamide co-crystal structure at 1.18 Å resolution. (Both structure solved by direct methods.) showing the bound cyclic polyamide with electron density contoured at the 1.0 \(\sigma\) level.
groove by 4 Å. The polyamide bends the DNA strand >18º toward the major groove as shown in Figure 5.3, and shortening of the overall length of the helix by ~1 Å (Figure 5.2).

Py/Im Polyamides linked by a GABA or substituted GABA can adopt either of two possible conformations shown in Figure 5.4. In conformation A, the amino group is directed toward the minor-groove wall of DNA with the potential for steric clash with the deoxyribose sugar. In conformation B the amine is directed up and out of the minor groove forcing the β-methylene to the floor of the minor groove with the potential for steric interaction with the nucleobases within van der Waals contact distance of the C2 hydrogen of adenine. We observe conformation B in our high resolution X-ray structure at both ends of the complex (Figure 5.4). It is possible that there is an intrinsic preference for conformation A, which relieves the β-methylene interaction with the floor of the minor groove.

**Figure 5.3** Analysis of native DNA (yellow) compared to polyamide complexed DNA (blue). Chart on the top left shows variation in the minor groove width for native DNA (yellow) and polyamide-complexed DNA (blue) over the central core sequence 5'-GGCC-3’. Chart on the bottom left shows variation in the major groove width for native DNA (yellow) and polyamide complexed DNA (blue) over the central core sequence 5'-GGCC-3’. Overlay of the Curves calculated geometric helix model from each structure showing a DNA bend of > 18º in the polyamide/DNA complex compared to native DNA.
However, for turn substitution at the $\alpha$-position interaction with the minor-groove wall becomes the dominant steric interaction, leading to conformational inversion. Figure 5.4 presents a view of the complex looking down the minor groove directly at the polyamide turn linkage. From this view, significant van der Waals interactions can be observed between the outside face of the pyrrole-imidazole strands and the walls of the minor groove, which form a deep binding pocket for the cycle. Approximately 40% of the polyamide surface area is buried leaving only the top of the methyl groups on the heterocycles, the amide carbonyl oxygens, and the chiral $\alpha$-ammonium turn solvent exposed. In addition the turn unit introduces conformational constraints preventing slipped or linear binding modes.

The conformational constraints imposed by the turn linkage result in ring placement that is an intermediate of ring-over-ring and ring-over-amide between adjacent PyPyImIm strands. This alignment allows the ring pairs to remain in phase with the nucleobases as the polyamide adopts an isohelical conformation complementary to the DNA helix. This is highlighted by comparison to the 2:1 structure in which the rings lie over the carboxamide linkages of the adjacent strand. The conformational constraints

![Figure 5.4 Conformation of the $\alpha$-amino substituted GABA turn. (top), Two possible Conformations A and B are shown with conformation A directing the $\beta$-methylene up and away from the minor groove floor while orienting the $\alpha$-ammonium toward the minor groove wall. Conformation B presents the $\beta$-methylene down toward the minor-groove floor while orienting the $\alpha$-ammonium up and out of the minor-groove, relieving possible steric interaction with the sugar-phosphate backbone (minor-groove wall). (bottom), View looking down the DNA minor-groove, showing the ($R$)-$\alpha$-amine-$\gamma$-turn conformation observed in the X-ray crystal structure, which matches that of conformation B. Electron density map is contoured at the 1.0 $\sigma$ level.](image-url)
imposed by the turn and inability of the ligand to slip into a possibly more preferred orientation may impact the overall DNA structure by inducing bending and other allosteric distortions accommodated by the plasticity of DNA. In addition, the turn-constrained cycle may have a major entropic driving force leading to substantial pre-organization, increased affinity, and increased specificity by locking out unproductive conformations and alternate binding modes. The van der Waals interactions between rings may also lead to cooperativity in the binding process. In addition, we find a shell of highly ordered water molecules around the α-ammonium substituent and a water-mediated hydrogen bond from the ammonium to the N3 lone-pair of the adenine under the turn. The hydration pattern around the turn is highly conserved at both ends of the structure and the water mediated hydrogen bonds are within ~2.7–2.9 Å from the ammonium to water to the adenine lone-pair (Figure 5.5a).

The amide NH's and imidazole lone-pairs form a continuous series of direct hydrogen bonds to the floor of the DNA minor-groove, while the imidazoles impart specificity for the exocyclic amine of guanine through relief of a steric interaction and a G(N2-hydrogen)-Im (lone pair) hydrogen bond. The amides linking the aromatic rings and the turns contribute hydrogen bonds to the purine N3 and pyrimidine O2 lone pairs. All amides are within hydrogen bonding distance of a single DNA base (~ 3.0 Å average, Figure 5.9). In all there are 10 direct amide hydrogen bonds (average distance = 2.97 Å), 4 direct imidazole hydrogen bonds (2 terminal average distance = 3.27 Å and 2 internal average distance = 3.05 Å), and 2 (R)-α-ammonium turn water-mediated hydrogen bonds (average distance amine to water = 2.75 Å and average distance from the water to adenine N3 = 2.98 Å) to the floor of the DNA minor groove with at least one interaction for all 12 DNA base-pairs in the 6 bp binding site for a total of 16 hydrogen bond interactions between the cyclic polyamide and the floor of the DNA minor-groove. These 16 hydrogen bonds utilize every heteroatom of the polyamide presented to the floor of the DNA minor-groove, which exactly matches the total number of Watson-Crick hydrogen bonds between all the DNA base pairs in the 6 bp binding site. In addition to these 16 hydrogen bonds, we find unique weak interactions in the form of lone pair-π interactions21,22 between the center of the leading imidazole ring and the lone pair of the adjacent deoxyribose O4’ oxygen (Figure 5.5d and 5.5e). Interestingly this interaction is only observed for the terminal imidazole aromatic ring and analysis of qualitative electrostatic potential surfaces substantiates the electropositive nature of the imidazole under these conditions (Figure 5.13).23
Figure 5.5 Direct and water-mediated non-covalent molecular recognition interactions. a) Geometry of the \(\alpha\)-amino turn interacting with the AT base pair through water-mediated hydrogen bonding interactions. Structural basis for the turn preference for AT versus GC is demonstrated by the \(\beta\)-methylene conformational preference, which points down toward the DNA minor-groove floor within van der Waals contact distance of the adenine base. b) Isolated view of one half of the macrocyclic-polyamide showing hydrogen bond distances made to the DNA minor groove floor by the imidazoles and amides of compound 1. c) Im-Py pair showing the mechanism for GC specificity. d) Interaction of the O4’ oxygen of a deoxyribose sugar with the terminal imidazole aromatic ring through a lone pair-π interaction. The sugar conformation is C2’-endo at the N-terminal imidazole of the polyamide with the sugar oxygen lone-pair pointing directly to the centroid of the imidazole ring. The distance between the sugar oxygen and the ring centroid is 2.90 Å, which is less than the sum of the van der Waals radii to any atom in the imidazole ring. Electrostatic potential maps calculated at the HF/3-21g* level of theory show the slightly electropositive nature of the imidazole ring under these conditions (Figure 5.13). e) View of the O4’ deoxyribose oxygen atom looking through the imidazole ring showing the ring centroid superimposed on the oxygen atom. All distances are reported in angstroms (Å) and all electron density maps are contoured at the 1.0 \(\sigma\) level (Im = imidazole and Py = pyrrole).

5.3 Conclusion

The crystal structure presented highlights the molecular recognition of turn-linked polyamides in the minor-groove of DNA and provides insight into the allosteric modulation of
B-form DNA by hairpin oligomers. The DNA structural distortion induced upon polyamide minor-groove binding provides an allosteric model for disrupting DNA:transcription factor interfaces in the promoters of selected genes. The ability of DNA to undergo short and long-range allosteric effects coupled with DNA binding by proteins can have influence over important processes such as modulation of eukaryotic gene networks. Allosteric communication along and through the DNA helix forms the basis for cooperative interactions among transcription factor regulatory networks such as the interferon-β enhanceosome. The potential for allosteric control over the transcriptional machinery provides a powerful concept for the design of small molecules that can bind to distinct locations on DNA with the possibility of modulating transcription factor activity.

5.4 Experimental Section

5.4.1 General
Chemicals and solvents were purchased from Sigma-Aldrich and Hampton Research and were used without further purification. Water (18 MΩ) was purified using a Millipore MilliQ purification system. Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250 x 4.6 mm, 5 μM), a diode array detector, and the mobile phase consisted of a gradient of acetonitrile (MeCN) in 0.1% (v/v) aqueous trifluoroacetic acid (TFA). Preparative HPLC was performed on an Agilent 1200 system equipped with a solvent degasser, diode array detector, and a Phenomenex Gemini column (5 μm particle size, C18 110A, 250 x 21.2 mm, 5 micron). A gradient of MeCN in 0.1% (v/v) aqueous trifluoroacetic acid (TFA) was utilized as the mobile phase. UV-Vis measurements were made on a Hewlett-Packard diode array spectrophotometer (model 8452 A) and polyamide concentrations were measured in 0.1% (v/v) aqueous TFA using an extinction coefficient of 69200 M⁻¹cm⁻¹ at λ max near 310 nm. Matrix-assisted LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on an Applied Biosystems Voyager DR Pro spectrometer using α-cyano-4-hydroxycinnamic acid as matrix.

5.4.2 Synthesis
Polyamide 1 was synthesized by standard solid-phase synthesis methods on oxime resin (Figure 5.6) and purified by reverse-phase high-performance liquid chromatography (Figure 5.7).

5.4.3 Cyclo-(-ImImPyPy-(R)α-BocHNγ-ImImPyPy-(R)α-H2Nγ-) (6)
Oxime resin (R3) was generated by manual solid-phase synthesis from Kaiser oxime resin (1 g, 0.48 mmol/g, Novabiochem) using previously described Boc-protected monomers. Boc-Py-OBt (2) (716 mg, 2 mmol) was dissolved in 2 mL of DMF and added to 1 g of oxime resin followed by 1 mL of DIEA. The reaction was left in a 37 °C shaker for 12 h. The resin was drained, washed with DMF 3x, DCM 3x, and the Boc group was removed upon treatment with 20% TFA/DCM for 30 min. After draining the resin and washing with DCM 3x followed by DMF 3x the second pyrrole residue was coupled in the same fashion as the first, with complete coupling after 2 h at 23 °C. Boc-Im-OH (3) (482 mg, 2 mmol) was dissolved in 2 mL of DMF and treated with 1.14 g (2 mmol) of PyBOP and 2 mL of DIEA. This solution was stirred for 5 minutes prior to addition to the resin vessel. Coupling was allowed to proceed for 2 h at 23 °C. The Boc-Im residue was deprotected using a 50% TFA/DCM solution for 30 min at room temperature followed by draining the resin and washing with DCM 3x and DMF 3x. A second imidazole residue was coupled following the exact same procedure as the first. The turn unit, α-Fmoc-γ-Boc-(R)-diaminobutyric acid (Fmoc-D-Dab-(Boc)-OH) (4) (660 mg, 1.5 mmol) was activated with PyBOP (855 mg, 1.5 mmol) in 2 mL of DMF and 1 mL of DIEA at 23 °C for 15 min prior to addition to the resin. Coupling was allowed to proceed for 2 h at 37 °C. After deprotection with 20% TFA/DCM for 30 min, the next two pyrrole residues were attached in exactly the same manner as previously described using Boc-Py-OBt. The last two imidazoles were added in the same fashion as the previous two. The final turn unit, α-Boc-γ-Fmoc-(R)-diaminobutyric acid (Boc-D-Dab-(Fmoc)-OH) (5) (660 mg, 1.5 mmol) was activated with PyBOP (855 mg, 1.5 mmol) in 2 mL of DMF and 1 mL of DIEA at 23 °C for 15 min prior to addition to the resin. Coupling was allowed to proceed for 2 h at 37 °C. After Fmoc deprotection with 25% piperidine/DMF for 3x5 min, the resin was washed with DMF 6x and 1:1 DMF/DIEA 3x. Next, the resin was diluted with 10 mL of DMF and stored in a 37 °C shaker on medium speed for 24 h. The resin was filtered off and the DMF concentrated to 1 mL volume, taken up in 9 mL of H_2O (0.1% TFA), and purified by preparative reverse-phase HPLC to give cyclo-(ImImPyPy-(R)α-H2Nγ-ImImPyPy-(R)α-H2Nγ-γH2Nγ-γ) as a fluffy white solid in 0.1% overall yield (0.480 μmol). Cyclo-(ImImPyPy-(R)α-BocHα-ImImPyPy-(R)α-H2Nγ-γH2Nγ-γ) (6) MALDI-TOF MS (m/z): calc’d for C_{57}H_{68}N_{24}O_{12} [M+Na]^+ 1303.53, found 1303.36.

5.4.4 Cyclo-(ImImPyPy-(R)α-H2Nγ-ImImPyPy-(R)α-H2Nγ-γ) (1)
A solution of cyclo-(ImImPyPy-(R)α-BocHα-ImImPyPy-(R)α-H2Nγ-γH2Nγ-γ) (6) (0.400 μmol) in anhydrous TFA/DCM (1:1, 500 μL) was stirred at 23 °C for 5 min prior to being taken up in 9.5 mL of H_2O
(0.1% TFA), and purified by preparative reverse-phase HPLC to give cyclo-(-ImImPyPy-(R)-α-H2Nγ-ImImPyPy-(R)-α-H2Nγ-ImImPyPy-(R)-α-H2Nγ-ImImPyPy-(R)-α-H2Nγ-) 1 as a fluffy white solid in 90% yield (0.360 μmol). Cyclo-(-ImImPyPy-(R)-α-H2Nγ-ImImPyPy-(R)-α-H2Nγ-ImImPyPy-(R)-α-H2Nγ-ImImPyPy-(R)-α-H2Nγ-) 1 MALDI-TOF MS (m/z): calc’d for C52H61N24O10 [M+H]+ 1181.50, found 1181.55; calc’d for C52H60N24O10Na [M+Na]+ 1203.48, found 1203.37; calc’d for C52H60N24O10K [M+K]+ 1219.46, found 1219.33.

5.4.5 Oligonucleotide Purification and Crystallization

Oligonucleotides were purchased HPLC purified from Trilink Biotechnologies (San Diego, CA). Prior to use, oligonucleotides were de-salted using a Waters Sep-Pak cartridge (5g, C-18 sorbent). The Sep-Pak was pre-washed with acetonitrile (25 mL, 3x) followed by MilliQ water (25 mL, 3x). The oligonucleotide was dissolved in 5 mL of 2.0 M NaCl and loaded directly onto the sorbent followed by a wash with 5 mL of 2.0 M NaCl and 250 mL of MilliQ water. The oligonucleotide was eluted with acetonitrile:water (1:1) and lyophilized to dryness. Single strand DNA was quantitated by UV-Vis spectroscopy and incubated with a 2:1 ratio of DNA to polyamide prior to crystallization. Crystals were obtained after 2-8 weeks from a solution of 0.5 mM duplex DNA, 0.65 mM polyamide, 21% 2-methyl-2,4-pentanediol (MPD), 35 mM calcium acetate, 10 mM Tris pH 7.5 equilibrated in sitting drops against a reservoir of 35% MPD at 4 °C. Crystals were collected in Hampton nylon CryoLoops (10 micron, 0.1 mm) and flash cooled to 100 K prior to data collection (Figure 5.8).

5.4.6 UV-visible analysis

DNA/polyamide complex formation was verified prior to structure solution by UV-Visible spectroscopy. Crystals were collected in Hampton nylon CryoLoops (10 micron, 0.1 mm) and washed with crystallization buffer 3 times prior to dissolution in 50 μL of MilliQ water. UV-Visible spectroscopy of the dissolved crystals confirmed the presence of polyamide and DNA duplex in a 1:1 stoichiometry (Figure 5.8).

5.4.7 Data collection, Structure Determination, and Refinement

Polyamide-DNA crystals grew in space group P1 with unit cell dimensions a = 22.50, b = 25.14, c = 29.09, α = 66.53, β = 79.28, γ = 79.57, and one polyamide-duplex DNA complex in the asymmetric unit. This data set was collected at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 12-2 with a MAR Research imaging plate detector at wavelength 0.97 Å. DNA only crystals grew in space group C2 (C 1 2 1) with unit cell dimensions a = 31.827, b = 25.636, c = 34.173, α = 90.00,
\[ \beta = 116.72, \gamma = 90.00 \] and one DNA strand in the asymmetric unit. This data set was collected at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 11-1 with a MAR Research imaging plate detector at wavelength 0.999 Å (Table 5.1).

Data was processed with MOSFLM\textsuperscript{30} and SCALA\textsuperscript{31} from the CCP4 suite of programs. Both crystals were solved by direct methods using the SHELX suite of programs (SHELXD).\textsuperscript{32,33} Model building and structure refinement was done with Coot\textsuperscript{34} and REFMAC5.\textsuperscript{35} The final polyamide-DNA complex was refined to an R factor of 9.8 % and an R\textsubscript{free} of 13.6 %. The final DNA structure was refined to an R factor of 10.9 % and an R\textsubscript{free} of 14.3 %. Anisotropic B factors were refined in the final stages and riding hydrogens included (Table 5.1).

5.4.8 Structure Analysis and Figure Preparation

DNA helical parameters were calculated using the program Curves and 3DNA.\textsuperscript{36,37} Molecular electrostatic potential maps were calculated at the HF/3-21g* level using the Gamess program (Figure 5.13).\textsuperscript{38-40} Distance measurements and least squares fitting procedures for ring-centroid measurements were performed using UCSF Chimera\textsuperscript{41} and Mercury.\textsuperscript{42} Structural figures were prepared using UCSF Chimera.\textsuperscript{41}

5.5 Notes and References


5.6 Spectra, Data Statistics, and Supplemental Information

Figure 5.6 Solid-phase synthesis of cyclic polyamide 1 using Kaiser oxime resin and the commercially available building blocks presented above. Reagents and conditions: (i) Boc-Py-OBt 2, DIEA, DMF; (ii) 20% TFA/DCM; (iii) Boc-Py-OBt 2, DIEA, DMF; (iv) 20% TFA/DCM; (v) Boc-Im-OH 3, PyBOP, DIEA, DMF; (vi) 50% TFA/DCM; (vii) Boc-Im-OH 3, PyBOP, DIEA, DMF; (viii) 50% TFA/DCM; (ix) Fmoc-D-Dab(Boc)-OH 4, PyBOP, DIEA, DMF; (x) 20% TFA/DCM; (xi) Boc-Py-OBt 2, DIEA, DMF; (xii) 20% TFA/DCM; (xiii) Boc-Py-OBt 2, DIEA, DMF; (xiv) 20% TFA/DCM; (xv) Boc-Im-OH 3, PyBOP, DIEA, DMF; (xvi) 50% TFA/DCM; (xvii) Boc-Im-OH 3, PyBOP, DIEA, DMF; (xviii) 50% TFA/DCM; (xix) Boc-D-Dab-(Fmoc)-OH 5, PyBOP, DIEA, DMF; (xx) piperidine, DMF; (xxi) DMF, 37°C, 24 h; (xxii) 50% TFA/DCM.
Figure 5.7 Polyamide analytical data. Analysis of polyamides 6 and 1 by analytical RP-HPLC and MALDI-TOF MS. Wavelength shown is at 310 nm.

MALDI-TOF MS (m/z): calc’d for
C_{52}H_{106}N_{24}O_{29}Na [M+Na]^+ 1303.53; found 1303.36.

MALDI-TOF MS (m/z): calc’d for
C_{32}H_{61}N_{16}O_{10} [M+H]^+ 1181.50; found 1181.55.
Table 5.1 Data collection and refinement statistics.

<table>
<thead>
<tr>
<th>Data collection</th>
<th>DNA-Polyamide</th>
<th>DNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P1</td>
<td>C2</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, and c, Å</td>
<td>22.50</td>
<td>31.83</td>
</tr>
<tr>
<td></td>
<td>25.14</td>
<td>25.64</td>
</tr>
<tr>
<td></td>
<td>29.09</td>
<td>34.17</td>
</tr>
<tr>
<td>α, β, and γ, °</td>
<td>66.53</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>79.28</td>
<td>116.72</td>
</tr>
<tr>
<td></td>
<td>79.57</td>
<td>90</td>
</tr>
<tr>
<td>Wavelength</td>
<td>0.970</td>
<td>0.999</td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>26.47 - 1.18</td>
<td>15.91 - 0.98</td>
</tr>
<tr>
<td>$R_{merge}$</td>
<td>12.4 (15.4)</td>
<td>5.0 (7.5)</td>
</tr>
<tr>
<td>$\omega$</td>
<td>12.5 (8.9)</td>
<td>18.7 (8.3)</td>
</tr>
<tr>
<td>Completeness, %$^*$</td>
<td>94.6 (90.7)</td>
<td>89.5</td>
</tr>
<tr>
<td>Redundancy</td>
<td>4.0</td>
<td>2.5</td>
</tr>
<tr>
<td>Refinement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution, Å</td>
<td>26.47 - 1.18</td>
<td>15.91 - 0.98</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>16,139</td>
<td>12,125</td>
</tr>
<tr>
<td>$R_{work}/R_{free}$</td>
<td>9.9 / 13.8</td>
<td>10.8 / 14.4</td>
</tr>
<tr>
<td>No. of atoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>441</td>
<td>297</td>
</tr>
<tr>
<td>Polyamide</td>
<td>86</td>
<td>–</td>
</tr>
<tr>
<td>Calcium</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>Water</td>
<td>262</td>
<td>137</td>
</tr>
<tr>
<td>B factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>5.6</td>
<td>6.5</td>
</tr>
<tr>
<td>Polyamide</td>
<td>5.2</td>
<td>–</td>
</tr>
<tr>
<td>Calcium</td>
<td>11.7</td>
<td>33.2</td>
</tr>
<tr>
<td>Water</td>
<td>19.2</td>
<td>20.1</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths, Å</td>
<td>0.026</td>
<td>0.033</td>
</tr>
<tr>
<td>Bond angles, °</td>
<td>2.57</td>
<td>3.44</td>
</tr>
</tbody>
</table>

$^*$Free $R$ calculated against 5% of the reflections randomly removed.

$^*$Highest-resolution shell is shown in parentheses.
Figure 5.8 Single crystal of the polyamide-DNA complex. a) Single crystal frozen in a cryoloop for data collection. b) UV-Vis of the single crystal polyamide-DNA complex dissolved in 50 μL of water.
Figure 5.9 Hydrogen bond map of polyamide-DNA complex.
Figure 5.10 Polyamide/DNA complex map. Schematic diagram of polyamide/DNA crystal structure showing the sugar conformation at each position and calcium ion coordination.
Figure 5.11 DNA structure map. Schematic diagram of native DNA crystal structure showing the sugar conformation at each position and the electron density map of the DNA asymmetric unit contoured at the 1.0 \( \sigma \) level. The structure on the right is oriented 5' to 3' from top to bottom and corresponds to the right hand strand in the structure map.
Figure 5.12 Hydrogen bond map of DNA crystal structure and electron density of the AT base pair contoured at the 2.0 σ level.
Figure 5.13 Molecular electrostatic potential maps of compounds 7-10 and complex 11. Complex 11 coordinates were taken directly from the polyamide/DNA crystal structure and used without geometry optimization. All ab initio calculations reported here were performed using HF/3-231G* as implemented in the Gamess program.37-39 Full geometry optimization was performed on all structures except for complex 11 whose coordinates correspond to those of the crystal structure. Electrostatic potential surfaces were generated by mapping the electrostatic potentials onto surfaces of molecular electron density (0.002 electron/Å³) and color-coding, using the Chimera program.40 The molecular electrostatic potential energy values range from -25 kcal/mol for values of negative potential (red) to +25 kcal/mol for values of positive potential (blue). This range was chosen to emphasize the variations in the aromatic region and some regions of the electrostatic potential associated with heteroatoms may lie beyond the ±25 kcal/mol range.
<table>
<thead>
<tr>
<th>Parameter*</th>
<th>CC/GG</th>
<th>CA/TG</th>
<th>AG/CT</th>
<th>GG/CC</th>
<th>GC/GC</th>
<th>CC/GG</th>
<th>CT/AG</th>
<th>TG/CA</th>
<th>GG/CC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shift, Å</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>-0.58</td>
<td>0.21</td>
<td>0.65</td>
<td>-0.34</td>
<td>0.00</td>
<td>0.34</td>
<td>-0.56</td>
<td>-0.21</td>
<td>0.56</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.00</td>
<td>0.63</td>
<td>-1.96</td>
<td>-1.35</td>
<td>0.13</td>
<td>1.43</td>
<td>1.81</td>
<td>-1.78</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Slide, Å</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>0.69</td>
<td>2.80</td>
<td>0.76</td>
<td>0.24</td>
<td>-0.37</td>
<td>0.24</td>
<td>0.78</td>
<td>2.60</td>
<td>0.68</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-0.47</td>
<td>0.59</td>
<td>0.31</td>
<td>0.94</td>
<td>-0.02</td>
<td>0.84</td>
<td>0.38</td>
<td>1.51</td>
<td>0.29</td>
</tr>
<tr>
<td><strong>Rise, Å</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>3.30</td>
<td>3.26</td>
<td>3.06</td>
<td>2.69</td>
<td>4.48</td>
<td>2.99</td>
<td>3.08</td>
<td>3.26</td>
<td>3.30</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>3.33</td>
<td>3.08</td>
<td>3.26</td>
<td>3.28</td>
<td>3.04</td>
<td>3.29</td>
<td>3.29</td>
<td>3.08</td>
<td>3.43</td>
</tr>
<tr>
<td><strong>Tilt, °</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>1.82</td>
<td>-0.81</td>
<td>3.57</td>
<td>-5.48</td>
<td>0.00</td>
<td>5.49</td>
<td>3.57</td>
<td>0.81</td>
<td>-1.82</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>4.14</td>
<td>1.42</td>
<td>8.98</td>
<td>-4.81</td>
<td>0.04</td>
<td>5.05</td>
<td>3.58</td>
<td>-2.03</td>
<td>-3.89</td>
</tr>
<tr>
<td><strong>Roll, °</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>0.80</td>
<td>-3.73</td>
<td>9.85</td>
<td>-2.96</td>
<td>-4.11</td>
<td>-2.96</td>
<td>9.85</td>
<td>-8.73</td>
<td>9.80</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>11.79</td>
<td>7.20</td>
<td>13.46</td>
<td>1.07</td>
<td>-0.20</td>
<td>2.57</td>
<td>16.09</td>
<td>6.54</td>
<td>11.05</td>
</tr>
<tr>
<td><strong>Twist, °</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>27.24</td>
<td>50.86</td>
<td>21.53</td>
<td>33.56</td>
<td>52.36</td>
<td>33.56</td>
<td>21.53</td>
<td>50.86</td>
<td>27.24</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>36.26</td>
<td>29.92</td>
<td>33.52</td>
<td>38.71</td>
<td>33.84</td>
<td>38.32</td>
<td>31.32</td>
<td>37.22</td>
<td>38.20</td>
</tr>
</tbody>
</table>

*Relationship between the bases composing the base pair.

**Figure 5.14** Comparison of Local base-pair step parameters for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.
**Local base-pair parameters and sugar pucker**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>C-G</th>
<th>G-C</th>
<th>A-T</th>
<th>G-C</th>
<th>G-C</th>
<th>C-G</th>
<th>C-G</th>
<th>T-A</th>
<th>G-C</th>
<th>G-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shear, Å</td>
<td>0.04</td>
<td>0.16</td>
<td>0.02</td>
<td>-0.33</td>
<td>-0.32</td>
<td>0.32</td>
<td>0.33</td>
<td>-0.02</td>
<td>-0.16</td>
<td>-0.04</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.15</td>
<td>0.10</td>
<td>0.16</td>
<td>-0.07</td>
<td>-0.17</td>
<td>0.11</td>
<td>0.10</td>
<td>-0.04</td>
<td>-0.37</td>
<td>0.05</td>
</tr>
<tr>
<td>Stretch, Å</td>
<td>-0.21</td>
<td>-0.20</td>
<td>-0.13</td>
<td>-0.19</td>
<td>0.02</td>
<td>0.02</td>
<td>-0.19</td>
<td>-0.13</td>
<td>-0.20</td>
<td>-0.21</td>
</tr>
<tr>
<td>DNA</td>
<td>-0.14</td>
<td>-0.12</td>
<td>-0.07</td>
<td>-0.10</td>
<td>-0.15</td>
<td>-0.13</td>
<td>-0.11</td>
<td>0.13</td>
<td>0.26</td>
<td>-0.09</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.07</td>
<td>0.08</td>
<td>0.13</td>
<td>0.31</td>
<td>0.86</td>
<td>0.86</td>
<td>0.31</td>
<td>0.13</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>Stagger, Å</td>
<td>0.16</td>
<td>0.09</td>
<td>-0.10</td>
<td>0.08</td>
<td>0.07</td>
<td>-0.02</td>
<td>-0.09</td>
<td>-0.21</td>
<td>0.03</td>
<td>0.17</td>
</tr>
<tr>
<td>Buckle, °</td>
<td>-5.52</td>
<td>-2.64</td>
<td>0.23</td>
<td>13.52</td>
<td>22.43</td>
<td>22.43</td>
<td>13.52</td>
<td>-0.23</td>
<td>2.64</td>
<td>5.52</td>
</tr>
<tr>
<td>DNA</td>
<td>-5.30</td>
<td>-9.09</td>
<td>5.62</td>
<td>-1.10</td>
<td>-2.90</td>
<td>1.00</td>
<td>1.61</td>
<td>-0.42</td>
<td>11.29</td>
<td>5.28</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-16.32</td>
<td>-9.30</td>
<td>-4.01</td>
<td>-9.51</td>
<td>-12.64</td>
<td>-12.64</td>
<td>-9.51</td>
<td>-4.01</td>
<td>-9.30</td>
<td>-16.32</td>
</tr>
<tr>
<td>DNA</td>
<td>2.09</td>
<td>-2.60</td>
<td>1.17</td>
<td>0.49</td>
<td>4.88</td>
<td>4.88</td>
<td>0.49</td>
<td>1.17</td>
<td>-2.60</td>
<td>2.09</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-3.69</td>
<td>-0.93</td>
<td>4.73</td>
<td>-1.94</td>
<td>-4.39</td>
<td>-4.48</td>
<td>-0.99</td>
<td>6.11</td>
<td>0.60</td>
<td>2.76</td>
</tr>
<tr>
<td>Sugar pucker</td>
<td>C1'-exo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C1'-exo</td>
<td>C2'-endo</td>
<td>C1'-exo</td>
<td>C1'-exo</td>
<td>C2'-endo</td>
<td>C3'-exo</td>
<td>C1'-exo</td>
</tr>
<tr>
<td>DNA</td>
<td>C1'-exo</td>
<td>C3'-exo</td>
<td>C2'-endo</td>
<td>C1'-exo</td>
<td>C1'-exo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C1'-exo</td>
<td></td>
</tr>
<tr>
<td>PA/DNA</td>
<td>C3'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td></td>
</tr>
</tbody>
</table>

*Relationship between the bases composing the base pair.*

**Figure 5.15** Comparison of Local base-pair parameters and sugar conformations for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.
Figure 5.16 Comparison of Local base-pair helical parameters for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.
Chapter 6: Structural Elucidation of a $\beta$-amino-$\gamma$-linked Cyclic Polyamide-DNA Complex and RNA Binding Studies
Abstract

Chemical control of dysregulated gene expression poses a significant challenge at the frontier of chemical biology. Py-Im polyamides are a class of small molecules that can be programmed to bind a broad repertoire of DNA sequences, disrupt transcription factor-DNA interfaces, and modulate gene expression pathways in cell culture experiments. Detailed structural information thus far has been limited to moderate resolution X-ray structures of unlinked 2:1 binding polyamides and NMR NOESY-restrained molecular dynamics models, with atomic resolution X-ray structures remaining elusive. Structural elucidation of polyamide-DNA complexes is fundamental to understanding the recognition properties at the molecular level and this chapter reports a high resolution (0.95 Å) X-ray crystal structure of a cyclic Py-Im polyamide bound to the central six base pairs of the sequence d(5’-CCAGTACTGG-3’)_2, revealing the hydrophobic, electrostatic, and shape selective recognition interactions responsible for DNA binding. Additionally, a structural basis for the allosteric modulation of transcription-factor DNA interfaces with β-amino turn linked cyclic polyamides is reported and a combination of biophysical, structural, and modeling studies are presented which explain the inability of polyamides to bind double helical A-form RNA.
6.1 Introduction

Pyrrole-imidazole polyamides are a class of small molecules that bind the minor groove of DNA sequence specifically. Encoded by side-by-side arrangements of N-methylpyrrole (Py) and N-methylimidazole (Im) carboxamide monomers, Im/Py pairs distinguish G•C from C•G base pairs, whereas Py/Py pairs are degenerate for T•A and A•T. Polyamides have been shown to bind a broad repertoire of DNA sequences, permeate cell membranes and traffic to the nucleus, access chromatin, and disrupt protein-DNA interactions. Polyamide inhibition of transcription factor-DNA binding of HIF-1α, androgen receptor (AR), and AP-1 has been exploited for controlling expression of medically relevant genes such as VEGF, PSA, TGF-β1 and LOX-1 in cell culture experiments. X-ray crystallography has provided structures of unlinked 2:1 binding polyamides to a resolution of 2.00 Å, providing valuable insight into the polyamide-DNA molecular recognition process. Structural studies of hairpin polyamides bound to the nucleosome core particle have also provided structural proof that polyamides can bind biologically relevant higher order structure, however a combination of resolution limits and high B-factors prevented a detailed picture of the polyamide-DNA interactions beyond confirmation of the binding location. Much insight has been gleaned from NMR studies where NOESY-restrained molecular dynamics models have provided structures of 1:1 and 6-ring cyclic polyamides. Despite these successes, atomic resolution X-ray structures of this important class of compounds and in general minor groove binders have remained elusive.

The ability of DNA to undergo bending, twisting, and stretching motions as well as the long-range propagation (allosteric effect) of these perturbations coupled with DNA recognition by proteins and small molecules can have profound influences over important processes such as gene transcription and modulation of eukaryotic gene networks. Allosteric communication along and through the DNA helix has been the subject of intense study and forms the basis for cooperative interactions among transcription factor regulatory networks such as the interferon-β enhanceosome, where transcription factor binding induced DNA conformational changes led to cooperative enhancer occupancy. This potential for short and long range allosteric control over the transcriptional machinery provides a powerful concept for the design of small molecules that can bind to topographically distinct locations on DNA with the possibility of modulating transcription factor-DNA binding by allosteric perturbation of the DNA structure. Recent studies (Chapter 3) of cyclic polyamide and analogous hairpin polyamides revealed they possessed high DNA binding affinities and could regulate endogenous androgen receptor-activated gene expression (prostate
specific antigen) in cell culture, from which we infer cell permeability. Additionally, In vitro ADMET studies of cycle 1 and hairpin polyamides revealed favorable drug-like properties for both classes of compounds and excellent metabolic stability.

Reported, is the atomic resolution structure (0.95 Å resolution) of an 8-ring cyclic polyamide in complex with double helical DNA. The cyclic polyamide 1 is comprised of two antiparallel ImPyPyPy strands capped by (R)-β-amino-γ turn units. Polyamide 1, which codes for the sequence 5’-WGWWCW-3’ was co-crystallized with the palindromic DNA oligonucleotide sequence 5’-CCIAGTACTGG-3’ 10 base pairs in length (Fig. 6.1). We observe significant allosteric structural perturbations of the DNA helix induced upon binding of substituted GABA (γ-aminobutyric acid) turn-linked polyamides in the DNA minor groove. In addition to amide and imidazole recognition with the DNA minor groove floor, a detailed view of the β-amino-γ-turn conformation and hydration reveals a network of well-ordered water-mediated interactions between the polyamide and DNA. Significantly, we find that a conformational inversion occurs at the turn position upon moving the amino substituent from the α to β positions of the GABA turn. The allosteric modulation of the DNA structure induced by polyamide binding is also shown and a structural basis for the inability of polyamides to bind A-form

![Figure 6.1](image_url)
RNA is presented with UV-melting temperature data.

### 6.2 Overall Structure

The structure of cyclic polyamide 1 in complex with d(5'-CCiAGTACTGG)2 was solved by direct methods to 0.95 Å resolution with synchrotron radiation (Fig. 6.1). One cyclic polyamide bound to a single DNA duplex is present in the asymmetric unit of the crystal in the P41212 space group. The final structure was refined anisotropic and unrestrained to an \( R \)-factor of 11.2% and an \( R_{\text{free}} \) of 12.4% (Table 6.1). The average B-factor for the polyamide was 6.7 Å² and 7.2 Å² for DNA. The asymmetric unit contains one full polyamide-complexed DNA double helix. In the DNA complex, the aromatic amino acids are bound with an N- to C-orientation of each ImPyPyPy strand of the cycle adjacent to the 5' to 3' direction of the DNA. The conformational constraints imposed by the turn unit result in ring placement that is ring-over-ring as opposed to ring-over-amide. Greater than 40% of the polyamide surface area is buried leaving only the top of the methyl groups on the heterocycles, the amide carbonyl oxygens, and the chiral \( \beta \)-ammonium turn solvent exposed. Alternate conformations are observed for 7 of the 18 nucleotides of the DNA duplex. The cyclic pyrrole-imidazole polyamide 1 was co-crystallized with the palindromic DNA oligonucleotide sequence shown in Figure 6.1. The polyamide selectively binds to the sequence, 5'-WGTACW-3', and previous studies have demonstrated that the equilibrium binding constants are sub-nanomolar and outside the measurement range of DNase I footprinting methods (see Chapter 3).
6.3 Overall structure of DNA-polyamide complex

The cyclic pyrrole-imidazole polyamide is bound in an antiparallel head-to-tail turn-linked fashion with the N- to C-terminal orientation of each PyPyPyIm strand of the polyamide directly adjacent to a DNA strand oriented in a 5’ to 3’ direction. Figure 6.1 shows the overall structure of the complex with the electron density map contoured to the 1.0 σ level. Figure 6.2b shows a view of the complex looking directly down the minor groove at the polyamide turn linkage. From this view it can be seen that significant stabilization of the complex is derived from van der Waals
interactions between the outside face of the pyrrole-imidazole heterocyclic strands and the walls of the minor groove, which form a deep binding pocket for the polyamide. Greater than 40% of the polyamide surface area is buried and not solvent exposed, leaving only the top of the methyl groups of the heterocycles, the amide carbonyl oxygens, and the chiral β-amino turn linkage solvated. The turn linkage adds a conformational constraint to the ends of the polyamides preventing the heterocycle strands from slipping past each other as observed in the slipped orientations found in some 2:1 binders such as distamycin.

6.4 Turn conformation

The β-methine conformational preference is puckered up and away from the DNA minor-groove floor, aligning the β-ammonium along the groove. Conformation A (left) is the conformation observed in the previously determined α-amino turn X-ray crystal structure. Conformation B (right) shows the preferred conformation for the β-amino turn determined by
X-ray crystallography in this report. The hydration pattern around the turn is highly conserved at both ends of the structure and there are two water-mediated hydrogen bonds within 2.79 - 2.87 Å from the ammonium to the DNA minor-groove floor (Fig. 6.2c). The amide NH’s and imidazole lone-pairs form a continuous series of direct hydrogen bonds to the floor of the DNA minor-groove, while the imidazoles impart specificity for the exocyclic amine of guanine through relief of steric interaction and a G(N2-hydrogen)-Im (lone pair) hydrogen bond (Fig. 6.2c-d). The amides linking the aromatic rings and the turns contribute hydrogen bonds to the purine N3 and pyrimidine O2 lone pairs where they all are within hydrogen bonding distance of a single DNA base. In total there are 10 direct amide hydrogen bonds (2.86–3.08 Å), 2 direct imidazole hydrogen bonds (3.15 and 3.16 Å), and 4 (R)-β-ammonium turn water-mediated hydrogen bonds (two per turn, average distance from amine to water = 2.79 Å–2.87 Å) to the floor of the DNA minor groove with at least one interaction for all 12 DNA base-pairs in the 6 bp binding site. There are a total of 16 hydrogen bond interactions between the cyclic polyamide and the floor of the DNA minor-groove, utilizing every hydrogen-bond donor and acceptor of the ligand (Fig. 6.2).

6.5 Allosteric Perturbations

Polyamide binding induces large structural changes in DNA and Figure 6.3 shows a

<table>
<thead>
<tr>
<th>bp step</th>
<th>Roll, °</th>
<th>Twist, °</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DNA</td>
<td>PA/DNA</td>
</tr>
<tr>
<td>CC/GG</td>
<td>9.03</td>
<td>7.70</td>
</tr>
<tr>
<td>CA/TG</td>
<td>-7.95</td>
<td>8.79</td>
</tr>
<tr>
<td>AG/CT</td>
<td>9.87</td>
<td>11.19</td>
</tr>
<tr>
<td>GT/AC</td>
<td>-0.74</td>
<td>2.85</td>
</tr>
<tr>
<td>TA/TA</td>
<td>3.26</td>
<td>7.23</td>
</tr>
<tr>
<td>AC/GT</td>
<td>-0.74</td>
<td>2.47</td>
</tr>
<tr>
<td>CT/AG</td>
<td>9.89</td>
<td>10.59</td>
</tr>
<tr>
<td>TG/CA</td>
<td>-7.96</td>
<td>9.57</td>
</tr>
<tr>
<td>GG/CC</td>
<td>9.03</td>
<td>3.58</td>
</tr>
</tbody>
</table>

**Figure 6.4** Allosteric distortion upon polyamide binding. (left) DNA bending is observed for polyamide-bound DNA (blue structure) versus unbound-DNA (yellow structure). (right) Helical parameters for DNA in the absence and presence of polyamide showing an increase in positive roll and significant changes in twist angles upon polyamide binding. (Polyamide has been removed from the blue complex for clarity)
slice through the short axis of the DNA helix showing the minor and major groove geometry at the center of the polyamide binding site for uncomplexed and complexed DNA. Figure 6.3 reveals a >4 Å widening of the DNA minor groove upon polyamide binding and a compression of the major groove by more than 4 Å. Additionally, Figure 6.3 shows a major perturbation in the major groove depth upon polyamide binding converting the wide shallow surface of the major groove from a functionally exposed protein recognition domain to a narrow deep cleft less likely to accommodate the width of a standard protein alpha-helical domain or beta-sheet from a transcription factor. Figure 6.4 shows the polyamide induced bending of the DNA helix. The helix is bent toward the major groove by >15° resulting in major groove compression. The base-pair step parameters in Figure 6.4 show a large positive roll throughout the polyamide binding site which contributes to the significant bend in the DNA helix. Additionally, polyamide binding induces a more uniform helical twist resulting in less variability as the base-pair step changes. The helical twist values for polyamide bound DNA range from 29.68-35.93°. Values for the helical twist are highly sequence dependent in native DNA and range from 21.04 to 50.50° depending on step sequence. Major perturbations in the DNA base pair buckle and opening are also observed upon polyamide binding. At the central 4 base pairs of the binding site the buckle is significantly reduced upon binding and the base pairs are opened toward the DNA major groove with the largest variations occurring at the central AT base pairs. [Note: For a full set of helical parameters and definitions see Figures 6.6, 6.7, and 6.8.]

### Table 6.2 Buckle and opening values.

<table>
<thead>
<tr>
<th>bp step</th>
<th>Buckle, °</th>
<th>Opening, °</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DNA</td>
<td>PA/DNA</td>
</tr>
<tr>
<td>C•G</td>
<td>-6.41</td>
<td>-10.59</td>
</tr>
<tr>
<td>C•G</td>
<td>-0.43</td>
<td>-7.08</td>
</tr>
<tr>
<td>A•T</td>
<td>-1.63</td>
<td>6.15</td>
</tr>
<tr>
<td>G•C</td>
<td>10.75</td>
<td>-2.23</td>
</tr>
<tr>
<td>T•A</td>
<td>13.15</td>
<td>0.37</td>
</tr>
<tr>
<td>A•T</td>
<td>-13.15</td>
<td>1.17</td>
</tr>
<tr>
<td>C•G</td>
<td>-10.74</td>
<td>2.39</td>
</tr>
<tr>
<td>T•A</td>
<td>1.63</td>
<td>-2.93</td>
</tr>
<tr>
<td>G•C</td>
<td>0.44</td>
<td>16.58</td>
</tr>
<tr>
<td>G•C</td>
<td>6.42</td>
<td>23.00</td>
</tr>
</tbody>
</table>

6.6 Solvation

The structure has a cell volume of 134,162 Å³ and a Matthews Coefficient of 2.24 with a solvent content of 51%. There are 130 out of 239 water molecules within 3.0 Å of the polyamide-DNA complex and 76 of the 130 water molecules localized around the DNA phosphates. The solvent exposed surface of the polyamide is hydrated by 22 of the 130 waters found within 3.0 Å of the complex.
These waters cluster and form bridges across the carbonyl oxygens of adjacent amides linking ring pairs, resulting in stripes of well ordered water across the polyamide surface. Additionally, 6 waters hydrate the polyamide ammonium turns (3 at each turn) with 4 of the 6 anchoring the polyamide to the floor of the DNA minor groove through bridging hydrogen bonds.

### 6.7 RNA Binding Studies

The minor groove binding of polyamides to B-form DNA has been extensively studied, however the ability of polyamides to bind double helical RNA has received little attention. There are two major differences in helical RNA versus DNA. First, thyamine (T) is replaced by uracil (U) presenting the addition of a 5’ methyl group to the major groove of the helix. However, the hydrogen

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>dsDNA sequence</th>
<th>dsRNA sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5′-CCAGTACTGG-3′</td>
<td>5′-CCAGUACUGG-3′</td>
</tr>
<tr>
<td></td>
<td>3′-GTCATGACC-5′</td>
<td>3′-GGCUAGACCC-5′</td>
</tr>
</tbody>
</table>

Table 6.3 Polyamide-DNA and Polyamide-RNA melting temperatures.

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>Tm / °C</th>
<th>ΔTm / °C</th>
<th>Tm / °C</th>
<th>ΔTm / °C</th>
</tr>
</thead>
<tbody>
<tr>
<td>—</td>
<td>46.1 (±0.8)</td>
<td>—</td>
<td>60.4 (±0.6)</td>
<td>—</td>
</tr>
<tr>
<td>1</td>
<td>82.8 (±0.6)</td>
<td>36.6 (±1.0)</td>
<td>59.9 (±0.8)</td>
<td>-0.5 (±1.0)</td>
</tr>
<tr>
<td>2</td>
<td>72.3 (±0.5)</td>
<td>26.2 (±1.0)</td>
<td>60.7 (±0.5)</td>
<td>0.3 (±0.8)</td>
</tr>
<tr>
<td>3</td>
<td>52.8 (±0.4)</td>
<td>6.6 (±0.9)</td>
<td>59.7 (±0.3)</td>
<td>-0.6 (±0.7)</td>
</tr>
</tbody>
</table>

Figure 6.5 a. Crystal structure of DNA-polyamide complex showing shape complementary and favorable hydrophobic interactions with the sugar-phosphate backbone. b. Coordinates of the cyclic polyamide docked within van der Waals radius of the putative binding site on a model of ideal A-form double helical RNA.
bonding functionality of the minor groove remains identical to that of B-form DNA. The second and most important difference is the addition of a 2′-OH on the sugar resulting in ribose as opposed to deoxyribose for DNA. This extra hydroxyl has a profound effect on the overall helical RNA structure and rigidity primarily due to the enforcement of a C3′-endo ribose sugar pucker. This pucker forces the RNA helix into an A or A′-form conformation due to the steric incompatibility of the 2′-OH with a DNA B-form conformation, which prefers a C2′-endo sugar conformation. The conformational rigidity leads to less sequence dependent microstructure than DNA and a dramatically different minor and major groove geometry. Additionally, the DNA helix has been shown to be highly conformationally mobile in contrast to the RNA helix. The structure of A-form RNA has an 11-fold helix with a narrow, deep major groove and shallow, wide minor groove in stark contrast to B-form DNA. The base pairs of A-form RNA are inclined and drastically displaced from the helix axis causing an overall expansion of the helix width, which in turn leads to a dramatically shallow curvature of the minor groove. The criteria required for polyamide binding and, in general, small molecule binding relies on the minimization of water exposed hydrophobic surfaces, the complementary pairing of buried hydrogen bond donors and acceptors, the maximization of van der Waals interactions, the solvation or neutralization of all charges, and the maximization of attractive and minimization of repulsive interactions. Our results show that the polyamides in this study provide a large thermal stabilization to DNA as opposed to RNA, which does not have an increased melting temperature for any of the compounds studied. The thermal stabilization for the DNA duplex ranges from 4 ºC for the unlinked polyamide (2:1 complex at saturating concentrations) to 23 ºC for the cyclic polyamide.

6.8 Conclusion

The crystal structure presented highlights the molecular recognition of β-amino turn-linked polyamides in the minor-groove of DNA and provides insight into the allosteric modulation of B-form DNA by Py-Im polyamides. The DNA structural distortion induced upon polyamide minor-groove binding provides an allosteric model for disrupting DNA:transcription factor interfaces in the promoters of selected genes. The ability of DNA to undergo short and long-range allosteric effects coupled with DNA binding by proteins can have influence over important processes such as modulation of eukaryotic gene networks. The potential for allosteric control over the transcriptional machinery provides a powerful concept for the design of small molecules that can bind to distinct locations on DNA with the possibility of modulating transcription factor activity. The RNA binding
studies demonstrate that this class of Py-Im polyamides are completely selective for binding dsDNA over dsRNA.

6.9 Experimental

6.9.1 Synthesis

Polyamides 1–3 were synthesized by standard solution-phase synthesis methods presented in Chapter 2 and 3 of this thesis. Synthetic deoxyoligonucleotides were purchased HPLC purified from Trilink Biotechnologies and desalted using a 5 gram sep-pak C18 cartridge (Waters) followed by lyophilization to dryness.

6.9.2 Oligonucleotide purification and Crystallization

Oligonucleotides were purchased HPLC purified from Trilink Biotechnologies (San Diego, CA). Prior to use the oligonucleotides were de-salted using a 5 gram sep-pak C18 cartridge (Waters) and lyophilized to dryness.

Single stranded DNA was quantitated by UV-Vis spectroscopy and incubated with a 2:1 ratio of DNA to polyamide prior to crystallization. Crystals were obtained from a solution of 0.6 mM duplex DNA, 0.75 mM polyamide, 24% 2-methyl-2,4-pentanediol (MPD), 35 mM calcium acetate, 10 mM Tris pH 7.5 equilibrated in sitting drops against a reservoir of 35% MPD at 4°C and crystals were flash cooled at 100 K prior to data collection. DNA crystals grew in space group P4₁2₁2 with unit cell dimensions \( a = 39.8270 \text{ Å}, \ b = 39.8270 \text{ Å}, \ c = 84.5718 \text{ Å}, \ \alpha = 90.00^\circ, \ \beta = 90.00^\circ, \ \gamma = 90.00^\circ \).

6.9.3 Data collection, Structure determination, and refinement

Polyamide-DNA crystals grew in space group P4₁2₁2 with unit cell dimensions \( a = 39.8270 \text{ Å}, \ b = 39.8270 \text{ Å}, \ c = 84.5718 \text{ Å}, \ \alpha = 90.00^\circ, \ \beta = 90.00^\circ, \ \gamma = 90.00^\circ \), and one polyamide-duplex DNA complex in the asymmetric unit. This data set was collected at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 12-2 with a MAR Research imaging plate plate detector at wavelength 0.97 Å.

Data was processed with Mosflm and scaled with the CCP4 program suite. Solution of both structures was obtained using SHELXD direct methods. All atoms were visible in initial maps from direct methods solution. Refinement was performed using Refmac37 and model building using Coot.38
6.9.4 Structure Analysis

DNA helical parameters were calculated using the program Curves\textsuperscript{39} and 3DNA.\textsuperscript{40} Figures were prepared and measurements made using UCSF Chimera.\textsuperscript{41}

6.10 Notes and References


6.11 Supplemental Information

### Local base-pair step parameters

<table>
<thead>
<tr>
<th>Parameter*</th>
<th>CC/GG</th>
<th>CA/TG</th>
<th>AG/CT</th>
<th>GT/AC</th>
<th>TA/TA</th>
<th>AC/GT</th>
<th>CT/AG</th>
<th>TG/CA</th>
<th>GG/CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shift, Å</td>
<td>-0.58</td>
<td>0.23</td>
<td>0.58</td>
<td>-0.73</td>
<td>0.00</td>
<td>0.73</td>
<td>-0.58</td>
<td>-0.23</td>
<td>0.58</td>
</tr>
<tr>
<td>DNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.17</td>
<td>0.45</td>
<td>1.72</td>
<td>0.74</td>
<td>-0.04</td>
<td>0.78</td>
<td>1.75</td>
<td>-0.00</td>
<td>0.34</td>
</tr>
<tr>
<td>Slide, Å</td>
<td>0.78</td>
<td>2.87</td>
<td>0.90</td>
<td>0.29</td>
<td>-0.10</td>
<td>0.29</td>
<td>0.80</td>
<td>2.87</td>
<td>0.78</td>
</tr>
<tr>
<td>DNA</td>
<td>-0.12</td>
<td>0.86</td>
<td>0.26</td>
<td>-0.15</td>
<td>0.90</td>
<td>-0.21</td>
<td>0.28</td>
<td>1.29</td>
<td>0.13</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>3.28</td>
<td>3.29</td>
<td>3.10</td>
<td>3.29</td>
<td>3.66</td>
<td>3.29</td>
<td>3.10</td>
<td>3.29</td>
<td>3.28</td>
</tr>
<tr>
<td>Rise, Å</td>
<td>3.28</td>
<td>3.18</td>
<td>3.59</td>
<td>3.20</td>
<td>3.26</td>
<td>3.24</td>
<td>3.33</td>
<td>3.06</td>
<td>3.19</td>
</tr>
<tr>
<td>DNA</td>
<td>3.32</td>
<td>-1.37</td>
<td>-3.91</td>
<td>2.23</td>
<td>0.00</td>
<td>-2.22</td>
<td>3.90</td>
<td>1.37</td>
<td>-3.33</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>5.81</td>
<td>0.74</td>
<td>-6.44</td>
<td>0.18</td>
<td>-0.16</td>
<td>0.40</td>
<td>7.94</td>
<td>0.64</td>
<td>-5.68</td>
</tr>
<tr>
<td>Tilt, °</td>
<td>9.03</td>
<td>-7.55</td>
<td>9.87</td>
<td>-0.74</td>
<td>3.66</td>
<td>-0.74</td>
<td>9.88</td>
<td>-7.96</td>
<td>9.03</td>
</tr>
<tr>
<td>DNA</td>
<td>7.70</td>
<td>8.79</td>
<td>11.19</td>
<td>2.85</td>
<td>7.23</td>
<td>2.47</td>
<td>10.59</td>
<td>9.57</td>
<td>3.58</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>33.68</td>
<td>30.23</td>
<td>35.16</td>
<td>35.80</td>
<td>48.48</td>
<td>35.80</td>
<td>21.04</td>
<td>50.49</td>
<td>28.26</td>
</tr>
<tr>
<td>Roll, °</td>
<td>28.26</td>
<td>50.50</td>
<td>21.04</td>
<td>35.80</td>
<td>48.48</td>
<td>35.80</td>
<td>21.04</td>
<td>50.49</td>
<td>28.26</td>
</tr>
<tr>
<td>DNA</td>
<td>33.68</td>
<td>30.23</td>
<td>35.16</td>
<td>35.80</td>
<td>48.48</td>
<td>35.80</td>
<td>21.04</td>
<td>50.49</td>
<td>28.26</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>28.26</td>
<td>50.50</td>
<td>21.04</td>
<td>35.80</td>
<td>48.48</td>
<td>35.80</td>
<td>21.04</td>
<td>50.49</td>
<td>28.26</td>
</tr>
</tbody>
</table>

* Relationship between the bases composing the base pair.

*aDNA corresponds to 0.73 Å structure of duplex DNA solved by Rees and coworkers (PDB 1D8G, 5'-CCAGTACTGG-3').

Figure 6.6 Comparison of Local base-pair step parameters for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.
### Figure 6.7: Comparison of Local base-pair parameters and sugar conformations for DNA with and without polyamide complexed.

Parameters were calculated using 3DNA.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>C-G</th>
<th>C-G</th>
<th>A-T</th>
<th>G-C</th>
<th>T-A</th>
<th>A-T</th>
<th>C-G</th>
<th>T-A</th>
<th>G-C</th>
<th>G-C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shear, Å</td>
<td>0.15</td>
<td>0.32</td>
<td>0.02</td>
<td>-0.19</td>
<td>-0.09</td>
<td>0.09</td>
<td>0.19</td>
<td>-0.02</td>
<td>-0.32</td>
<td>-0.15</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.15</td>
<td>0.19</td>
<td>0.13</td>
<td>-0.12</td>
<td>-0.13</td>
<td>0.13</td>
<td>0.12</td>
<td>-0.10</td>
<td>-0.16</td>
<td>-0.16</td>
</tr>
<tr>
<td>Stretch, Å</td>
<td>-0.15</td>
<td>-0.16</td>
<td>-0.11</td>
<td>-0.07</td>
<td>-0.03</td>
<td>-0.03</td>
<td>-0.07</td>
<td>-0.11</td>
<td>-0.16</td>
<td>-0.15</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-0.14</td>
<td>-0.10</td>
<td>-0.11</td>
<td>-0.07</td>
<td>-0.10</td>
<td>-0.08</td>
<td>-0.10</td>
<td>-0.13</td>
<td>-0.04</td>
<td>-0.10</td>
</tr>
<tr>
<td>Stagger, Å</td>
<td>0.09</td>
<td>0.12</td>
<td>0.08</td>
<td>0.25</td>
<td>0.00</td>
<td>0.00</td>
<td>0.25</td>
<td>0.08</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.27</td>
<td>-0.02</td>
<td>-0.10</td>
<td>0.09</td>
<td>-0.02</td>
<td>0.04</td>
<td>0.06</td>
<td>-0.14</td>
<td>0.07</td>
<td>0.61</td>
</tr>
<tr>
<td>Buckle, ⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>-6.41</td>
<td>-0.43</td>
<td>-1.63</td>
<td>10.75</td>
<td>13.15</td>
<td>-13.15</td>
<td>-10.74</td>
<td>1.63</td>
<td>0.44</td>
<td>6.42</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-10.59</td>
<td>-7.08</td>
<td>0.15</td>
<td>-2.23</td>
<td>0.37</td>
<td>1.17</td>
<td>2.39</td>
<td>-2.95</td>
<td>16.58</td>
<td>23.00</td>
</tr>
<tr>
<td>Propeller, ⁶</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-16.87</td>
<td>-5.31</td>
<td>-10.92</td>
<td>-4.06</td>
<td>-14.69</td>
<td>-16.26</td>
<td>-4.66</td>
<td>-7.96</td>
<td>-5.00</td>
<td>-17.05</td>
</tr>
<tr>
<td>Opening, ⁷</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>0.99</td>
<td>-1.15</td>
<td>1.35</td>
<td>1.26</td>
<td>-0.27</td>
<td>-0.27</td>
<td>1.27</td>
<td>1.34</td>
<td>-1.13</td>
<td>1.00</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.68</td>
<td>1.71</td>
<td>3.84</td>
<td>0.78</td>
<td>-4.87</td>
<td>-4.88</td>
<td>-1.10</td>
<td>3.78</td>
<td>0.87</td>
<td>0.15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sugar pucker</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>C4'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C1'-exo</td>
<td>C2'-endo</td>
<td>O4'-endo</td>
<td>C2'-endo</td>
<td>C4'-endo</td>
<td>C2'-endo</td>
<td>C3'-exo</td>
</tr>
<tr>
<td>O4'-endo</td>
<td>C1'-exo</td>
<td>C3'-exo</td>
<td>C2'-endo</td>
<td>O4'-endo</td>
<td>C2'-endo</td>
<td>C1'-exo</td>
<td>C4'-endo</td>
<td>C1'-exo</td>
<td>C2'-endo</td>
<td>C1'-exo</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>C3'-endo</td>
<td>C1'-exo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>O4'-endo</td>
<td>C2'-endo</td>
<td>O4'-endo</td>
<td>C2'-endo</td>
<td>C2'-endo</td>
<td>C3'-endo</td>
</tr>
</tbody>
</table>

⁵ Relationship between the bases composing the base pair.

⁶ DNA corresponds to 0.73 Å structure of duplex DNA solved by Rees and coworkers (PDB 1D8G, 5'-CCAGTACTGG-3').

---

**Figure 6.7** Comparison of Local base-pair parameters and sugar conformations for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.
### Local base-pair helical parameters

<table>
<thead>
<tr>
<th>Parameter*</th>
<th>CC/GG</th>
<th>CA/TG</th>
<th>AG/CT</th>
<th>GT/AC</th>
<th>TA/TA</th>
<th>AG/GT</th>
<th>CT/AG</th>
<th>TG/CA</th>
<th>GG/CC</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-displacement, Å</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>-0.42</td>
<td>3.84</td>
<td>-1.31</td>
<td>0.57</td>
<td>-0.41</td>
<td>0.57</td>
<td>-1.32</td>
<td>3.84</td>
<td>-0.42</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-1.36</td>
<td>-0.07</td>
<td>-1.21</td>
<td>-0.76</td>
<td>0.42</td>
<td>-0.81</td>
<td>-1.07</td>
<td>0.59</td>
<td>-0.34</td>
</tr>
<tr>
<td>Y-displacement, Å</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>1.83</td>
<td>-0.36</td>
<td>-2.71</td>
<td>1.50</td>
<td>0.00</td>
<td>-1.50</td>
<td>2.71</td>
<td>0.36</td>
<td>-1.84</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>0.98</td>
<td>-0.68</td>
<td>1.45</td>
<td>1.36</td>
<td>0.05</td>
<td>-1.34</td>
<td>1.59</td>
<td>1.60</td>
<td>-1.52</td>
</tr>
<tr>
<td>Inclination, °</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>17.85</td>
<td>-9.25</td>
<td>25.08</td>
<td>-1.20</td>
<td>3.96</td>
<td>-1.20</td>
<td>25.12</td>
<td>-9.26</td>
<td>17.85</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>12.56</td>
<td>16.43</td>
<td>17.68</td>
<td>5.14</td>
<td>11.58</td>
<td>4.47</td>
<td>16.74</td>
<td>16.10</td>
<td>6.29</td>
</tr>
<tr>
<td>Tilt, °</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>-6.56</td>
<td>1.59</td>
<td>9.52</td>
<td>-3.62</td>
<td>0.00</td>
<td>3.60</td>
<td>-9.90</td>
<td>-1.59</td>
<td>6.59</td>
</tr>
<tr>
<td>PA/DNA</td>
<td>-9.78</td>
<td>-1.38</td>
<td>13.33</td>
<td>-0.33</td>
<td>0.26</td>
<td>-0.73</td>
<td>-12.56</td>
<td>-1.20</td>
<td>9.97</td>
</tr>
</tbody>
</table>

*Relationship between the bases composing the base pair.

DNA corresponds to 0.73 Å structure of duplex DNA solved by Rees and coworkers (PDB 1D8G, 5'-CCAGTACTGG-3').

---

**Figure 6.8** Comparison of Local base-pair helical parameters for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.
Chapter 7: Programmable Oligomers Targeting 5’-GGGG-3’ in the

Minor Groove of DNA and NF-κB Binding Inhibition

The text of this chapter was taken in part from a manuscript coauthored with Julie A. Poposki, Michael A. Marques and Peter B. Dervan* (Caltech)

Abstract

A series of hairpin oligomers containing benzimidazole (Bi) and imidazopyridine (Ip) rings were synthesized and screened to target 5’-WGGGGW-3’, a core sequence in the DNA-binding site of NF-κB, a prolific transcription factor important in biology and disease. Five Bi and Ip containing oligomers bound to the 5’-WGGGGW-3’ site with high affinity. One of the oligomers (Im-Im-Im-Im-γ-Py-Bi-Py-Bi-β-Dp) was able to inhibit DNA binding by the transcription factor NF-κB.
7.1 Introduction

DNA-binding polyamides based on the architecture of the natural products netropsin and distamycin A have emerged as a promising class of gene modulators. These molecules are capable of distinguishing the four Watson-Crick base pairs in the DNA minor groove and have been the subject of intense study along with other classes of minor groove binders for interfering with specific protein–DNA interfaces. Sequence-specific recognition of the minor groove of DNA by polyamides arises from the pairing of three different aromatic amino acids, pyrrole (Py), imidazole (Im), and hydroxypyrrole (Hp). The direct readout, or information face, on the inside of the crescent-shaped polyamide may be programmed by the incremental change of atoms on the corners of the ring pairs presented to the DNA minor groove floor. Stabilizing and importantly, destabilizing interactions with the different edges of the four Watson-Crick bases are modulated by shape complementarity and specific hydrogen bonds. For example, the Im/Py pair distinguishes G•C from C•G, T•A, and A•T. Im presents a lone pair of electrons to the DNA minor groove and can accept a hydrogen bond from the exocyclic amine of guanine. Additionally, the Hp/Py pair distinguishes T•A from A•T, G•C, and C•G. Hp projects an exocyclic OH group toward the minor groove floor that is sterically accommodated in the cleft of the T•A base pair, preferring to lie over T not A. Since the development of these pairing rules based on the five-membered heterocyclic carboxamides Py, Im, and Hp, hundreds of DNA sequences have been successfully coded for using polyamides. Yet, due to sequence dependent alterations in the shape of the DNA minor groove, there remain DNA sequences that prove difficult to target with high affinity and specificity.

There has been an ongoing effort to broaden the repertoire of heterocycles with improved DNA recognition, stability, and cellular trafficking profiles. We recently reported that the benzimidazole ring can be an effective platform for modular paired recognition elements in the minor groove of DNA. The benzimidazole 6–5 bicyclic ring structure, while having slightly different curvature from the classic five-membered N-methyl pyrrole–carboxamides, presents an ‘inside edge’ with a similar atomic readout to the DNA minor groove floor, effectively mimicking Py, Im, and Hp (Figure 7.1). The imidazopyridine/pyrrole pair Ip/Py distinguishes G•C from C•G and the hydroxybenzimidazole/pyrrole pair Hz/Py distinguishes T•A from A•T, providing a solution to the unanticipated hydroxypyrrole instability limitation. This second generation solution to DNA recognition can be elaborated further, deleting incrementally almost all carboxamide linkages in the backbone of the hairpin motif creating an oligomer comprised of four dimer units capable of binding the site 5’-GTAC-3’, a sequence formally containing all four Watson-Crick base pairs.
A key strategic issue for small-molecule gene regulation is interfering with protein–DNA interfaces in the promoter of important genes. The new oligomer architecture has been successful in targeting the hypoxia inducible factor (HIF-1) binding site on the promoter of the VEGF gene.\textsuperscript{11} This recent success spurred our interest in using these new oligomers to target another prolific transcription factor, NF-κB, important in biology and disease.\textsuperscript{12} Guanine rich sequences are highly prevalent and partially conserved in the promoter region of NF-κB driven genes (Figure 7.2).\textsuperscript{12a,13} Previous studies from our laboratory have established the ability of polyamides, targeting the sequence 5’-GGGACT-3’, to inhibit DNA binding by NF-κB, however targeting the

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Figure 7.1 Structures of the (a) pyrrole-benzimidazole internal dimer (-Py-Bi-), (b) imidazole-imidazopyridine internal dimer (-Im-Ip-), and (c) imidazole-imidazopyridine cap (Im-Ip-) in comparison with their respective five-membered ring systems. Hydrogen-bonding surfaces to the DNA minor-groove floor are bolded.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{Figure 7.2 Crystal structure of the p50/p65 NF-κB heterodimer bound to the DNA duplex 5’-TG-GGGACTTCC-3’\textsuperscript{13}. The p50 and p65 monomers are represented as gold and dark blue ribbons, respectively. a) Top view looking down the DNA double helix. b) Side view showing the 5’-GGGG-3’ oligomer binding site. GC rich regions are shown in red and AT rich regions are shown in light blue. c) Sequence of DNA bound to NF-κB.}
\end{figure}
four consecutive guanines was less successful due to the modest affinity of pyrrole/imidazole polyamides for this sequence.\textsuperscript{14,15} The question arises whether a second generation oligomer architecture can target 5’-GGGG-3’ recognition sequences with improved affinity.

We report the synthesis, DNA binding properties, and NF-κB:DNA binding inhibition properties of a new class of DNA binding oligomers targeted at the DNA sequence 5’-WGGGGW-3’ containing 4-consecutive guanines (Figure 7.3). Oligomers vary by incorporation of benzimidazole–pyrrole into various positions in the parent polyamide 1, resulting in oligomers 2–6. From previous studies, we expect the BiPy dimer to be a good mimic for the traditional Py-Py recognition elements.\textsuperscript{16} The Ip-Im moiety is introduced as a new mimic for the traditional Im-Im recognition motif (Figure 7.1). Quantitative DNase I footprinting titration experiments\textsuperscript{17} were used to determine the binding affinities and specificities against single base-pair mismatch sites of oligomers 2–6 (Figure 7.3) as compared to their parent polyamide 1. We found that in oligomers 2–6, the 6–5 fused rings are effective mimics of their respective five-membered ring systems and that these oligomers target the binding-site 5’-WGGGGW-3’ without loss of affinity as compared to parent hairpin polyamide 1. In addition, we report the inhibition of DNA binding by the transcription factor NF-κB using this new class of minor groove binding oligomers (Figures 7.3 and 7.4).

7.2 Results and Discussion

7.2.1 Heterocycle Synthesis

Dimeric units Im-Ip-OH (9) and Boc-Im-Ip-OH (12) were synthesized by oxidative condensation of aldehydes 7 and 10, respectively, with previously reported diaminopyridine 8. Mixing 7 and 8 in DMF at 80 °C for 1 h followed by 12 h of heating at 100 °C in the presence of FeCl\textsubscript{3} and air afforded 9 after purification by precipitation from water and saponification using a mixture of KOH (4 M) in MeOH at 40 °C (Scheme 7.1). Nitroimidazole (10) was synthesized from 7 using a mixture of H\textsubscript{2}SO\textsubscript{4} + SO\textsubscript{3} and neat red fuming nitric acid.\textsuperscript{18} 10 was then added to a mixture of 8 in nitrobenzene and refluxed at 140 °C open to the atmosphere overnight to provide NO\textsubscript{2}-Im-Ip-OMe (11) after precipitation from water. Compound 11 was reduced using Pd/C in the presence of hydrogen followed by Boc protection using a mixture of Boc\textsubscript{2}O and DMAP in DMF to provided the final dimer Boc-Im-Ip-OH (12) after saponification and precipitation (Scheme 7.1).

7.2.2 Oligomer Synthesis

Oligomers 1–6 were synthesized using manual solid phase synthesis methodology on
commercially available β-Ala-PAM resin as previously described (Scheme 7.2). Starting from base resin (BR1), monomeric and dimeric heterocyclic units were appended onto the resin in stepwise fashion using HBTU activation. Couplings were allowed to proceed for several h between

Figure 7.3 Postulated hydrogen-bonding models for the 1:1 polyamide–DNA complexes with their matched sequence and the ball-and-stick representation for compounds 1 and 6 over the 6-base-pair matched binding site (variable region W = A or T). (a) Im-Im-Im-γ-Py-Py-Py-β-Dp (1), (b) Im-Im-Im-Im-γ-Py-Bi-Py-Bi-β-Dp (2), Im-Im-Im-Ip-γ-Py-Bi-Py-Bi-β-Dp (3), Im-Im-Ip-Ip-γ-Py-Bi-Bi-β-Dp (4), Im-Ip-Ip-Im-γ-Py-Bi-Py-Bi-β-Dp (5), Im-Ip-Ip-Ip-γ-Py-Bi-Bi-β-Dp (6).
and 40 °C. Unreacted amines were acylated between coupling rounds using acetic anhydride. Deprotection of the Boc-protected amines was accomplished using 80% TFA in DCM. After completion of solid phase synthesis, the resin was treated with dimethylaminopropylamine (Dp) and the oligomers were purified by reverse-phase preparatory HPLC: Im-Im-Im-Im-γ-Py-Py-Py-Py-β-Dp (1), Im-Im-Im-Im-γ-Py-Py-Bi-Py-Bi-β-Dp (2), Im-Im-Im-Ip-γ-Py-Bi-Py-Bi-β-Dp (3), Im-Im-Ip-Im-γ-Py-Bi-Py-Bi-β-Dp (4), Im-Ip-Im-Im-γ-Py-Bi-Py-Bi-β-Dp (5), and Im-Ip-Im-Ip-γ-Py-Bi-Py-Bi-β-Dp (6). Oligomers were characterized using MALDI-TOF mass spectrometry.

7.2.3 DNA affinity and sequence specificity

Quantitative DNase I footprinting titrations were carried out for oligomers 1–6 on the PCR product of plasmid pEF16 (Figures 7.5, 7.6, and 7.7). Plasmid pEF16 was constructed containing two designed match sites (5’-XGGGGT-3’ X = A, T) and two mismatch sites (5’-AGGGAT-3’ and 5’-AGAGGT-3’) (Figures 7.5, 7.6, and 7.7). The first two match sites 5’-TGGGGT-3’ and 5’-AGGGGT-3’ were included to determine if there was an energetic penalty associated with a 5’-AG-3’ step. Previously, DNA sequences containing multiple 5’-AG-3’ transitions have proven more difficult to target at high affinity and it has been shown that changes in flanking sequence can affect binding at a proximal site. The second two binding sites, 5’-AGGGAT-3’ and 5’-AGAGGT-3’, were designed to elucidate the energetic penalty for the loss of a favorable hydrogen bond between the exocyclic amine of guanine and the lone pair nitrogen on the oligomer in question. Control compound 1 bound both 4-G match sequences (5’-XGGGGT-3’, X = A, T) with comparably low affinity (\( K_a \approx 10^8 \text{ M}^{-1} \)), showing no bias for either site (Table 7.1). Compound 1 distinguished against mismatch sequences (5’-AGGGAT-3’ and 5’-AGAGGT-3’) with roughly 10-fold specificity (Table 7.1). Oligomer 2 demonstrated a large increase in affinity (\( K_a \approx 10^{10} \text{ M}^{-1} \)) as compared to 1 for both match sites, but showed lower specificity (4-fold) over the mismatch sequence.
sites. Oligomer 5 showed a moderate increase in affinity \((K_a \approx 10^9 \text{ M}^{-1})\) and demonstrated minor selectivity (3-fold) over the mismatch \(5'\text{-atGGGAt-3'}\), however showed an 11-fold specificity over the mismatch \(5'\text{-atGAGGt-3'}\). Oligomer 4 demonstrated high affinity for the 4-G match sequences \((K_a \approx \text{mid } 10^9 \text{ M}^{-1})\) with reasonable 5-fold selectivity over the mismatch sequences \((5'\text{-AGGGAT-3'}\) and \(5'\text{-AGGAGT-3'}\)). Compounds 3 and 6 bound all designed sequences with similar affinity. Thermodynamic data for oligomers 1–6 are summarized in Table 7.1.

In general, the global increase in affinity for these novel oligomers is not altogether unexpected. In contrast to the 5-membered heterocyclic carboxamides, the 6–5 fused benzimidazole analogues have a larger hydrophobic surface, likely promoting their placement in the DNA minor groove.\(^{21}\) Aromatic stacking and van der Waals interactions are also contributing factors.\(^{22,23}\) Furthermore, the benzimidazole derivatives are a more rigid structure with a lower degree of rotational freedom. Such pre-organization may decrease the entropic cost of DNA complexation.

### 7.2.4 NF-κB Electrophoretic Mobility Gel Shift Assay

The guanine rich region of the NF-κB binding element has been shown to be important in protein–DNA recognition (Figure 7.2).\(^{13,15,24,25}\) Polyamides targeted to this region have previously been shown to inhibit NF-κB binding.\(^{15}\) The possibility of steric or allosteric inhibition exists due to numerous contacts between the p50 protein and the phosphate backbone or direct protein–DNA contacts in the major groove, opposite the minor groove polyamide binding site.\(^{13,15}\) To test whether the second generation oligomer architecture would be successful at targeting \(5'\text{-GGGG-3'}\) in a biologically relevant context we employed an NF-κB electrophoretic mobility gel shift assay. The
NF-κB binding inhibition properties of compounds 1–6 were screened against a 40 bp DNA probe containing the 5’-WGGGGW-3’ sequence, which is part of an element from the intronic enhancer of the immunoglobulin κ light-chain gene recognized and bound by NF-κB (Figure 7.8). In the initial screen each oligomer was tested at two concentrations, 10 and 100 nM, for the ability to interfere with protein binding (Figure 7.8). At the higher concentration, 2 and 4 demonstrate a clear decrease in band intensity, whereas compounds 1, 3, 5, and 6 have little effect. This result agrees with the footprinting data, as protein inhibition appears to scale with compound affinity for the 5’-WGGGGW-3’ site. At the 10 nM concentration compound 2 is the only compound to show significant protein inhibition, reducing band intensity by more than 80% (Figure 7.8). Full titration of compound 2 over a concentration range of 500 pM–500 nM established an EC<sub>50</sub> of 15.7 nM for NF-κB inhibition as shown in Figure 7.9. In addition, the identity of the NF-κB band was confirmed by antibody supershift and the data are shown in Figure 7.10.

Scheme 7.2. Representative solid-phase synthesis of polyamide 6 along with a table of the amino acid building blocks used for the synthesis. Reagents and conditions: (i) 80% TFA/DCM; (ii) Boc-Py-Bi-OH (13), HBTU, DIEA, DMF; (iii) Ac<sub>2</sub>O, DIEA, DMF; (iv) repeat (i–iii); (v) 80% TFA/DCM; (vi) Boc-c-OH (14), HBTU, DIEA, DMF; (vii) Ac<sub>2</sub>O, DIEA, DMF to provide BR1; (viii) 80% TFA/DCM; (ix) Boc-Im-Ip-OH (12), HBTU, DIEA, DMF; (x) 80% TFA/DCM; (xi) Im-Ip-OH, HBTU, DIEA, DMF; (xii) dimethylaminopropylamine (Dp), 80 °C 2 h; (xiii) preparative HPLC to give 6.
Figure 7.5 Quantitative DNaseI footprinting experiments in the hairpin motif for polyamides 1 and 2 respectively, on the 278 bp, 5'-end-labeled PCR product of plasmid pEF16: lane 1, intact DNA; lanes 2 and 3, G and A sequencing 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 concentration, respectively. Each footprinting gel is accompanied by the following: (top) ball-and-stick models of the compound bound to its target DNA sequence; and (bottom) Binding isotherms for the four designed sites. $\theta_{\text{app}}$ values were obtained according to published methods. Imidazoles and pyrroles are shown as filled and non-filled circles, respectively; dimer units are represented as rectangles containing either Bi enclosed in a circle representing benzimidazole or Ip in a black box representing imidazopyridine; beta alanine is shown as a diamond; the gamma-aminobutyric acid turn residue is shown as a semicircle connecting the two subunits.
Figure 7.6 Quantitative DNaseI footprinting experiments in the hairpin motif for polyamides 3 and 4 respectively, on the 278 bp, 5'-end-labeled PCR product of plasmid pEF16: lane 1, intact DNA; lanes 2 and 3, G and A sequencing 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 concentration, respectively. Each footprinting gel is accompanied by the following: (top) ball-and-stick models of the compound bound to its target DNA sequence; and (bottom) Binding isotherms for the four designed sites. \( \theta_{\text{app}} \) values were obtained according to published methods. Imidazoles and pyrroles are shown as filled and non-filled circles, respectively; dimer units are represented as rectangles containing either Bi enclosed in a circle representing benzimidazole or Ip in a black box representing imidazopyridine; beta alanine is shown as a diamond; the gamma-aminobutyric acid turn residue is shown as a semicircle connecting the two subunits.
Figure 7.7 Quantitative DNaseI footprinting experiments in the hairpin motif for polyamides 5 and 6 respectively, on the 278 bp, 5'-end-labeled PCR product of plasmid pEF16: lane 1, intact DNA; lanes 2 and 3, G and A sequencing 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 concentration, respectively. Each footprinting gel is accompanied by the following: (top) ball-and-stick models of the compound bound to its target DNA sequence; and (bottom) Binding isotherms for the four designed sites. $\theta_{app}$ values were obtained according to published methods. Imidazoles and pyrroles are shown as filled and non-filled circles, respectively; dimer units are represented as rectangles containing either Bi enclosed in a circle representing benzimidazole or Ip in a black box representing imidazopyridine; beta alanine is shown as a diamond; the gamma-aminobutyric acid turn residue is shown as a semicircle connecting the two subunits.
7.3 Conclusion

While traditional Py, Im, and Hp containing polyamides have been successful at recognizing hundreds of pre-determined DNA sequences with high affinity and specificity, a host of target sequences such as 5'-GGGG-3' have proven difficult to code for using polyamides. A series of novel oligomers containing the 6–5 fused benzimidazole (Bi) and imidazopyridine (Ip) heterocycles were developed. These oligomers, composed of 5-membered heterocyclic carboxamides, demonstrated a substantial increase in affinity (10- to 100-fold) for the 5'-GGGG-3' sequence. The marked increase in affinity could be attributed to a combination of oligomer properties including a larger hydrophobic surface, a high degree of ligand pre-organization, or differential solvation/desolvation effects. The ability of this class of new oligomers to inhibit protein–DNA binding was demonstrated by the inhibition of NF-κB. We are encouraged by the fact that these oligomers demonstrate improved affinity for guanine rich DNA sequences and future work directed toward improving sequence specificity and examination of the nuclear trafficking ability is a priority.

7.4 Experimental

7.4.1 General

*N,N*-Dimethylformamide (DMF), *N,N*-diisopropylethylamine (DIEA), *N,N*-dimethylaminopropylamine (Dp), triethylamine (TEA), nitrobenzene (NO₂Ph), 2-formyl-N-

Table 7.1 Affinities of 5’-GGGG-3’ binding oligomers $K_a$ (M⁻¹)ₐ,b

<table>
<thead>
<tr>
<th>Polyamide</th>
<th>5’-atGGGGt-3’</th>
<th>5’-aatGGGGt-3’</th>
<th>5’-atGGGAt-3’</th>
<th>5’-atGAGGt-3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.4±1.0 x 10⁸</td>
<td>2.6±1.1 x 10⁸</td>
<td>2.3±0.8 x 10⁷</td>
<td>≤ 1.0 x 10⁷</td>
</tr>
<tr>
<td>2</td>
<td>1.9±1.4 x 10¹⁰</td>
<td>2.0±1.1 x 10¹⁰</td>
<td>4.8±1.1 x 10⁹</td>
<td>3.6±0.9 x 10⁹</td>
</tr>
<tr>
<td>3</td>
<td>2.6±0.6 x 10⁹</td>
<td>2.7±0.4 x 10⁹</td>
<td>2.6±0.2 x 10⁹</td>
<td>2.9±0.1 x 10⁹</td>
</tr>
<tr>
<td>4</td>
<td>4.4±0.9 x 10⁹</td>
<td>4.1±1.2 x 10⁹</td>
<td>8.1±2.3 x 10⁹</td>
<td>8.2±1.3 x 10⁹</td>
</tr>
<tr>
<td>5</td>
<td>2.6±0.5 x 10⁹</td>
<td>2.9±0.2 x 10⁹</td>
<td>8.6±2.1 x 10⁹</td>
<td>2.5±0.8 x 10⁹</td>
</tr>
<tr>
<td>6</td>
<td>1.6±0.2 x 10⁹</td>
<td>1.1±0.4 x 10⁹</td>
<td>1.3±0.1 x 10⁹</td>
<td>1.1±0.1 x 10⁹</td>
</tr>
</tbody>
</table>

a) Values reported are the mean values from at least three DNase I footprinting titration experiments, with the standard deviation given in parentheses.
b) Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂ at pH 7.0.
Figure 7.8  
a) Match DNA sequence with the p50 protein of NF-κB overlapping the oligomer binding site. Asterisks indicate the location of radio-labeled nucleotides in the probe sequence.

\[
\begin{align*}
5' &\text{ AATTACATGCAGTTGA} &\text{GGGG} &\text{ATTTCCCAAGGATGGCAAGCT}^* \\
3' &\text{TTAGTACGTCAACT} &\text{CCCC} &\text{TAGAAGGTCCGTACGTTCGA} \\
\end{align*}
\]

b) Mismatch DNA sequence.

\[
\begin{align*}
5' &\text{AATTACATGCAGTTGA} &\text{GGTT} &\text{ATTTCCCAAGGATGGCAAGCT}^* \\
3' &\text{TTAGTACGTCAACT} &\text{CCAA} &\text{TGGAGGTCGTACGTTCGA} \\
\end{align*}
\]

c) Gel shift screen for compounds 1–6 at concentrations of 10 and 100 nM.

d) Plot of relative NF-κB inhibition for compounds 1–6.
methylimidazole, red fuming nitric acid, 1,3-dichloro-4-nitropyridine, 30% bromine in acetic acid, palladium acetate Pd(OAc)$_2$, and 10% palladium on carbon were purchased from Aldrich. Boc-$\beta$-alanine-(4-carbonylaminomethyl)-benzyl-ester-copoly(styrene–divinylbenzene) resin (Boc-$\beta$-Pam-resin), dicyclohexylcarbodiimide (DCC), hydroxybenzotriazole (HOBt), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyuronium hexafluorophosphate (HBTU), $N,N$-dimethylaminopyridine (DMAP), and Boc-$\beta$-alanine were purchased from NOVA Biochem. Trifluoroacetic acid (TFA) was

**Figure 7.9** (Top) ball-and-stick model for 2 and EC$_{50}$ value from gel shift experiment. (Middle) representative NF-$\kappa$B titration gel ($n = 3$) for 2. (Bottom) binding isotherm for 2.

**Figure 7.10** Antibody supershift on match DNA. The NF-$\kappa$B band is shifted in the presence of anti-p50 or anti-p65 antibody.
purchased from Halocarbon. All other solvents were reagent grade from EM. Oligonucleotide inserts were synthesized by the Biopolymer Synthesis Center at the California Institute of Technology. Precoated silica gel plates 60F_{254} for TLC and silica gel 60 (40 μm) for flash chromatography were from Merck. Glycogen (20 mg/ml), dNTPs (PCR nucleotide mix), and all enzymes, unless otherwise stated, were purchased from Boehringer–Mannheim. pUC19 was purchased from New England Biolabs and deoxyadenosine [γ-32P] triphosphate was provided by ICN. Calf thymus DNA (sonicated, deproteinized) was obtained from Amersham Pharmacia. DNase I (7500 U/ml, FPLC pure) was purchased from Roche. AmpliTaq DNA polymerase was from Perkin-Elmer and used with the provided buffers. Tris–HCl, DTT, RNase-free water, and 0.5 M EDTA were from United States Biochemical. Calcium chloride, potassium chloride, and magnesium chloride were purchased from Fluka. Tris–borate–EDTA was from Gibco and bromophenol blue was from Acros. All reagents were used without further purification. NMR spectra were recorded on a Varian spectrometer at 300 MHz in DMSO-d₆ or CDCl₃ with chemical shifts reported in parts per million relative to residual solvent. UV spectra were measured on a Hewlett-Packard (model 8452 A) diode array spectrophotometer. High resolution FAB and EI mass spectra were recorded at the Mass Spectroscopy Laboratory at the California Institute of Technology. Matrix-assisted, laser desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS) was conducted at the Mass Spectroscopy Laboratory at the California Institute of Technology.

### 7.4.2 Heterocycle Synthesis

Heterocyclic building blocks Boc-Im-OH and Boc-Py-Bi-OH were synthesized as reported. Im-Im-OH (1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-imidazole-2-carboxylic acid, CAS [464892-44-2]) and Boc-Im-Im-OH (4-(4-(tert-butoxycarbonylamino)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-imidazole-2-carboxylic acid, CAS [502170-49-2]), are commercially available.

#### 7.4.2.1 1-methyl-4-nitro-1H-imidazole-2-carbaldehyde (NO₂-Im-CHO) (10)

A cooled flask (0 °C) of 1-methyl-2-imidazole-carboxaldehyde (7) (8g, 72.6 mmol, Aldrich) was treated dropwise with a precooled (0 °C) solution of red fuming nitric acid (75 ml) in conc. H₂SO₄•SO₃ (30%) (75 ml). The mixture was warmed to room temperature and stirred for 12 h open to the atmosphere. Next, the mixture was poured over ice, neutralized with solid sodium carbonate, extracted 4 times with dichloromethane, dried over anhydrous sodium sulfate, and concentrated in
vacuo to give a brownish-yellow oil. The oil was recrystallized from iPrOH/Et₂O or EtOH/Et₂O to give 1-methyl-4-nitro-1H-imidazole-2-carbaldehyde (10) as a tan crystalline solid (4.5 g, 40% Yield). TLC (1:1 EtOAc/Hex) Rₐ = 0.4; ¹H NMR (300 MHz, DMSO-d₆) δ 9.74 (s, 1H), 8.71 (s, 1H), 3.99 (s, 3H); ¹³C (75 MHz, DMSO-d₆) δ 182.31, 146.14, 140.56, 127.33, 35.70; HR-MS (EI⁺): calculated for C₇H₅O₃N₃: 155.0330; found: 155.0350.

7.4.2.2 2-(1-Methyl-4-nitroimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (NO₂-Im-Ip-OMe) (11).

1-Methyl-4-nitro-1H-imidazole-2-carbaldehyde (2.01 g, 13.0 mmol) (10) and methyl 5,6-diaminopyridine-2-carboxylate (8) (2.17 g, 13.0 mmol) suspended in 120 ml of nitrobenzene was heated to 140 ºC for 48 h open to the atmosphere. The reaction mixture was cooled to 23 ºC and the precipitate collected by vacuum filtration. The solid was washed with diethyl ether and dried under high vacuum to provide 2-(1-methyl-4-nitroimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (11) (3.6 g, 92% Yield) as a powdery tan solid. ¹H NMR (300 MHz, DMSO-d₆) δ 14.02 (broad s, 1H), 8.03-8.06 (m, 2H), 4.27 (s, 3H), 3.91 (s, 3H); ¹³C (75 MHz, DMSO-d₆) δ 166.01, 146.44, 142.95, 136.59, 135.80, 130.36, 127.35, 123.81, 121.17, 121.04, 52.85, 37.43; HR-MS (EI⁺): calculated for C₁₂H₁₀N₆O₄: 302.0760; found: 302.0760.

7.4.2.3 2-{4-Amino-1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (H₂N-Im-Ip-OMe).

2-(1-Methyl-4-nitroimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (11) (3 g, 9.9 mmol) was dissolved in anhydrous DMF (150 ml) and the solution was degassed with Ar. After the addition of Pd/C (10 wt. %, 600 mg) the reaction mixture was purged 3 times with hydrogen and then left to stir at 23ºC for 9 h under a hydrogen balloon atmosphere. After filtering through a pad of Celite and washing with copious amounts EtOAc the filtrate was concentrated in vacuo to give 2-{4-amino-1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester, without further purification (2.7 g, 100% Yield). ¹³C NMR (75 MHz, DMSO-d₆) δ 166.17, 149.03, 148.21, 141.40, 131.50, 120.12, 107.77; HR-MS (EI⁺): calculated for C₁₂H₁₀N₆O₂: 272.1020; found: 272.1030.

7.4.2.4 2-{4-{(tert-Butoxy)carbonylamino]-1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (Boc-Im-Ip-OMe).
2-{4-amino-1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (11) (2.0 g, 7.3 mmol) dissolved in DMF (25 ml) was treated with Boc2O (5.3 g, 24.3 mmol), DIEA (5.2 ml), and DMAP (95 mg, 0.73 mmol). The reaction mixture was then heated to 80 ºC for 72 h, cooled to 23 ºC, and flashed through a plug of silica gel eluting with EtOAc to give a mixture of mono- and di-bocked (2-{4-[(tert-butoxy)carbonylamino]-1-methylimidazol-2-yl)-3-[(tert-butoxy) carbonylamino]-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester) products which were carried on for saponification.

7.4.2.5 2-{4-[(tert-Butoxy)carbonylamino]-1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic (Boc-Im-Ip-OH) (12).

2-{4-[(tert-Butoxy)carbonylamino]-1-methylimidazol-2-yl)-3-{[(tert-butoxy)carbonylamino]-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester dissolved in MeOH (10 ml) and NaOH (1 N, 25 ml) was heated to 50 ºC for 4 h. The reaction mixture was cooled to 0 ºC and the pH adjusted slowly to pH = 4 with 1 N HCl. The reaction mixture was then extracted with ethyl acetate (4 times), dried over anhydrous sodium sulfate, concentrated in vacuo, and dried under high vacuum to give 2-{4-[(tert-butoxy)carbonylamino]-1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic (12) (258 mg, 60% Yield) as a light yellow solid. 1H NMR (300 MHz, DMSO-d6) δ 13.19 (broad s, 2H), 9.53 (s, 1H), 8.00 (m, 2H), 7.36 (s, 1H), 4.15 (s, 3H), 1.47 (s, 9H); 13C (75 MHz, DMSO-d6) δ 166.61, 152.92, 147.81, 142.48, 138.38, 132.01, 128.91, 119.74, 113.50, 79.07, 35.35, 28.12; HR-MS (EI+) calculated for C16H18N6O4: 358.1390; found: 358.1370.

7.4.2.6 2-(1-Methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (Im-Ip-OMe).

1-Methylimidazole-2-carbaldehyde (7) (214 mg, 1.85 mmol) and methyl 5,6-diaminopyridine-2-carboxylate (8) (310 mg, 1.85 mmol) were suspended in 17 ml of DMF and heated to 80 ºC for 1 hour open to the atmosphere. Next, FeCl3•6H2O (24mg, 0.09 mmol) was added and the reaction mixture was heated to 100 ºC for 12 hrs while air was bubbled through the reaction mixture. The reaction mixture was then cooled to room temperature and poured over ice. The precipitate was collected by filtration and washed with cold diethyl ether. The material was dissolved in hot iso-propanol, cooled to room temperature and re-precipitated with diethyl ether. The solid was washed with diethyl ether and dried under high vacuum to provide 2-(1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (Im-Ip-OMe) (240 mg, 50% Yield) as a tan
solid. $^1$H NMR (300 MHz, DMSO-d$_6$) $\delta$ 7.99 (s, 2 H), 7.51 (s, 1 H), 7.20 (s, 1 H), 4.18 (s, 3 H), 3.88 (s, 3 H); $^{13}$C (75 MHz, DMSO-d$_6$) $\delta$ 147.7, 142.0, 137.1, 131.8, 129.6, 126.9, 119.9, 109.3, 52.8, 35.8; HR-MS (EI$^+$): calculated for C$_{12}$H$_{11}$N$_5$O$_2$: 257.0910; found: 257.0920.

7.4.2.7 2-(1-Methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid (Im-Ip-OH) (9).

2-(1-Methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid methyl ester (Im-Ip-OMe) (250 mg, 0.97 mmol) dissolved in MeOH (2 ml) and KOH (4 N, 3 ml) was heated to 50 ºC for 4 h. The methanol was removed in vacuo and the aqueous layer washed with EtOAc (2 x 10 mL) to remove any starting material and trace impurities. The pH of the aqueous layer was then adjusted slowly to pH = 4 with 1 N HCl upon which time a cloudy beige precipitate formed. The mixture was placed in a falcon tube and the precipitate concentrated by centrifugation. The supernatant was decanted and the solid dried under high vacuum to give 2-(1-methylimidazol-2-yl)-3H-imidazo[4,5-b]pyridine-5-carboxylic acid (9) (154 mg, 65% Yield) as a brown solid. $^1$H NMR (300 MHz, DMSO-d$_6$) $\delta$ 7.99 (s, 2 H), 7.51 (d, 1 H, $J = 0.9$ Hz), 7.20 (d, 1 H, $J = 0.9$ Hz), 4.18 (s, 3 H); HR-MS (EI$^+$): calculated for C$_{11}$H$_9$N$_5$O$_2$: 243.0750; found: 243.0740.

7.4.3 Oligomer Synthesis

Oligomers were synthesized on solid support using Boc-β-PAM resin (0.59 meq/g). Stepwise elongation of the oligomers was done according to previously published protocols.$^{19}$ The synthesis of compound 1 has been previously reported.$^{14}$

7.4.3.1 Preparation of Base Resin R-β-Bi-Py-Bi-Py-γ-NHBoc (BR1).

To a manual solid phase synthesis vessel was added Boc-β-PAM resin (0.3 g). The resin was washed with DMF (15 mL) and allowed to swell for 15 min while shaking at room temperature. The resin was then washed with DCM (~30 mL), followed by 80% TFA in DCM (~30 mL) to remove the Boc-group. The resin was then agitated at room temperature in 80% TFA/DCM for another 25 min to provide the deprotected resin bound amine (R–β-NH$_2$). Following Boc-deprotection, the resin was washed with DCM and 10% DIEA in DMF to neutralize and prepare for coupling. Simultaneously, in a separate reaction vessel, Boc-Py-Bi-OH (189 mg, 531 μM), HBTU (191 mg, 504 μM), DIEA (137 mg, 185 μL, 1.06 mM) and DMF (1.2 mL) was mixed and allowed to activate at room temperature for 25 min. This mixture was then added to the solid phase synthesis vessel containing R–β-NH$_2$. Coupling was allowed to proceed at room temperature with agitation for 3-6
h. Initial loading of the resin requires elongated coupling times. Following coupling, the resin was acylated by the addition of acetic anhydride to the mixture and shaking for 15 min. The addition of the next Boc-Py-Bi-OH (25) dimer was incorporated and deprotected as described above to provide the resin bound fragment (R-β-Bi-Py-Bi-Py-NH₂). To this fragment was added a preactivated mixture of Boc-γ-OH (180 mg, 885 μM), HBTU (319 mg, 841 μM), DIEA (229 mg, 308 μL, 1.77 mM). Coupling was allowed to proceed for 3 h at room temperature with agitation. The resin was then capped with acetic anhydride as described above to provide the base resin R-β-Bi-Py-Bi-Py-γ-NHBoc (BR1). BR1 was then washed with DCM followed by MeOH and Et2O. The resin was then dried under high vacuum and stored for subsequent use.

7.4.3.2 Im-Im-Im-Im-γ-Py-Bi-Py-Bi-β-Dp (2).

BR1 (50 mg) was added to a manual solid phase synthesis vessel. The resin was washed with DCM (~15 mL), followed by deprotection with 80% TFA in DCM. The resin was shaken at room temperature in the 80% TFA solution for 25 min. The resin was then drained, washed with DCM, and neutralized with 10% DIEA in DMF. A pre-activated mixture of Boc-Im-Im-OH (54 mg, 148 μM), HBTU (53 mg, 140 μM), DIEA (38 mg, 52 μL, 295 μM) and DMF (400 μL) was then added to the reaction vessel and coupling was allowed to proceed for 3 h at room temperature, followed by capping with acetic anhydride as described for BR1 to give R-β-Bi-Py-Bi-Py-γ-Im-Im-NHBoc. Following resin deprotection as described above, Im-Im-OH was activated as described for Boc-Im-Im-OH. Coupling of Im-Im-OH to the resin was allowed to proceed overnight at room temperature to provide R-β-Bi-Py-Bi-Py-γ-Im-Im-Im. The resin was treated with the cleavage protocol outlined below to provide Im-Im-Im-Im-γ-Py-Bi-Py-Bi-β-Dp (2) in 5% yield. MALDI-TOF-MS: calculated for C₅₈H₆₆N₂₃O₈: 1212.55; found 1212.50 [M+H]⁺.

7.4.3.3 Im-Ip-Im-Im-γ-Py-Bi-Py-Bi-β-Dp (5).

BR1 (50 mg) was added to a manual solid phase synthesis vessel and R-β-Bi-Py-Bi-Py-γ-Im-Im-NHBoc was prepared as described above for 2. Following deprotection, washing and neutralization as described above, a pre-activated mixture of Im-Ip-OH (14) (21.5 mg, 88.5 μM), HBTU (32 mg, 84 μM), DIEA (23 mg, 31 μL, 177 μM), DMF (400 μL) was added to the vessel containing R-β-Bi-Py-Bi-Py-γ-Im-Im-NH₂. Coupling was allowed to proceed overnight at room temperature to provide R-β-Bi-Py-Bi-Py-γ-Im-Im-Ip. The resin was treated with the cleavage protocol outlined below to provide Im-Ip-Im-Im-γ-Py-Bi-Py-Bi-β-Dp 5 in 2.2% yield. MALDI-TOF-MS: calculated
for \( \text{C}_{59}\text{H}_{64}\text{N}_{23}\text{O}_7 \): 1206.54; found 1206.60 [M+H]^+.

### 7.4.3.4 Im-Im-Im-Ip-γ-Py-Bi-Py-Bi-β-Dp (3).

**BR1** (50 mg, 0.81 meq/g) was added to a manual solid phase synthesis vessel. The resin was treated with the cleavage protocol outlined below to provide Im-Im-Im-γ-Py-Bi-Py-Bi-β-Dp 3 in 3% yield. MALDI-TOF-MS: calculated for \( \text{C}_{59}\text{H}_{64}\text{N}_{23}\text{O}_7 \): 1206.50; found 1206.50 [M+H]^+.

### 7.4.3.5 Im-Im-Ip-Im-γ-Py-Bi-Py-Bi-β-Dp (4).

**BR1** (70 mg, 0.59 meq/g) was added to a manual solid phase synthesis vessel. The resin was washed with DCM (~15 mL), followed by deprotection with 80% TFA in DCM. The resin was then drained, washed with DCM, neutralized with 50% DIEA in DCM, and washed with DMF. A pre-activated mixture of Boc-Im-OH (50 mg, 207 \( \mu \)mol), HBTU (79 mg, 208 \( \mu \)mol), DIEA (53 mg, 72 \( \mu \)L, 413 \( \mu \)mol) and DMF (900 \( \mu \)L) was then added to the reaction vessel and coupling was allowed to proceed for 12 h at room temperature, followed by capping with acetic anhydride as described for **BR1** to give \( \text{R-β-Bi-Py-Bi-Py-γ-Im-NHBoc} \). Following resin deprotection as described above, Boc-Im-Ip-OH (12) was activated as described for Boc-Im-OH. Coupling of 12 to the resin was allowed to proceed overnight at room temperature to provide \( \text{R-β-Bi-Py-Bi-Py-γ-Im-Im-NHBoc} \). Following resin deprotection as described above, Im-CCl\(_3\) (2-Trichloroacetyl-1-methylpyrrole) (47 mg, 207 \( \mu \)mol) and DIEA (53 mg, 72 \( \mu \)L, 413 \( \mu \)mol) were dissolved in NMP (900 \( \mu \)L) and added to the reaction vessel. Coupling of Im-CCl\(_3\) to the resin was allowed to proceed overnight at 32 \( ^\circ \)C to provide \( \text{R-β-Bi-Py-Bi-Py-γ-Im-Im} \). The resin was treated with the cleavage protocol outlined below to provide Im-Im-Ip-Im-γ-Py-Bi-Py-Bi-β-Dp 4 in 3% yield. MALDI-TOF-MS calculated for \( \text{C}_{59}\text{H}_{64}\text{N}_{23}\text{O}_7 \): 1206.54; found 1206.50 [M+H]^+.

### 7.4.3.6 ImIp-ImIp-γ-PyBi-PyBi-β-Dp (6).

**BR1** (50 mg) was added to a manual solid phase synthesis vessel. The resin was treated with the cleavage protocol outlined below to provide Im-Ip-Ip-γ-Py-Bi-Py-Bi-β-Dp 6 in 2% yield. MALDI-TOF-MS calculated for \( \text{C}_{60}\text{H}_{62}\text{N}_{23}\text{O}_6 \): 1200.52; found 1200.50 [M+H]^+.

### 7.4.4 Resin Cleavage Procedure

A sample of resin (20-100 mg) was washed with DCM followed by the addition of
dimethylaminopropylamine (Dp) (1 mL). The mixture was heated to 80 °C for 2 h with occasional agitation. The resin was then filtered and washed with 0.1% TFA in water (7 mL). The combined filtrate was collected and subjected to purification by reverse phase preparatory HPLC using a Waters C₁₈ column and 0.1% TFA/ACN solvent system. Appropriate fractions from the HPLC purification were checked for purity by analytical HPLC and characterized by MALDI-TOF spectroscopy. Pure fractions were then pooled, flash frozen using liquid nitrogen and lyophilized to a dry solid for later use.

7.4.5 Footprinting Experiments
Plasmids pEF16 was constructed using standard methods. DNase I footprint titrations were performed according to standard protocols.¹⁷

7.4.6 NF-κB Electrophoretic Mobility Shift Assay
7.4.6.1 Materials
Jurkat Nuclear Extract containing activated NF-κB was purchased from Active Motif (36013) and diluted as necessary just prior to use with Buffer C (20 mM Hepes pH 7.9, 420 mM KCl, 1.5 mM MgCl₂, 25% glycerol, 0.2 mM EDTA, 0.5 mM DTT, 0.2 mM PMSF, 1 μg/ml aprotinin). The final working concentration was 0.5-2.5 μg/μl. The identity of the NF-κB band was confirmed by antibody supershift (Figure 7.9) using antibodies against p50 (sc-7178, sc-114) and p65 (sc-372) from Santa Cruz Biotechnology.

7.4.6.2 Antibody Supershift
The NF-κB antibody supershift data is shown in Figure 7.9. NF-κB antibodies (sc-7178 and sc-114 against p50 and sc-372 against p65) came from Santa Cruz Biotechnology. The actin antibody was purchased from Sigma-Aldrich (A5441). Antibodies were diluted to 0.1 μg/μl just prior to use with ice cold PBS. Complete binding reactions contained 0.1 ng labeled probe, 10 mM Tris•HCl pH 7.5, 50 mM NaCl, 10% glycerol, 1% NP-40, 1 mM EDTA, 0.1 mg/ml poly(dI-dC), 1 μl antibody (0.1 μg), and 4 μg (2 μl) nuclear extract in a total volume of 10 μl. Nuclear extract was incubated with antibody at 4 °C for 1 hour followed by addition of binding buffer and probe. The complete reactions were allowed to incubate a further 30 min at room temperature, loaded onto pre-run 5% acrylamide, 5% glycerol gels, and resolved for 2 h at 150 volts. Gels were pre-run for 15-45 min prior to loading. The gel and running buffer was 24.8 mM Tris Base, 190 mM glycine, 1 mM 0.5
7.4.6.3 Sequence of Gel Shift Probes

Oligonucleotides were purchased from Integrated DNA Technologies. The size of both probes following labeling was 40 basepairs. The match probe contained an imbedded κB site from the intronic enhancer of the immunoglobulin κ light-chain gene (underlined below). As a control, the sequence of the mismatch probe contained a mutated κB that prevented NF-κB binding (changes bolded below).

**Match Probe:**

5’ AATTCATGCAGTTGAGGGGACTTTCCCAGGCATGCA 3’

**Match Complementary Sequence:**

5’ AGCTTGCATGCCTGGGAAAGTCCCCTCAACTGCATG 3’

**Mismatch Probe:**

5’ AATTCATGCAGTTGAGTTTTACTTTCCCAGGCATGCA 3’

**Mismatch Complementary Sequence:**

5’ AGCTTGCATGCCTGGGAAAGTAACTTTCCCAGGCATGCA 3’

7.4.6.4 Preparation of 3’ Labeled Probes

In separate reactions, 1.1 pmol of match or mismatch oligonucleotide was annealed to an equal amount of its complementary sequence by heating at 95 °C for 1 minute and cooling slowly to room temperature. The 3’ ends were radiolabeled with 32P using Sequenase Version 2.0 (Amersham) and α32P dATP (PerkinElmer). The resulting labeled probes were purified with G-50 Microspin Columns (Amersham). Probe concentration was estimated at 0.8 ng/μl assuming 100% recovery based on a calculated molecular weight of 24,662 g/mol. Just prior to use, labeled probes were diluted in water to 0.1 ng/μl.

7.4.6.5 Gelshift Screen

Dry HPLC purified aliquots of polyamide were dissolved in water and their concentration determined by measuring the absorbance at 310 nm (ε = 69,520 M⁻¹ C⁻¹). 10x solutions (1 μM and 100 nM) of each polyamide were prepared from serial dilutions of the concentrated stock. Binding reactions contained 0.1 ng labeled probe, 10 mM Tris•HCl pH 7.5, 50 mM NaCl, 10% glycerol, 1% NP-40, 1 mM EDTA, 0.1 mg/ml poly(dI-dC), 1 μl 10x polyamide (final concentration 10 nM or 100 nM),
and 3 μg (1-2 μl) nuclear extract in a total volume of 10 μl. Reactions were allowed to equilibrate at room temperature without nuclear extract for 3 h, followed by addition of protein. After a further 30 min at room temperature, the reactions were loaded and resolved on non-denaturing 5% acrylamide, 5% glycerol gels for 2-2.5 h at 150 volts. The gels were immediately dried and exposed to a phosphoimage storage plate for at least 8 h. The gel and running buffer was 24.8 mM Tris Base, 190 mM glycine, 1 mM 0.5 M EDTA. Gels were pre-run for 15 to 45 min prior to loading.

7.4.6.6 Gelshift Titration
Reactions were prepared as described above with a constant 0.1 ng probe and 3 μg nuclear extract, except that 10x solutions of polyamide were prepared to give a final concentration range of 500 pM to 500 nM.

7.4.6.7 Data Analysis
A Typhoon 8600 Variable Mode Imager was used to visualize the gels and band intensity was quantified using ImageQuant software version 5.1 from Molecular Dynamics. The fraction of DNA bound to NF-κB was calculated as the intensity ratio of the NF-κB shifted band to the sum total of all bands, shifted and unshifted. EC50 values were determined graphically as the polyamide concentration required to reduce NF-κB band intensity to half its value. The data was plotted vs. polyamide concentration and fit to the Hill equation (n=1) using Kaleidagraph software.

7.5 Notes and References


20. (a) Qu, X. G.; Ren, J. S.; Riccelli, P. V.; Benight, A. S.; Chaires, J. B. Enthalpy/entropy


Chapter 8: Fluorescent Sequence-Specific dsDNA Binding Oligomers

The text of this chapter was taken in part from a manuscript coauthored with Anne Viger and Peter B. Dervan* (Caltech)

Abstract

Sequence-specific detection methods for double-stranded DNA that obviate the need for denaturation would provide a powerful tool for bioorganic chemistry and genetics. As part of a sustained effort to develop sequence-specific fluorescent DNA detection methods, two programmable oligomers have been synthesized which target their respective sequences 5’-WTACGW-3’ and 5’-WGGGGW-3’ (W = A or T). The two oligomers were found to fluoresce weakly in the absence of DNA but showed significant fluorescence enhancement by the addition of match DNA. The fluorescence is shown to increase in a concentration-dependent manner, and the intensity varies depending on the number of mismatch sites incorporated into the DNA hairpins. This new class of oligomers provides a method to detect DNA sequences without denaturation and in the absence of conjugation to a dye molecule. This is a first step toward sequence-specific DNA-binding molecules containing a fluorescent switch integrated as part of the recognition modules.
8.1 Introduction

Sequence-specific detection methods for double-stranded DNA (dsDNA) that obviate the need for denaturation would provide useful tools for bioorganic chemistry and genetics. Previous efforts, such as molecular beacons or peptide nucleic acid-thiazole orange (PNA-TO) conjugates, require harsh denaturation conditions for hybridization to single-stranded DNA. Previous efforts from our laboratory for the sequence-specific detection of dsDNA have focused on pyrrole-imidazole (Py-Im) polyamide-fluorophore conjugates, such as tetramethylrhodamine (TMR) or thiazole orange (TO), that bind in the minor groove of DNA. TMR fluorescence was shown to be quenched when the fluorophore was covalently linked to the ring nitrogen of a pyrrole recognition element within a polyamide. Remarkably, fluorescence was restored in a sequence-dependent manner upon binding to dsDNA. Similarly, polyamide-TO intercalator conjugates also demonstrate fluorescence enhancement in the presence of match dsDNA.

Having established Py-Im polyamide-dye conjugates as a suitable platform for sequence-specific fluorescent dsDNA detection, we sought to develop a new class of fluorescent DNA binders wherein the fluorescent moiety is an integrated part of the recognition modules. We report here the design of sequence-specific fluorescent dsDNA-binding oligomers (Figure 8.1) which incorporate multiple 6-5 fused dimer recognition modules and show a marked fluorescent enhancement upon excitation at 340 nm in the presence of dsDNA.

8.2 Results and Discussion

Oligomer **O1** contains the chlorothiophene-benzimidazole (Ct-Bi-), pyrrole-imidazopyridine (-Py-Ip-), and imidazolehydroxybenzimidazole (-Im-Hz-) recognition modules, whereas oligomer **O2** contains imidazole-imidazopyridine (Im-Ip-) and two pyrrole-benzimidazole (-Py-Bi-) recognition modules. **Figure 8.1** Structure of oligomers. a) Oligomer O1 containing Ct-Bi-, -Py-Ip-, and -Im-Hz- recognition modules. b) Oligomer O2 containing Im-Ip and two -Py-Bi- recognition modules.
modules. The binding affinities of O1 and O2 targeted to two biologically important sequences, 5’-ATACGT-3’ (O1) and 5’-WGGGGW-3’ (O2), were determined to be $K_a = 1.6 \times 10^9$ M$^{-1}$ and $2.6 \times 10^9$ M$^{-1}$, respectively, by quantitative DNase I footprinting.$^9\text{--}^{11}$

A library of dsDNA hairpins containing six base-pair match and mismatch binding sites for O1 and O2 was used to investigate their emission properties (Figure 8.2). The dsDNA library for O1 and O2 contained match sites (1 and 8, respectively), single base-pair (bp) mismatch sites (2–5 and 10–12, respectively), double bp mismatch sites (6 and 13, respectively), and full mismatch sites (7 for both oligomers). The dsDNA 9 contains the 4-G match site of oligomer O2; however, the flanking sequence has been changed to emphasize the effect on binding. The presence of G•C bp under the tail is expected to lower the binding affinity of O2 as compared to that of dsDNA 8.$^{12}$

Oligomers O1 and O2 (1 μM concentration) were each incubated with an increasing concentration (1 nM to 1 μM) of dsDNA, and their emission spectra were recorded after excitation at 340 nm. The oligomers exhibited a marked increase in fluorescence upon addition of dsDNA containing their match site 1 and 8, respectively (Figures 8.3 and 8.4).$^{13}$ Oligomer O1 showed a moderate decrease in fluorescence intensity in the presence of dsDNA 2, but proved to be much more sensitive to the incorporation of single base-pair mismatches at the alternate positions in dsDNAs 3–5 (Figure 8.4a). The incorporation of multiple base-pair mismatches in dsDNAs 6 and 7 showed a significant reduction in fluorescence intensity for O1. Oligomer O2 exhibited a similar trend in sequence specificity, with a moderate decrease in fluorescence intensity observed upon incorporation of single base-pair mismatches (9–12) and a more significant decrease with multiple mismatches (13 and 7, Figure 8.4b).
8.3 Conclusion

Sequence-specific DNA binding molecules containing a fluorescent switch integrated as part of the recognition modules provides a method to detect DNA sequences without denaturation and in the absence of conjugation to a dye molecule. Fluorescent oligomers may be useful as site-specific chromosome paints for telomeric and centromeric repeats and could provide insight into cellular trafficking of DNA binding compounds.

8.4 Experimental

Calcium chloride, potassium chloride, and magnesium chloride were purchased from Fluka. Tris-HCl was purchased from United States Biochemical. All reagents were used without further purification. Water (18 MΩ) was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered. Reagent-grade chemicals were used as received, unless otherwise stated. Oligomers O1 and O2 were prepared by literature procedures.

UV spectra were measured on a Agilent model 8453 diode-array spectrophotometer (Figure 8.5). Fluorescence spectra were measured with a Jobin Yvon/SPEX Fluorolog spectrofluorimeter (Model FL3-11) equipped with a Hamamatsu R928 PMT. Samples were excited at 340 nm using 8 nm emission and excitation slits and luminescence was observed from 400 to 600 nm at room temperature.

All measurements were performed in TKMC buffer [10 mM Tris-HCl (pH 7.0), 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂]. The concentration of O1 or O2 was 1 μM and the volume of solution used was 50 μL for fluorescence measurements. The extinction coefficients for O1 and O2 were \( \varepsilon = 58,700 \text{ cm L mol}^{-1} \text{ at } 330 \text{ nm} \) and \( \varepsilon = 69,200 \text{ cm L mol}^{-1} \text{ at } 340 \text{ nm} \) respectively. The

Figure 8.3 Fluorescence emission spectra of O1 and O2 (1 μM) after 12 h incubation with their match binding site dsDNA (\( \lambda_{\text{ex}} \) 340 nm). a) Data for compound O1. b) Data for compound O2. The emission was shown to plateau beyond 1 equiv DNA. (See Section 8.6 Spectra and Supplemental Information for plots.)
concentration of hairpin DNA was varied from 1 nM to 1 μM. Solutions containing O1 and O2 in the presence of varying concentrations of hairpin DNA were allowed to equilibrate for 12 h prior to fluorescence measurements.

**Figure 8.4** Plot of dsDNA concentration versus normalized fluorescence for each dsDNA. a) Data for compound O1. b) Data for compound O2.

### 8.5 Notes and References


7. Approaches based on peptide-thiazole orange conjugates show enhanced fluorescence in the presence of calf thymus (CT) DNA; however, DNA sequence specificity has not been observed. Carreon, J. R.; Mahon, K. P., Jr.; Kelley, S. O. Thiazole orange-peptide conjugates: sensitivity of DNA binding to chemical structure. *Org. Lett.* 2004, 6, 517-519.


11. The sequence 5’-ATACGT-3’ is the hypoxia response element in the VEGF promoter and 5’-WGGGGW-3’ is a sequence within the NF-kB response element.9,10


13. Quenching in the absence of DNA could be attributed to intramolecular quenching or solvent assisted quenching. Studies are underway to elucidate the mechanism.

Figure 8.5 UV spectra for compounds O1 and O2.
Figure 8.6 Fluorescence emission as a function of equivalents of DNA for compounds O1 and O2.
Figure 8.7 Fluorescence emission spectra compound O1.
Figure 8.8 Fluorescence emission spectra compound O2.
Chapter 9: Polyamide/NCP Ligation and Profluorescent Azido-Carbostyrils

The research in this chapter on the NCP templated ligation of azide and alkyne containing polyamides was done in collaboration with Justin D. Cohen (Caltech).
The nucleosome core particle (NCP) represents a physiologically relevant fundamental repeating unit of chromatin. The ability to modulate the structure of chromatin using sequence specific DNA binding small molecules could have potential applications in gene regulation, biotechnology, physical biology, and macromolecular structure studies. The possibility of self-assembling small molecules inside living cells using biological architectures represents a novel approach to molecular recognition. Small molecule probes may be used to report on dynamic information, regulate structural changes, and provide stabilization for structure determination of inherently dynamic macromolecules. Sequence-specific polyamides are a modular platform and could be used as DNA-binding building blocks to template the self-assembly of larger polyamides that target higher order structures such as the NCP, the tetra-NCP, the chromatin fiber, or other diverse nanometer scale structures inside the cell. This study provides a proof-of-principle experiment demonstrating that polyamides can be self-assembled using a high-order nucleoprotein structure (NCP) as a template.
9.1 Introduction

9.1.1 Templated Dimerization of Polyamides

The ability to use double stranded DNA as a template for the ligation of two polyamides with bioorthogonal functionality has recently been demonstrated.\(^1,^2\) The 1,3-dipolar cycloaddition between an azide and alkyne was employed, resulting in a triazole linked tandem hairpin polyamide capable of targeting >10 base pairs of DNA with rate increases of 20,000-fold.\(^1,^2\) The ability to template the dimerization of two cell permeable molecules for the targeted downregulation of a gene represents a powerful strategy for the control of transcriptional regulated processes in biological systems. The Dervan laboratory has also shown that polyamides are capable of binding the nucleosome core particle (NCP) and that a polyamide clamp constructed from two 8-ring polyamides connected at the turns by a linker is capable of locking the DNA onto the histone octamer (Figure 9.1).\(^3,^4\) This NCP clamp was able to lock a complete turn of DNA onto the octamer significantly improving \textit{in vitro} stability. Unfortunately, NCP clamps are very high molecular weight branched oligomers (MW > 2500) and often have poor cell permeability profiles.\(^5,^6\) The ability to template the dimerization of two hairpin polyamides across the supergroove between the two gyres of DNA on the NCP would allow for the introduction of two lower molecular weight polyamides potentially obviating cell permeability and size issues.

Incorporation of reporting strategies into the ligation event such as a profluorescent linker, where fluorescence would be activated by the dimerization process, would allow for the direct monitoring of reaction kinetics and cellular localization. The clamp dimerization strategy is illustrated in Figure 9.2 and would provide an important tool for the study of nucleosome core

\textbf{Figure 9.1} High resolution crystal structure of an NCP bound polyamide clamp. Clamp is shown in red. DNA backbone shown in blue with minor and major grooves in gray. Side view (left) showing histone octamer core (orange, yellow, tan, and pink). Top view (right) with clamp bound in a minor supergroove on the proximal face. (PDB 1S32)
particle recognition and gene regulation in addition to serving as a macromolecular structural probe. The potential for gene regulation at the NCP level represents a unique physiologically relevant recognition platform distinct from that of linear DNA. In addition, NCP’s tend to be sensitive to radiation damage and this along with dynamic DNA dissociation significantly reduce crystal diffraction quality. Polyamide clamps provide a unique tool for structural biology due to their ability to increase the size, order, and resolution of NCP crystals, as evidenced by the 2 Å resolution structure in Figure 9.1, providing the potential for crystallization of larger physiologically relevant structures. One could dream of the possibility of crystallizing structures such as multiply linked histone octamers or NCP’s with bound transcription factors and clamps of this sort could be just the

Figure 9.2 Illustration of the overall clamp dimerization strategy. Nucleosomal DNA shown without histone octamer present for clarity. Initial polyamide binding is a reversible process along with unwinding and winding of the DNA from the histone octamer. Dimerization of the two polyamides via a bioorthogonal reaction templated by the NCP is irreversible resulting in a large, sequence specific, high affinity clamp. The clamp can serve to stabilize the NCP for structural studies or potentially downregulate gene expression. Direct kinetic information and cellular localization data can be obtained via dimerization induced profluorescent linker activation.
tool needed to gain structural insight into these complex biological architectures.

**Figure 9.3** a) Sequence of the 146 base pair fragment of α-satellite DNA used for crystallographic studies with the nucleosome clamp. In those structures, the clamp was found to bind in the homodimeric “supergroove” highlighted in yellow. Each of the other four supergrooves on the NCP are highlighted in purple, green, blue and red. a) Highlight of the sequences to which the nucleosome clamp was bound (left). At right is shown the two base pair mutation (in red) introduced so that the supergroove becomes heterodimeric. Located above each highlighted site is the polyamide designed to target that site. b) Analysis of the linker dependence of NCP templated ligations. Crystal structure view of the supergroove. The predicted distance between the amines of the turn are shown in green for binding at the original site, and in red if the sites are moved one base pair back. c) Modeling of the linkers and the calculated turn-to-turn length.
9.2 Results and Discussion

In order to ascertain whether the NCP could be used to template ligation reactions (Figure 9.3), a series of three azide containing polyamides and two alkyne containing polyamides were synthesized. Several linker lengths were used in order to examine the distance dependence of the reaction (Figure 9.4). Examination of the previous crystal structure data and computational modeling showed that only the azide containing the longest linker was expected to be capable of reacting. In addition, the previous study of DNA ligation had demonstrated that alkyne had reacted over 20,000 times faster than the alkyne containing an extra carbon unit and was expected to perform similarly in our experiments. The alkynes were purchased commercially and then coupled directly to the appropriate polyamides as described previously. The azides were synthesized using the scheme shown in Figure 9.5, 9.6, and 9.7.

The NCP was reconstituted as outlined in the supplemental information. A control sample containing only the 146 bp DNA and a second control containing

Figure 9.4 Analysis of the linker dependence of NCP templated ligations. Models and structure of the linkers with the calculated turn-to-turn length listed below each.

Figure 9.5 Synthesis of alkyl azido linkers.
only the histone octamer were prepared as well. Next, 200 pmol each of 4 and 5 was added to 40 pmol of the reconstituted NCP and the reaction was incubated for 5 h at 37°C. MALDI-TOF MS was used to analyze the reaction, and after 5 h peaks were observed corresponding to the expected mass of the ligation product ([M+Na]+ at 2687.77 and [M+H]+ at 2664.88, which is in agreement with the calculated mass of 2664.23) (Figure 9.8). Control samples were run under the same conditions containing either just the DNA, just the histone octamer, or only buffer. In all three cases the target mass was not observed upon addition of the polyamide azide and alkyne, indicating that the ligation reaction was dependent upon the presence of the fully reconstituted NCP (Figure 9.9). The lack of product observed in the DNA only control reaction is promising, indicating there is reaction specificity dependent on DNA geometrical constraints imposed by assembly of the NCP.

This is a DNA binding-site proximity-based reaction, dependent on the proper assembly of a higher order macromolecular structure with complete specificity over linear DNA. This study provides evidence that our reaction is taking place in a templated format and the sequence specificity of the polyamide reagents is ensuring that the reaction only takes place when the two targeted binding site sequences are properly aligned in 3-dimensional space.

The same ligation reaction was performed using all three azides and both alkynes. Of the six potential ligation reactions, product was only observed in the case of compounds 4 and 5. This result agrees with both computational work indicating the linkers for compounds 2 and 3 were too short to react, as well as prior work demonstrating that the alkyne used in 6 was significantly less reactive.

These results demonstrate the ability to perform templated ligation reactions on the NCP as confirmed by MALDI-MS analysis, however the small scale of the reaction prevented

Figure 9.6 Synthesis of azido-polyamides.
the quantitation of ligation yield. As previously discussed, a fluorescent reporter strategy utilizing pro-fluorescent probes could provide an alternative to MALDI-MS analysis, with far less background noise.

Coumarins have enjoyed widespread use as platforms for the discovery of fluorescent molecules (e.g. laser dyes). Reports of profluorescent coumarins, which upon chemical modification or reaction give rise to a fluorescent coumarin based molecules have been reported. Azido coumarins have been shown to react with alkynes via 1,3-dipolar cycloadditions giving rise to highly fluorescent products due to a change in donor acceptor properties of appended functionality (Figure 9.10). In addition, maleimide-functionalized coumarins are similarly non-fluorescent until conjugate addition with a thiol yields a fluorescent product. These reactions could potentially be used in a bioorthogonal profluorescent ligation strategy for templating the dimerization of polyamides on the nucleosome core particle as illustrated in Figure 9.2. The ability of the linker to act as a fluorescent switch offers the unique possibility of monitoring reaction progress and cellular localization using highly sensitive fluorescent techniques. To establish linker requirements for the coumarins, molecular modeling was done using Spartan ES software package and energy-minimized using an AM1 model, followed by ab initio calculations by means of the Hartree-Fock method using the 6-31G* Pople basis set. As shown in Figure 9.10, the fully compacted, shortest triazole-coumarin spans 13 Å, which is 2 Å longer than the 11.0 Å distance between the two 8-ring polyamide turn amines in the crystal structure (Figure 9.3). This posed a potential problem due to the rigidity of the of the coumarin system. When 6-ring hairpin polyamides were modeled in for the 8- ring polyamides on the

**Figure 9.7** Synthesis of alkynyl-polyamides.
crystal structure, the distance between turn amines increased to 17.8 Å. Replacement of 8-ring polyamides for 6-ring polyamides was a potential solution to the problem, but not ideal since all structural data to date has been generated for 8-ring polyamides bound to the NCP. After surveying the literature for potential profluorescent replacements the aza-analogues of coumarins (carbostyrils) were discovered. Carbostyrils\textsuperscript{4,11} exhibit photophysical properties similar to coumarins and can

Figure 9.8 DNA templated ligation on the NCP. The reaction of 4 and 5 with the NCP leads to dimer 7 which was observed by MALDI mass spectrometry. NCP templated ligation between different length azides and alkynes. The reaction product was only detected with polyamides 4 and 5. N.R. = no reaction.
be electronically tuned by substitution with electron withdrawing and donating functionality.\textsuperscript{8,12} Figure 9.11 shows the structure and linker distance comparison for carbostyrils versus coumarins. It appears that the carbostyril based linkers will be able to accommodate both 6- and 8-ring systems. In addition, synthetic ease, compact size, and a nitrogen handle for derivatization makes the carbostyrils an ideal candidate for use as a coumarin alternative. Preliminary efforts towards the synthesis of carbostyril (17) are presented in Figure 9.12.

Synthesis of carbostyril (17) (Figure 9.12) started with condensation of commercially available $N,N$-dimethyl-m-phenylenediamine (8) with dimethyl malonate at 200 °C to give 2,4-dihydroxy-7-(dimethylamino)quinoline (14) in 50% isolated yield after a single recrystallization. Chlorination of 14 under refluxing POCl$_3$ afforded 2,4-dichloro-7-(dimethylamino)quinoline (15) in 80% yield following recrystallization. Quinoline 15 was hydrolyzed under refluxing conditions in 6M HCl to afford 4-chloro-7-(dimethylamino)carbostyril (16) in 90% yield. Selective hydrolysis of the chlorine at position 2 is often explained by the observation that the 2 position is more reactive toward nucleophilic substitution than the 4 position. After comparing the proton NMR of compound 16 to that reported in the literature, an apparent discrepancy was realized. Since there was a possibility of displacement of chlorine at two different positions and NMR would not unambiguously resolve the regioselectivity issue, the compound was recrystallized from hot DMF.
and the crystal structure solved. The crystal structure and unit cell are shown in Figure 9.12. After unambiguously proving the identity of compound 16 it was subjected to azidification. Compound 16 was treated with excess sodium azide in d$_6$-DMSO and the reaction monitored by proton NMR. After heating for exactly 10.5 h at 120 ºC complete conversion to azide 17 was achieved by NMR. Compound 17 can be N-alkylated to provide a functional handle for conjugation to small molecules and macromolecules. Initial investigations into the photophysical properties of compound 17 are shown in Figure XX and demonstrate the ability of the probe to fluoresce upon cycloaddition with an alkyne substrate.

9.3 Conclusion

These studies have demonstrated the feasibility of the NCP templated ligation approach for the self assembly of polyamide dimers. Additionally, the development of a pro-fluorescent azido carbostyril provides a new tool for monitoring ligation reactions using fluorescence. This strategy offers an exciting opportunity for modifying gene expression in cells by the targeted self-assembly of polyamides on NCP’s.
9.4 Experimental

9.4.1 Materials and General Methods

Dicyclohexylcarbodiimide (DCC), 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), N-Hydroxybenzotriazole (HOBr), Fmoc-Dab(Boc)-OH and Boc-β-Ala-Pam resin were purchased from Peptides International. Benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate (PyBOP) was purchased from Novabiochem. Flouro-N,N,N',N'-tetramethylformamidinium hexafluorophosphate (TFFH) was purchased from Advanced ChemTech. O-(7-Azabenzotriazol-1-yl)-N,N,N',N'-tetramethyluronium hexafluorophosphate (HATU), 4-(dimethylamino)-pyridine (DMAP), N-hydroxysuccinimide (NHS), N,N-dimethylformamide (DMF), N-methylpyrrolidinone (NMP), N,N-dimethylpropylamine (Dp), N,N-diisopropylethylamine (DIEA), ethylene diamine, piperidine, and other miscellaneous chemicals were purchased from Aldrich and used without further purification. All other solvents were purchased from EM Sciences and were reagent grade. Trifluoroacetic acid (TFA) was purchased from Halocarbon.

1H NMR spectra were recorded using a 300MHz General Electric-QE NMR spectrometer. CDCl3 was obtained from Cambridge Isotope Laboratories. UV spectra were recorded in water using a Beckman Coulter DU 7400 Spectrophotometer. Matrix-assisted LA SER desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) was performed using an Applied Biosystems Voyager DE Pro Spectrometer. Electrospray ionization (ESI) mass spectrometry was performed by the Protein and Peptide Microanalytical Facility at the California Institute of Technology. Analytical High-Pressure Liquid Chromatography (HPLC) was performed with a Beckman Gold system using a Varian Microsorb-MV 100 C18 column (5µm particle size, 250 x 4.6mm). Preparative HPLC was performed using a Beckman Gold system with either a Waters Bondapak C18 column (15-

Figure 9.11 a) Coumarin linker length presented for comparison against carbostyril. b) Carbostyril linker distance shows a decrease in length of ~3.7 angstroms.
20 μm particle size, (25 x 100mm) or a Phenomenex Gemini C18 column (5 μm particle size, 250 x 21.2 mm). For both HPLC systems Solvent A was 0.1% (v/v) aqueous TFA and solvent B was acetonitrile. Analytical HPLC was done using a gradient of 1.85%/min of Solvent B starting from 0% over 35 min with a flowrate of 1.5mL/min. Preparative HPLC was typically done using a gradient of 1%/min of Solvent B for 20 min followed by a gradient of 0.3%/min for an additional 100 min at a flowrate of 8mL/min. Radioactive gels were imaged using a Molecular Dynamics 400S PhosphorImager.

Restriction endonucleases, deoxyribonucleotide triphosphates, DNAse I, Polynucleotide kinase (PNK), and glycogen were purchased from Roche. [α-32P]-Thymidine-5′-triphosphate (≥ 3000 Ci/mmol) and [α-32P]-Deoxyadenosine-5′-triphosphate (≥ 6000 Ci/mmol) were purchased from Perkin Elmer. [γ-32P]-Adenosine-5′-triphosphate (≥ 7000 Ci/mmol) was purchased from MP Biomedicals. Water was purified from a Millipore Mill-Q purification system for general use.

**Figure 9.12** a) Progress toward the synthesis of profluorescent azide 18. b) X-ray crystal structure and unit cell for structural proof of compound 16.
Figure 9.13 Photophysical properties of compound 17 before and after cycloaddition with alkynes 18 and 20.
Ultrapure RNAse/DNase free water from USB was used for biological work. All buffer reagents used were molecular biology grade. Buffers were sterilized using a Nalgene 0.2 μm cellulose filtration device.

9.4.2 Plasmids

pJDC1 and pJDC2 were constructed using 80mer oligonucleotides purchased from Integrated DNA Technologies. pUC19 plasmid was purchased from Sigma. JM109 Competent Cells (>108 cfu/μg) were purchased from Promega. A Rapid DNA ligation kit was purchased from Roche. Purification was done using a Promega Wizard Plus Midipreps DNA purification system. Sequence analysis was performed by the Sequence Analysis Facility at the Caltech.

9.4.3 Polyamide Synthesis

Polyamide synthesis was performed as previously reported. All polyamides were synthesized using Boc-β-Ala-PAM resin (~0.59 meq/g). The resin was initially swelled in DMF for 5 min in a glass reaction vessel fitted with a glass filter and stopcock. The vessel was drained and the resin washed twice with DCM. Deprotection with 80% TFA:DCM was performed for 20 min while the resin was shaken. Following deprotection, the resin was washed 2 x DCM, 1 x 4:1 DMF:DIEA, and 1 x DMF. Coupling of the Boc-Py-OBt pre-activated ester was performed using 1.8 eq of monomer in ~1 mL of NMP. Coupling of Boc-Im-OH, Boc-PyIm-OH, and other acids was done by first preactivating 1.8 eq of the monomer with 1.7 eq of HBTU, 5.4 eq of DIEA in ~2mL of NMP. The activation mixture was shaken for 20min before being filtered and added to the resin. Couplings were allowed to proceed for 2h except in the case of Im-OH which was allowed to react overnight. After each coupling step the resin was washed 2 x DMF and then 2 x DCM before the next deprotection step. Polyamides were cleaved from resin using 1.5 mL of Dp for 200 mg of resin at 55°C for eight h. Crude products were purified by preparative HPLC.

2: MALDI-TOF-MS calculated [M+H]⁺: 1348.63, observed 1348.54
3: MALDI-TOF-MS calculated [M+H]⁺: 1362.65, observed 1362.61
4: MALDI-TOF-MS calculated [M+H]⁺: 1376.66, observed 1376.73
5: MALDI-TOF-MS calculated [M+H]⁺: 1289.58, observed 1289.67
6: MALDI-TOF-MS calculated [M+H]⁺: 2664.23, observed 2664.88
9.4.4 Reconstitution of NCP

The NCP was reconstituted following established protocols. Briefly, 10 μg of the 146 bp DNA was 5’ radiolabelled as described above using Polynucleotide Kinase and purified using a Chroma Spin STE 10 column from BD Biosciences. Histone octamer which was obtained from the Luger lab was diluted to 1 μg / μL using 2 M NaCl TE Buffer. 5 μL of the DNA was added to 5 μL of 4M NaCl, 10 mM Tris, 1 mM EDTA solution to make a sample of 10 μL of DNA in 2 M NaCl TE buffer. Five samples were made up and 0, 0.6, 0.8, 1.0, or 1.2 μL of octamer (1 μg / μL) was added to each sample. The total volume was brought to 2 μL using 2M NaCl TE buffer. The samples were incubated for 1 hr before the addition of 12 μL, 6 μL, 6 μL, and 84 μL of dilution buffer (TE) with each addition followed by a 1 h incubation. The sample was heated to 37°C for 2 h before a final addition of 120 μL of dilution buffer to give a final NaCl concentration of .1 M. Reconstituted NCP samples were stored at 4°C. Non-denaturing PAGE was done on a 6% polyacrylamide TBE gel. 1 μL of each sample was diluted to 10 μL with 10 mM Tris, 20 mM NaCl, 10% glycerol, and .1% Igapal. The gel was run at 150 V for 20 min, dried and imaged.

9.4.5 NCP Templated Ligation Reactions

The NCP was reconstituted as described above with the sole difference that non-radiolabelled DNA was used. For each ligation reaction 200 pmol of each polyamide was lyophilized into an eppendorf tube and 40 pmol of the reconstituted NCP was added. Samples containing only the 146 bp DNA and or the histone octamer were prepared as negative controls. The reaction was incubated for 5 h at 37°C before MALDI-TOF analysis.

9.5 Notes and References


Appendix A: Next Generation Hairpin Polyamides with (R)-3,4-Diaminobutyric Acid Turn Unit

The text of this chapter was taken in part from a manuscript coauthored with Christian Dose, Michelle E. Farkas, and Peter B. Dervan* (Caltech)

Abstract

The characterization of a new class of pyrrole–imidazole hairpin polyamides with β-amino-γ-turn units for recognition of the DNA minor groove is reported. A library of eight hairpins containing (R)- and (S)-3,4-diaminobutyric acid (β-amino-γ-turn) has been synthesized, and the impact of the molecules on DNA-duplex stabilization was studied for comparison with the parent γ-aminobutyric acid (γ-turn) and standard (R)-2,4-diaminobutyric acid (α-amino-γ-turn)-linked eight-ring polyamides. For some, but not all, sequence compositions, melting temperature analyses have revealed that both enantiomeric forms of the β-amino-γ-turn increase the DNA-binding affinity of polyamides relative to the (R)-α-amino-γ-turn. The (R)-β-amine residue may be an attractive alternative for constructing hairpin polyamide conjugates. Biological assays have shown that (R)-β-amino-γ-turn hairpins are able to inhibit androgen receptor-mediated gene expression in cell culture similar to hairpins bearing the standard (R)-α-amino-γ-turn, from which we infer they are cell-permeable.
A.1 Introduction

The ability to modulate the expression of eukaryotic gene networks by small molecules is a challenge in the field of chemical biology. Hairpin pyrrole-imidazole polyamides are a class of programmable small molecules that bind to the minor groove of DNA with affinities similar to transcription factors and have been shown to inhibit gene expression in living cells by interfering with transcription factor/DNA interfaces.1 The DNA sequence specificity of polyamides arise from interactions of pairs of the aromatic amino acids N-methylpyrrole (Py), N-methylimidazole (Im), and N-methylhydroxypyrrole (Hp) with the edges of the Watson-Crick base pairs.2 The generality of the polyamide pairing rules has been demonstrated by numerous studies3 and applications of polyamide conjugates include DNA alkylations,4 DNA-templated ligations,5 sequence-specific DNA intercalators,6 fluorescent DNA paints,7 DNA nanoarchitectures,8 and transcription factor mimics.9 Efforts have been made to improve the DNA-binding properties of hairpin polyamides with modified turn units.10 Substitution of γ-aminobutyric acid (γ-turn) by (R)-2,4-diaminobutyric acid (α-amino-γ-turn) increases the DNA-binding affinity by ~15-fold.10b,11 In contrast, hairpins containing the opposite enantiomer, (S)-α-amino-γ-turn, result in diminished binding affinities. This decrease is most likely caused by an unfavorable steric clash of the amine residue with the DNA minor groove.10b Sugiyama and co-workers have introduced polyamides containing the α-hydroxy-γ-turn.10c These hairpins provide discrimination for A•T/T•A base pairs at the turn position, although a ~70-fold reduced DNA-binding affinity relative to analogue (R)-α-amino-γ-turn-linked polyamides has been observed.

Here we introduce a new class of hairpin polyamides which are linked by 3,4-diaminobutyric acid which results in a β-amine residue at the turn unit (β-amino-γ-turn) (Figure A.1). DNA-binding affinities of four different eight-ring polyamide core sequences (with incrementally increasing Im/Py pairs) have been investigated and were compared to analogue hairpins bearing the parent γ-turn and the standard (R)-α-amino-γ-turn. We show that, for certain series of hairpin polyamides, both enantiomers of the β-amino-γ-turn are able to increase the relative DNA-binding affinity. However, this is sequence context dependent. Biological assays revealed that hairpin polyamides bearing the (R)-β-amino-γ-turn are able to inhibit specific gene expression in cell culture, which is taken as evidence for cell permeability.

A.2 Results and Discussion

A.2.1 Thermal stabilization of DNA duplexes by hairpin polyamides
Hairpin polyamides 1-16 were synthesized with different Im/Py and Py/Py compositions targeted to the four DNA sequences with increasing G/C content 5'-TGTTCA-3', 5'-TGGTCA-3', 5'-TGGCGA-3', and 5'-TGGGGA-3' (Figure A.2). The energetics of DNA-binding properties of polyamides are typically characterized by quantitative DNase I footprint titrations. These measurements provide precise information regarding the affinity and specificity of DNA/polyamide complexes. Unfortunately, quantitative footprinting experiments revealed similar equilibrium association constants (K_a values ~ 2 x 10^{10} M^{-1}) for hairpins 1-8, reaching an upper limit of the standard procedure (see Section A.6 Supplemental Information). Prior results have shown that the increase in melting temperature (ΔT_m) of DNA duplexes bound by hairpin polyamides correlates with DNA-binding affinity and can be utilized to detect single base pair mismatched DNA/polyamide complexes. Accordingly, we have used melting temperature analysis for dissecting differences in DNA affinities of hairpin polyamides. Spectroscopic analyses were performed on 12mer DNA duplexes containing the appropriate match sequence in the absence and presence of polyamides in order to derive the desired ΔT_m values (Figure A.3). Table A.1 shows that all hairpins provided an increase in melting temperature, relative to the individual DNA duplexes, confirming the formation of DNA/polyamide complexes. As expected, spectroscopic analysis with (R)-α-amino-γ-turn hairpins revealed stronger stabilizations than the parent γ-turn analogues; for example, achiral polyamide 1 targeted to DNA sequence 5'-TGTCCA-3' resulted in a ΔT_m value of 15.9°C, while chiral hairpin (R)-α-2 led to a 3.6°C higher

Figure A.1 Schematic representation of hairpin polyamides with increased DNA-binding affinity caused by different γ-turn units. Hairpin polyamides targeted to DNA sequence 5'-TGGTCA-3' are shown as ball-and-stick models. Ball-and-stick representation legend: black and white circles represent N-methylimidazole and N-methylpyrrole units, respectively, half-circles represent γ-aminobutyric acid, white diamonds represent β-alanine units, and half-circles containing a cross represent 3-(dimethylamino)-1-propylamine (Dp) as tail.
melting temperature ($\Delta T_{m} = 19.5^\circ C$). Remarkably, melting temperature analyses in the presence of $\beta$-amino-$\gamma$-turn hairpins (S)-$\beta$-3 ($\Delta T_{m} = 20.9^\circ C$) and (R)-$\beta$-4 ($\Delta T_{m} = 22.2^\circ C$) revealed higher $\Delta T_{m}$ values compared to those for the $\alpha$-series hairpin (R)-$\alpha$-2 ($\Delta T_{m} = 19.5^\circ C$). The same trend was observed for hairpins 5-8 targeted to DNA sequence 5'-TGGTCA-3' (Table A.1, Figure A.2). First, it is noteworthy that both the (R)- and (S)-$\beta$-amino-$\gamma$-turn generated higher melting temperatures than the standard (R)-$\alpha$-amino-$\gamma$-turn. Second, the enhancement (relative to achiral hairpins) observed for the (R)-$\beta$-series is almost twice that of the (R)-$\alpha$-series targeted to DNA sequences 5'-TGTTCA-3' and 5'-TGGTCA-3'; for example, polyamide (R)-$\beta$-8 provided a $\Delta \Delta T_{m}$ value of 6.9°C, while the $\alpha$-series (R)-$\alpha$-6 led to a $\Delta \Delta T_{m}$ value of 3.5°C relative to achiral hairpin 5. Interestingly, by further increasing the amounts of Im/Py pairs in the polyamides, significantly less DNA duplex stabilizations have been observed. For example, achiral polyamide 9 and chiral hairpin (R)-$\alpha$-10 targeted to DNA sequence 5'-TGGGCA-3' yielded $T_{m}$ values of 8.6 and 13.2°C, while the $\beta$-series

Figure A.2 Chemical structures for hairpins 1-16 targeted to DNA sequences: (A) 5'-TGTTCA-3', (B) 5'-TGGTCA-3', (C) 5'-TGGGCA-3', and (D) 5'-TGGGGA-3'.
(S)-β-11 and (R)-β-12 led to ΔT_m values of 13.3 and 13.6°C, respectively (Table A.1). Even lower melting temperatures were observed for hairpins 13-16 designed to bind DNA sequence 5'-TGGGGA-3'. Both β-amino-γ-turns, as in (S)-β-15 (ΔT_m = 6.7°C) and (R)-β-16 (ΔT_m = 6.8°C), resulted in significantly lower ΔT_m values than the α-series analogue (R)-α-14 (ΔT_m = 9.1°C). These results imply that the impact of polyamide turn units on DNA-duplex stabilization is sequence context dependent.

The general increase in DNA-binding affinity for polyamides containing the (R)-α-substituted γ-turn, relative to achiral hairpins, is most likely caused by a superposition of favorable noncovalent interactions of the positively charged substituent and conformational preferences of the turn unit. The (R)-α-amino-γ-turn can exist in two different conformations, one orienting the α-ammonium in a pseudoequatorial position (Figure A.4A), which directs the substituent toward the wall of the minor groove with the potential of steric interactions. The alternate conformation places the α-amine residue in a pseudoaxial position out of the minor groove, orienting the β-methylene to the floor of the double helix (Figure A.4B). Modeling of the (S)-β-amino-γ-turn conformations suggests that the -ammonium in a pseudoaxial position is directed out of the minor groove (Figure A.4C) relieving the potential steric interactions with the wall in comparison to the α-series. In contrast, the (S)-β-amine in a pseudoequatorial orientation is following the curvature of the minor groove (Figure A.4D). The possibility for favorable noncovalent interactions should exist in both conformations without the detriment of steric interactions. As shown in Figure A.4E, the pseudoequatorial β-amine residue of the (R)-β-amino-γ-turn is well accommodated in the DNA minor groove, while the pseudoaxial position should result in a steric clash of the substituent with the groove floor (Figure A.4F). Previous results have shown that polyamides constructed with several continuous Im/Py pairs are overcurved with respect to the DNA minor groove, significantly influencing the DNA-binding affinity and sequence specificity. We assume that this curvature affects the alignment of the turn units in the DNA minor groove.
This is supported by the observation that the presence of fewer continuous Im’s improves the DNA affinity of β-amino-γ-turns while diminishing the affinity for α-amino-γ-turns, and vice versa. However, illustrative modeling is not sufficient to explain the sequence context dependence of chiral hairpin polyamides, highlighting the pressing need for high-resolution structural studies.

### A.2.2 Sequence specificity at the turn position

Hairpin polyamides containing the γ-turn have been shown to possess an equal preference for A•T/T•A over G•C/C•G base pairs at the turn position, presumably for steric reasons.\(^\text{16}\) Sugiyama’s α-hydroxy-γ-turns have been demonstrated to discriminate A•T versus T•A at the turn position.\(^\text{10a}\) In order to study the sequence specificity for polyamides 5-16, we performed melting temperature analyses in the presence of DNA duplexes bearing a T•A base pair at the turn position. Experiments involving hairpins 1-4 have been omitted due to the palindromic core sequence specified by the polyamides. As shown in Table A.1, most γ-turn and (R)-α-amino-γ-turn hairpins provided similar ΔTₘ values for T•A and A•T base pairs. In contrast, significantly lower thermal stabilizations for T•A over A•T base pairs were observed for β-amino-γ-turn-linked polyamides targeting DNA sequences 5’-TGGTCA-3’

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>A•T</th>
<th>T•A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5’-CGA TGGTCA AGC-3’</td>
<td>5’-CGA TGGTCA AGC-3’</td>
</tr>
<tr>
<td></td>
<td>Tₘ/°C, ΔTₘ/°C</td>
<td>Tₘ/°C, ΔTₘ/°C</td>
</tr>
<tr>
<td>—</td>
<td>54.0 (±0.2)</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>69.9 (±0.3)</td>
<td>15.9</td>
</tr>
<tr>
<td>—</td>
<td>73.5 (±0.2)</td>
<td>19.5</td>
</tr>
<tr>
<td>—</td>
<td>74.9 (±0.2)</td>
<td>20.9</td>
</tr>
<tr>
<td>—</td>
<td>76.2 (±0.2)</td>
<td>22.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>A•T</th>
<th>T•A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5’-CGA TGGTCA AGC-3’</td>
<td>5’-CGA TGGTCA AGC-3’</td>
</tr>
<tr>
<td></td>
<td>Tₘ/°C, ΔTₘ/°C</td>
<td>Tₘ/°C, ΔTₘ/°C</td>
</tr>
<tr>
<td>—</td>
<td>57.2 (±0.1)</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>70.6 (±0.2)</td>
<td>13.4</td>
</tr>
<tr>
<td>—</td>
<td>74.1 (±0.3)</td>
<td>16.9</td>
</tr>
<tr>
<td>—</td>
<td>76.1 (±0.2)</td>
<td>18.9</td>
</tr>
<tr>
<td>—</td>
<td>77.5 (±0.3)</td>
<td>20.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>A•T</th>
<th>T•A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5’-CGA TGGGCA AGC-3’</td>
<td>5’-CGA TGGGCT AGC-3’</td>
</tr>
<tr>
<td></td>
<td>Tₘ/°C, ΔTₘ/°C</td>
<td>Tₘ/°C, ΔTₘ/°C</td>
</tr>
<tr>
<td>—</td>
<td>60.2 (±0.2)</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>68.8 (±0.2)</td>
<td>8.6</td>
</tr>
<tr>
<td>—</td>
<td>73.4 (±0.2)</td>
<td>13.2</td>
</tr>
<tr>
<td>—</td>
<td>73.5 (±0.1)</td>
<td>13.3</td>
</tr>
<tr>
<td>—</td>
<td>73.8 (±0.1)</td>
<td>13.6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>A•T</th>
<th>T•A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5’-CGA TGGGGA AGC-3’</td>
<td>5’-CGA TGGGGT AGC-3’</td>
</tr>
<tr>
<td></td>
<td>Tₘ/°C, ΔTₘ/°C</td>
<td>Tₘ/°C, ΔTₘ/°C</td>
</tr>
<tr>
<td>—</td>
<td>57.5 (±0.1)</td>
<td>—</td>
</tr>
<tr>
<td>—</td>
<td>60.9 (±0.1)</td>
<td>3.4</td>
</tr>
<tr>
<td>—</td>
<td>66.6 (±0.1)</td>
<td>9.1</td>
</tr>
<tr>
<td>—</td>
<td>64.2 (±0.1)</td>
<td>6.7</td>
</tr>
<tr>
<td>—</td>
<td>64.3 (±0.3)</td>
<td>6.8</td>
</tr>
</tbody>
</table>

**Table A.1** Melting temperatures of DNA/polyamide complexes for A•T and T•A base pairs at the turn position of hairpin polyamides.\(^a\)

\(^a\) All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses (n.d. = not determined). ΔTₘ values are given as Tₘ(DNA/polyamide) – Tₘ(DNA).
Figure A.4 Illustrative models of different turn conformations for hairpin polyamides containing the (R)-α-amino-γ-turn (A and B), (S)-β-amino-γ-turn (C and D), and the (R)-β-amino-γ-turn (E and F) bound to the minor groove of DNA (dark gray = carbons, white = hydrogen, blue = nitrogen, red = oxygen).
(5-8) and 5’-TGGGCA-3’ (9-12). Even more diminished duplex stabilizations were observed in presence of C•G and G•C base pairs (see Section A.6 Supplemental Information). These observations suggest that polyamides containing β-amino-γ-turns prefer A•T > T•A >> C•G > G•C base pairs at the turn position. However, sequence specificity studies by thermal denaturation measurements require binding enthalpies ($\Delta H_b$) of DNA/polyamide complexes in order to determine equilibrium association constants. One could also imagine using six-ring hairpin polyamides with lower DNA-binding affinities in order to discriminate sequence specificities at the turn position by quantitative DNase I footprint titration methods.

**A.2.3 Acetylated chiral hairpin polyamides**

Several approaches have been reported wherein the (R)-α-amino-γ-turn was utilized as a position for synthetic modifications of hairpin polyamides. It has been shown that acetylation of the (R)-α-amine in six-ring hairpin polyamides results in ~15-fold reduced DNA-binding affinity. To study the tolerance of synthetic modifications for eight-ring polyamides containing the -amino-γ-turns, we examined acetylated hairpins 17-19 by melting temperature analysis (Figure A.5). Indeed, hairpin 17 containing the acetylated (R)-α-amino-γ-turn yielded a markedly lower $\Delta T_m$ value (12.8°C) than nonacetylated

![Figure A.5](image-url)
analogue 6 ($\Delta T_m = 16.9 \, ^\circ C$, Table A.2). Even more pronounced was the decrease in DNA duplex stabilization for acetylated (S)-$\beta$-amino-$\gamma$-turn hairpin 18 leading to a $\Delta T_m$ value of 11.7°C. Remarkably, the opposite enantiomer (R)-$\beta$-19 resulted in significantly less destabilization ($\Delta T_m = 17.8°C$). All hairpins lose the positive charge at the turn unit by acetylation. This implies that the cationic state of the amine residue is not the only contribution impacting the energetics of the DNA/polyamide complexes, as evidenced by the differences in melting temperatures between hairpins 17-19. Increased steric demands of the acetylated substituents may also be responsible for differing binding affinities, due to the restricting DNA minor groove and alternate conformations of the $\gamma$-turn units (Figure A.4).

### A.2.4 Biological assay for cell permeability

Hairpin polyamide conjugates bearing the standard (R)-$\alpha$-amino-$\gamma$-turn have been shown to modulate the expression of certain gene pathways in living cells by interfering with transcription factor/DNA interfaces.1 Recently, a hairpin designed to bind DNA sequence 5'-AGAACA-3', found in the androgen response element (ARE), has been demonstrated to inhibit androgen receptor-mediated expression of prostate specific antigen (PSA) in LNCaP cells (Figure A.6).1b We utilized this cell culture transcription assay to investigate the cell permeability of (R)-$\beta$-amino-$\gamma$-turn hairpins because small structural changes within polyamides can influence nuclear uptake properties.18 Hairpin polyamide 21 was examined in comparison to the previously used (R)-$\alpha$-amino-$\gamma$-turn hairpin.
20 (Figure A.7A). Chiral polyamides 22 and 23, designed to target different DNA sequences, have been used as controls. Melting temperature analyses for polyamide conjugates 20-23 confirmed the results obtained for hairpins 1-4, revealing highest DNA-duplex stabilizations for (R)-β-amino-γ-turn hairpins (see Section A.6 Supplemental Information). The induction of PSA mRNA by dihydrotestosterone (DHT) in LNCaP cells for matched and mismatched polyamides 20-23 was measured by quantitative real-time RT-PCR. As shown in Figure A.7B, hairpin 21 provided significant inhibition of AR-mediated expression of PSA mRNA, KLK2, FKBP5, and TMPRSS2 mRNA (see Section A.6 Supplemental Information) which supports cell-permeable properties for (R)-β-amino-γ-turn hairpins.
A.3 Conclusion

Herein we have introduced (R)- and (S)-β-amino-γ-turn hairpin polyamides. Eight new polyamides targeting different DNA-binding motifs have been synthesized, and their impact on DNA duplex stabilization in relation to hairpins containing the parent γ-turn and the standard (R)-α-amino-γ-turn was investigated. It was found that changing the turn unit from the (R)-α-amino-γ-turn to either enantiomeric forms of the β-amino-γ-turn increases the relative DNA-binding affinity of polyamides targeted to 5′-TGTTCA-3′ and 5′-TGGTCA-3′ but not to 5′-TGGGCA-3′ and 5′-TGGGGA-3′ sequences, rendering the impact of α-amino-substituted γ-turns sequence context dependent. Acetylation of the (S)-β-amino-γ-turn has been demonstrated to significantly impact the DNA-binding affinity but has minimal effect for the (R)-β-amino-γ-turn, which makes the (R)-β-amino residue attractive for synthetic modifications and conjugate design. Upper limits presented by DNase I footprinting titrations of high affinity binders rendered melting temperature analysis a more practical choice for dissecting improvements by structural changes of new turn units in hairpin polyamides. Due to the strong thermal stabilizations, reported for eight-ring hairpin polyamides 1-8 targeted to 5′-TGTTCA-3′ and 5′-TGGTCA-3′ sequences, it is not unreasonable to estimate that the DNA-binding equilibrium association constants are markedly higher than \(K_a \sim 2 \times 10^{10} \text{ M}^{-1}\). Biological experiments have demonstrated that (R)-β-amino-γ-turn hairpins possess biological activity to inhibit AR-mediated gene expression within a human cancer cell line and may have similar uptake properties as polyamides bearing the standard (R)-α-amino-γ-turn. Ongoing work is focused on the use of the next generation hairpins in biological investigations as well as turn unit sequence specificity and high-resolution crystallographic studies for DNA/chiral hairpin polyamide complexes. These efforts will be reported in due course.

A.4 Experimental

A.4.1 General

Chemicals and solvents were purchased from Sigma-Aldrich and were used without further purification. Boc-γ-Abu-OH was purchased from Novabiochem. (R)-2,4-Fmoc-Dbu(Boc)-OH and Boc-β-Ala-PAM resin were purchased from Peptides International. (R)-3,4-Cbz-Dbu(Boc)-OH and (S)-3,4-Cbz-Dbu(Boc)-OH were purchased from Senn Chemicals AG. All DNA oligomers were purchased HPLC purified from Integrated DNA Technologies. Water (18 MΩ) was purified using a Millipore MilliQ purification system. The pH of buffers was adjusted using a Beckman 340 pH/temp meter. Analytical HPLC was performed on a Beckman Gold system equipped with a
diode array detector using a Phenomenex Gemini column (5 μm particle size, C18 110A, 250 × 4.6 mm, 5 μm). Preparative HPLC was performed on a Beckman Gold system equipped with a single-wavelength detector monitoring at 310 nm using a Phenomenex Gemini column (5 μm particle size, C18 110A, 250 × 21.2 mm, 5 μm). For both analytical and preparative HPLC, solvent A was 0.1% (v/v) aqueous trifluoroacetic acid (TFA) and solvent B was acetonitrile. Solvent gradients were adjusted as needed. Polyamide concentrations were measured in 0.1% (v/v) aqueous TFA on a Hewlett-Packard diode array spectrophotometer “Model 8452 A” and were determined by using an extinction coefficient of 69200 M⁻¹·cm⁻¹ at λ_max near 310 nm. Matrix-assisted, LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on an Applied Biosystems Voyager DR Pro spectrometer using α-cyano-4-hydroxycinnamic acid as matrix.

A.4.2 Synthesis of polyamides
Polyamide monomers and oligomers were synthesized as described previously.19 All β-amino-γ-turn hairpins were synthesized by performing the following procedure: the polyamide was cleaved from the resin with 3-(dimethylamino)-1-propylamine, purified by preparative HPLC, and characterized by MALDI-TOF MS, UV-vis spectroscopy, and analytical HPLC. A 500 nmol fraction of the Cbz-protected hairpin polyamide was dissolved in a 9:1 mixture (500 μL) of TFA and trifluoromethanesulfonic acid (TFMSA). After 5 min reaction time, the solution was flash-frozen by liquid N₂ and overlaid with N,N'-dimethylformamide (1 mL). The thawed solution was diluted with 20% aqueous acetonitrile (8 mL), purified by preparative HPLC, and characterized by MALDI-TOF MS, UV-vis spectroscopy, and analytical HPLC. Acetylated polyamides 17-19 were synthesized by performing the following procedure: A 500 nmol fraction of the polyamide was dissolved in N,N'-dimethylformamide (900 μL) and a 9:1 mixture of pyridine/acetic anhydride (100 μL) was added. After 5 min reaction time, the solution was diluted with 10% aqueous TFA (8 mL), purified by preparative HPLC, and characterized by MALDI-TOF MS, UV-vis spectroscopy, and analytical HPLC. Polyamide conjugates 20-23 were synthesized as described previously.20 Polyamide 1: MALDI-TOF [M + H]⁺ calculated for C₅₈H₇₂N₂₁O₁₀⁺ = 1222.6, observed = 1222.7. Polyamide 2: MALDI-TOF [M + H]⁺ calculated for C₅₈H₇₃N₂₂O₁₀⁺ = 1237.6, observed = 1237.8. Polyamide 3: MALDI-TOF [M + H]⁺ calculated for C₅₈H₇₃N₂₂O₁₀⁺ = 1237.6, observed = 1237.8. Polyamide 4: MALDI-TOF [M + H]⁺ calculated for C₅₈H₇₃N₂₂O₁₀⁺ = 1237.6, observed = 1237.8. Polyamide 5: MALDI-TOF [M + H]⁺ calculated for C₅₇H₇₁N₂₂O₁₀⁺ = 1223.6, observed = 1223.5. Polyamide 6: MALDI-TOF [M + H]⁺ calculated for C₅₇H₇₂N₂₂O₁₀⁺ = 1238.6, observed = 1238.6.
Polyamide 7: MALDI-TOF [M + H]$^+$ calculated for C$_{57}$H$_{72}$N$_{23}$O$_{10}$$^+$ = 1238.6, observed = 1238.5.
Polyamide 8: MALDI-TOF [M + H]$^+$ calculated for C$_{57}$H$_{72}$N$_{23}$O$_{10}$$^+$ = 1238.6, observed = 1238.5.
Polyamide 9: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{70}$N$_{23}$O$_{10}$$^+$ = 1224.6, observed = 1224.8.
Polyamide 10: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{71}$N$_{24}$O$_{10}$$^+$ = 1239.6, observed = 1239.6.
Polyamide 11: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{71}$N$_{24}$O$_{10}$$^+$ = 1239.6, observed = 1239.5.
Polyamide 12: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{71}$N$_{24}$O$_{10}$$^+$ = 1239.6, observed = 1239.6.
Polyamide 13: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{70}$N$_{23}$O$_{10}$$^+$ = 1224.6, observed = 1224.6.
Polyamide 14: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{71}$N$_{24}$O$_{10}$$^+$ = 1239.6, observed = 1239.7.
Polyamide 15: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{71}$N$_{24}$O$_{10}$$^+$ = 1239.6, observed = 1239.4.
Polyamide 16: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{71}$N$_{24}$O$_{10}$$^+$ = 1239.6, observed = 1239.5.
Polyamide 17: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{74}$N$_{23}$O$_{11}$$^+$ = 1280.6, observed = 1280.6.
Polyamide 18: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{74}$N$_{23}$O$_{11}$$^+$ = 1280.6, observed = 1280.7.
Polyamide 19: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{74}$N$_{23}$O$_{11}$$^+$ = 1280.6, observed = 1280.6.
Polyamide 20: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{74}$N$_{23}$O$_{11}$$^+$ = 1280.6, observed = 1280.6.
Polyamide 21: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{74}$N$_{23}$O$_{11}$$^+$ = 1280.6, observed = 1280.6.
Polyamide 22: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{74}$N$_{23}$O$_{11}$$^+$ = 1280.6, observed = 1280.6.
Polyamide 23: MALDI-TOF [M + H]$^+$ calculated for C$_{56}$H$_{74}$N$_{23}$O$_{11}$$^+$ = 1280.6, observed = 1280.6.

A.4.3 UV Absorption Spectrophotometry

Melting temperature analysis was performed on a Varian Cary 100 spectrophotometer equipped with a thermocontrolled cell holder possessing a cell path length of 1 cm. The buffer for the spectroscopic measurements was chosen to match as closely as possible the conditions of DNase I footprinting experiments. We used 10 mM sodium cacodylate since the temperature dependence of Tris-HCl makes it poorly suited for melting temperature analyses. A degassed aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM MgCl$_2$, and 5 mM CaCl$_2$ at pH 7.0 was used as analysis buffer. DNA duplexes and hairpin polyamides were mixed in 1:1 stoichiometry to a final concentration of 2 μM for each experiment. Prior to analysis, samples were heated to 90 °C and cooled to a starting temperature of 25°C with a heating rate of 5°C/min for each ramp. Denaturation profiles were recorded at λ = 260 nm from 25 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.
A.4.4 Molecular Modeling

DNA/polyamide models are based on coordinates derived from NMR structure studies using standard bond length and angles. The molecular graphics images are nonminimized and have been created by introducing ammonium residues to the appropriate position of the turn unit using the UCSF Chimera package from the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIH P41 RR-01081).

A.4.5 Measurement of Androgen-Induced PSA mRNA

Experiments were performed as described previously.

A.5 Notes and References


13. For quantitative footprinting experiments, the DNA concentrations of equilibrium mixtures should be at least 10-fold less than the total association constant of the DNA/ligand complex in order to ensure the approximation \[\text{[ligand]}_{\text{free}} = \text{[ligand]}_{\text{total}}\] for numerical analysis. However, the concentration of the labeled DNA fragment specified by the standard DNA/polyamide footprinting protocol is \(~5\text{ pM}\). Consequently, DNA-association constants become compressed and hence unreliable for comparison studies at \(K_a\) values \(\geq 2 \times 10^{10} \text{ M}^{-1}\). (a) Brenowitz, M.; Senear, D. F.;


A.6 Supplemental Information

Construction of Plasmids pCDMF-1 and pCDMF-2: Oligonucleotides were purchased from Integrated DNA Technologies. The plasmids pCDMF-1 and pCDMF-2 were constructed by annealing the oligonucleotides: 5'-AGCTGCGGCTCGAGACGGCTAACCCATCGAGACGGCTACCCCATCGAGAGGATC-3' and 5-GATCGATCTCTCT-GATGGGGTAGCCGTCTCGAGGATAGCCGTCTCGATGGGCTA GCGCTTCAGTGGGTTA-GCCGCTTCGAGACGCCG-3'; 5'-AGCTGCGAGACGGCTCGAGACGGCTTGAACATCGAGACGGCTCGAGACGGCTACCCCATCGAGAGGATC-3' and 5-GATCGAGCCGTCTCGAGCCGTCTCGAGGATAGCCGTCTCGATGGGCTA GCGCTTCAGTGGGTTA-GCCGCTTCGAGACGCCG-3', respectively, followed by ligation into the BamHI/HindIII restriction fragment of pUC19 using T4 DNA ligase. The plasmid was then transformed into Escherichia coli JM109 competent cells. Ampicillin-resistant white colonies were selected from 25 mL Luria–Bertani (LB) agar plates containing 50 mg/mL ampicillin treated with XGAL and isopropyl-β-D-thiogalactopyranoside (IPTG) solutions and grown overnight at 37°C. Cells were harvested the following day and purification of the plasmid was performed with a Wizard Plus Midiprep DNA purification kit (Promega). DNA sequencing of the plasmid insert was performed by the sequence analysis facility at the California Institute of Technology.

DNase I Footprinting Titrations: Polyamide equilibrations and DNase I footprint titrations were conducted on the 5' end-labeled PCR product of pCDMF-1 and pCDMF-2 according to standard protocols. DNA was incubated with polyamides or water (control) for 12 h at room temperature prior to reaction with DNase I.

### Table A.3
Melting temperatures of DNA/polyamide complexes for all four base pair variations at the turn position of hairpin polyamides.\(^a\)

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>A+T</th>
<th>T+A</th>
<th>C+G</th>
<th>G+C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>T(_m)/°C</td>
<td>ΔT(_m)/°C</td>
<td>T(_m)/°C</td>
<td>ΔT(_m)/°C</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>—</td>
<td>57.2 (±0.1)</td>
<td>—</td>
<td>55.8 (±0.1)</td>
<td>—</td>
</tr>
<tr>
<td>(5)</td>
<td>70.6 (±0.2)</td>
<td>13.4</td>
<td>69.0 (±0.3)</td>
<td>13.2</td>
</tr>
<tr>
<td>(6)</td>
<td>74.1 (±0.3)</td>
<td>16.9</td>
<td>72.9 (±0.2)</td>
<td>17.1</td>
</tr>
<tr>
<td>(7)</td>
<td>76.1 (±0.2)</td>
<td>18.9</td>
<td>73.2 (±0.1)</td>
<td>17.4</td>
</tr>
<tr>
<td>(8)</td>
<td>77.5 (±0.3)</td>
<td>20.3</td>
<td>74.2 (±0.1)</td>
<td>18.4</td>
</tr>
</tbody>
</table>

\(^a\) All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses (n.d. = not determined). ΔT\(_m\) values are given as T\(_m\) (DNA/polyamide) − T\(_m\) (DNA).
Table A.4 Equilibrium association constants for hairpin polyamides determined by quantitative DNase I footprint titrations.a

<table>
<thead>
<tr>
<th>Polyamides</th>
<th>$A\cdot T$</th>
<th>$T\cdot A$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$K_a$ / M$^{-1}$</td>
<td>$K'_a$ / M$^{-1}$</td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.0 (±0.8) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.6 (±0.6) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.1 (±0.1) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.7 (±0.3) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(4)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.3 (±0.7) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.1 (±0.5) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.4 (±0.3) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.3 (±0.3) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(8)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>(9)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.d.</td>
<td>1.5 (±0.2) x 10$^{10}$</td>
</tr>
<tr>
<td>(10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.d.</td>
<td>3.0 (±0.4) x 10$^9$</td>
</tr>
<tr>
<td>(11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.d.</td>
<td>5.9 (±0.9) x 10$^9$</td>
</tr>
<tr>
<td>(12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.8 (±0.2) x 10$^{10}$</td>
<td>n.d.</td>
</tr>
<tr>
<td>(13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.d.</td>
<td>6.6 (±1.8) x 10$^9$</td>
</tr>
<tr>
<td>(14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.d.</td>
<td>9.4 (±3.0) x 10$^7$</td>
</tr>
<tr>
<td>(15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5'-CGA TGGTCA AGC-3'</td>
<td>5'-CGA TGGTCT AGC-3'</td>
<td></td>
</tr>
<tr>
<td></td>
<td>n.d.</td>
<td>2.1 (±0.6) x 10$^8$</td>
</tr>
<tr>
<td>(16)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a Equilibrium association constants reported are mean values from at least three quantitative DNase I footprint titration experiments. Standard deviations are shown in parentheses. b Equilibrium association constants have been reported previously11 (n. d. = not determined).
Figure A.8 Quantitative DNase I footprint titration experiments for polyamides 2, 3, and 4 on the 285 base pair, 5’ end-labeled PCR product of plasmid pCDMF-2: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction.
Figure A.9 Quantitative DNase I footprint titration experiments for polyamides 6, 7, and 8 on the 285 base pair, 5’ end-labeled PCR product of plasmid pCDMF-2: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction.
Figure A.10 Quantitative DNase I footprint titration experiments for polyamides 10, 11, and 12 on the 285 base pair, 5’ end-labeled PCR product of plasmid pCDMF-2: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction.
Figure A.11 Quantitative DNase I footprint titration experiments for polyamides 14, 15, and 16 on the 293 base pair, 5’ end-labeled PCR product of plasmid pCDMF-1: lanes 1-11, 100 nM, 30 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction.
Table A.5 Melting temperatures of polyamides targeted to DNA-sequence 5’-AGAACA-3’ in complex with DNA.\textsuperscript{a}

\begin{table}[h]
\begin{center}
\begin{tabular}{lcc}
\hline
DNA sequence = 5’-TTGC AGAACA GCAA-3’ & \\
Polyamides & \(T_m/^\circ\text{C}\) & \(\Delta T_m/^\circ\text{C}\) \\
\hline
— & 60.1 (±0.2) & — \\
IPA-\begin{tikzpicture}[thick, scale=0.6]
  \draw[fill=black] (0,0) circle (0.1);
  \node at (0,0) {\textbullet};
  \draw[fill=black] (0,-0.5) circle (0.1);
  \node at (0,-0.5) {\textbullet};
  \draw[fill=black] (0,-1) circle (0.1);
  \node at (0,-1) {\textbullet};
\end{tikzpicture}_{\text{IPA}} & 74.4 (±0.2) & 14.3 \\
IPA-\begin{tikzpicture}[thick, scale=0.6]
  \draw[fill=black] (0,0) circle (0.1);
  \node at (0,0) {\textbullet};
  \draw[fill=black] (0,-0.5) circle (0.1);
  \node at (0,-0.5) {\textbullet};
  \draw[fill=black] (0,-1) circle (0.1);
  \node at (0,-1) {\textbullet};
\end{tikzpicture}_{\text{IPA}} & 76.3 (±0.2) & 16.2 \\
IPA-\begin{tikzpicture}[thick, scale=0.6]
  \draw[fill=black] (0,0) circle (0.1);
  \node at (0,0) {\textbullet};
  \draw[fill=black] (0,-0.5) circle (0.1);
  \node at (0,-0.5) {\textbullet};
  \draw[fill=black] (0,-1) circle (0.1);
  \node at (0,-1) {\textbullet};
\end{tikzpicture}_{\text{IPA}} & 64.6 (±0.1) & 4.5 \\
IPA-\begin{tikzpicture}[thick, scale=0.6]
  \draw[fill=black] (0,0) circle (0.1);
  \node at (0,0) {\textbullet};
  \draw[fill=black] (0,-0.5) circle (0.1);
  \node at (0,-0.5) {\textbullet};
  \draw[fill=black] (0,-1) circle (0.1);
  \node at (0,-1) {\textbullet};
\end{tikzpicture}_{\text{IPA}} & 66.9 (±0.1) & 6.8 \\
\hline
\end{tabular}
\end{center}
\end{table}

\textsuperscript{a} All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses. \(\Delta T_m\) values are given as \(T_m(\text{DNA/polyamide}) - T_m(\text{DNA})\).
Figure A.12 A) Chemical structures and ball-and-stick models of matched and mismatched polyamides 20-23, respectively, targeted to 5′-AGAACA-3′. B) Inhibition of DHT-induced PSA, KLK2, FKBP5, and TMPRSS2 expression by 20-23 measured by quantitative real-time RT-PCR.
Appendix B: Aprendica ADMET Report

(Supplemental Information Continued from Chapter 3)
Figure B.1 Polyamides 1 and 5 were subjected to preclinical ADMET testing by contract service at Apredica (Watertown, MA). Shown in Chapter 3 of this thesis (Table 3.2-3.7) are summaries of the ADMET results taken directly from the final report provided by Apredica. The full ADMET report, which includes experimental conditions, is contained on the following pages.
Aprelica Study Number: CIT-001

ADMET Properties of Test Agents

Final Report

Sponsor: California Institute of Technology
Division of Chemistry and Chemical Engineering
1200 E California Blvd; MC 164-30
Pasadena, CA 91125 (USA)

Sponsor’s Representative
Daniel Harki, Ph.D.
Phone: (626) 395-6032
Email: harki@caltech.edu

Test Facilities: Aprelica
313 Pleasant Street
Watertown, MA 02472
Telephone Number: 617.812.1911

Study Director
Jon Gilbert
In Vitro ADMET, Bioanalysis
617.812.1911x122
jon@aprelica.com

Approved: ____________________________

Date: 12/2/08
TABLE OF CONTENTS

1 Objective ...................................................................................................................... 3

1.1 Regulatory Guidelines ............................................................................................... 3

2 Test Methods .................................................................................................................. 4

3 Test Methods .................................................................................................................. 4

3.1 Analytical Methods ...................................................................................................... 4

3.1.1 Method development ............................................................................................. 4

3.1.2 Analysis .................................................................................................................. 4

3.2 In vitro ADME-Tox Experimental Conditions ............................................................ 4

3.2.1 Caco-2 monolayer permeability experimental conditions .................................. 4

3.2.2 Cytotoxicity experimental conditions .................................................................. 4

3.2.3 Fluorescent cytochrome P450 inhibition experimental conditions ....................... 5

3.2.4 Microsomal stability experimental conditions ....................................................... 5

3.2.5 Plasma stability experimental conditions ............................................................... 5

3.2.6 Plasma protein binding experimental conditions ................................................... 5

3.2.7 hERG FastPatch experimental conditions ............................................................ 5

4 Results ............................................................................................................................. 6

4.1 Analytical ...................................................................................................................... 6

4.1.1 Method development ............................................................................................. 6

4.2 In vitro ADME-Tox Summary ...................................................................................... 7

4.2.1 Caco-2 permeability summary .............................................................................. 7

4.2.2 Cytotoxicity summary ........................................................................................... 7

4.2.3 Fluorescent CytoicpIC50 summary ..................................................................... 7

4.2.4 hERG FastPatch summary .................................................................................. 7

4.2.5 Microsomal intrinsic clearance summary ............................................................ 8

4.2.6 Plasma half-life summary ..................................................................................... 8

4.2.7 Plasma protein binding summary ......................................................................... 8

4.3 In vitro ADME-Tox Individual Data ............................................................................ 10

4.3.1 Caco-2 permeability individual data ..................................................................... 10

4.3.2 Cytotoxicity individual data ................................................................................ 10

4.3.3 hERG FastPatch individual data .......................................................................... 11

4.3.4 Fluorescent cytochrome P450 inhibition individual data ....................................... 13

4.3.5 Microsomal intrinsic clearance individual data ..................................................... 16

4.3.6 Plasma half-life individual data ............................................................................ 19

4.3.7 Plasma protein binding individual data ................................................................. 22

5 References .................................................................................................................... 23

6 Storage and Retention of Records ................................................................................. 24

7 Appendices ..................................................................................................................... 25

7.1 Appendix A. Standard Apredica Methods ................................................................ 25

7.2 Appendix B. Sample Spectra and Chromatograms of the Test Agents ..................... 29
1 Objective
The objective of this study is to determine the ADMET properties of test agents.

1.1 Regulatory Guidelines
This study was not conducted under US FDA Good Laboratory Practice Regulations (GLPs). Standard operating procedures of Aprenda were used throughout the study.

2 Test Articles

<table>
<thead>
<tr>
<th>Aprenda ID</th>
<th>Client ID</th>
<th>Physical Form</th>
<th>Submitted FW</th>
<th>Parent MW</th>
<th>Stock solutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>Solid</td>
<td>1399</td>
<td>1399</td>
<td>50 mM DMSO</td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>Solid</td>
<td>1407.26</td>
<td>1178.5</td>
<td>50 mM DMSO</td>
</tr>
</tbody>
</table>

Test agent powders were stored at -20 °C. Stock solutions were stored at -20 °C.
3 Test Methods
Testing was performed at Apredica in Watertown, MA.

3.1 Analytical Methods

3.1.1 Method development
The signal was optimized for each compound by ESI positive or negative ionization mode. A MS2 scan was used to identify the precursor ion and a product ion analysis was used to identify the best fragment for analysis and to optimize the collision energy. An ionization ranking was assigned indicating the compound’s ease of ionization.

3.1.2 Analysis
Samples were analyzed by LC/MS/MS using either an Agilent 6410 mass spectrometer coupled with an Agilent 1200 HPLC and a CTC PAL chilled autosampler, all controlled by MassHunter software (Agilent), or an ABI2000 mass spectrometer coupled with an Agilent 1100 HPLC and a CTC PAL chilled autosampler, all controlled by Analyst software (ABI). After separation on a C18 reverse phase HPLC column (Agilent, Waters, or equivalent) using an acetonitrile-water gradient system, peaks were analyzed by mass spectrometry (MS) using ESI ionization in MRM mode.

3.2 In vitro ADME-Tox Experimental Conditions
Additional protocol details are given in Appendix A.

3.2.1 Caco-2 monolayer permeability experimental conditions

<table>
<thead>
<tr>
<th>Apredica ID</th>
<th>Client ID</th>
<th>Test conc.</th>
<th>Assay time</th>
<th>Direction</th>
<th>Reference compounds</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>10 μM</td>
<td>2 hr</td>
<td>A-&gt;B</td>
<td>warfarin</td>
<td></td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>10 μM</td>
<td>2 hr</td>
<td>B-&gt;A</td>
<td>ranitidine</td>
<td>LC/MS/MS</td>
</tr>
</tbody>
</table>

3.2.2 Cytotoxicity experimental conditions

<table>
<thead>
<tr>
<th>Apredica ID</th>
<th>Client ID</th>
<th>Test conc.</th>
<th>Assay time</th>
<th>Cell lines</th>
<th>Readout</th>
<th>Reference compound</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>0.4</td>
<td>48 hr</td>
<td>HepG2</td>
<td>Neutral</td>
<td>chlorpromazine</td>
<td>fluorescent plate reader</td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>0.07</td>
<td>48 hr</td>
<td>NIH/3T3</td>
<td>red</td>
<td>propranolol</td>
<td></td>
</tr>
<tr>
<td>CIT-001-03</td>
<td></td>
<td>100, 40, 16, 6.4, 2.6, 1.0, 0.4, 0.16, 0.07</td>
<td>48 hr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Page 4 of 34
### 3.2.3 Fluorescent cytochrome P450 inhibition experimental conditions

<table>
<thead>
<tr>
<th>Aprelica ID</th>
<th>Client ID</th>
<th>Test conc.</th>
<th>Cyp assays</th>
<th>Reference compound</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>50, 16.7, 5.6, 0.6</td>
<td>Cyp1A2/CEC, Cyp2C8/DBP, Cyp2C9/DBF, Cyp2C19/DBF</td>
<td>α-naphthoflavone, ketoconazole, sulphaphenazole, tranylcypromine</td>
<td>fluorescent reader</td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>0.02 μM</td>
<td>Cyp3A4/DBF</td>
<td>quinidine, ketoconazole</td>
<td>plate reader</td>
</tr>
</tbody>
</table>

### 3.2.4 Microsomal stability experimental conditions

<table>
<thead>
<tr>
<th>Aprelica ID</th>
<th>Client ID</th>
<th>Test conc.</th>
<th>Microsome source</th>
<th>Protein conc.</th>
<th>Incubation</th>
<th>Ref. comp.</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>5 μM</td>
<td>Human and rat</td>
<td>0.3 mg/mL</td>
<td>37 °C</td>
<td>verapamil</td>
<td>LC/MS/MS</td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>5 μM</td>
<td>Human and rat</td>
<td>30, 40, and 60 min</td>
<td>warfarin</td>
<td>LC/MS/MS</td>
<td></td>
</tr>
</tbody>
</table>

### 3.2.5 Plasma stability experimental conditions

<table>
<thead>
<tr>
<th>Aprelica ID</th>
<th>Client ID</th>
<th>Test conc.</th>
<th>Plasma source</th>
<th>Incubation</th>
<th>Reference compounds</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>10 μM</td>
<td>Human and rat</td>
<td>120 min</td>
<td>propantheline</td>
<td>LC/MS/MS</td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>10 μM</td>
<td>Human and rat</td>
<td>37 °C</td>
<td>propantheline</td>
<td>LC/MS/MS</td>
</tr>
</tbody>
</table>

### 3.2.6 Plasma protein binding experimental conditions

<table>
<thead>
<tr>
<th>Aprelica ID</th>
<th>Client ID</th>
<th>Test conc.</th>
<th>Plasma species</th>
<th>Incubation</th>
<th>Sep. method</th>
<th>Ref. compound</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>10 μM</td>
<td>Human and rat</td>
<td>4 hr</td>
<td>equilibrium</td>
<td>warfarin, atenolol</td>
<td>LC/MS/MS</td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>10 μM</td>
<td>Human and rat</td>
<td>37 °C</td>
<td>dialysis</td>
<td>warfarin, atenolol</td>
<td>LC/MS/MS</td>
</tr>
</tbody>
</table>

### 3.2.7 hERG FastPatch experimental conditions

<table>
<thead>
<tr>
<th>Aprelica ID</th>
<th>Client ID</th>
<th>Test conc.</th>
<th>Medium</th>
<th>Incubation</th>
<th>Ref. comp.</th>
<th>Analytical method</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIT-001-01</td>
<td>DH-V-88</td>
<td>10, 3, 1, 0.3, 0.1</td>
<td>HEPES-aspartate buffer</td>
<td>5 min</td>
<td>E-4031</td>
<td>electrophysiology</td>
</tr>
<tr>
<td>CIT-001-02</td>
<td>DMC2-239</td>
<td>0.03 μM</td>
<td>HEPES-aspartate buffer</td>
<td>ambient temp.</td>
<td>E-4031</td>
<td>electrophysiology</td>
</tr>
</tbody>
</table>
4 Results

4.1 Analytical

4.1.1 Method development

| Client ID | MW  | Polari- | Precursor m/z | Product m/z | Collision energy (V) | Ionization classification*
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DH-V-88</td>
<td>1399</td>
<td>pos</td>
<td>700.2</td>
<td>231</td>
<td>26</td>
<td>2</td>
</tr>
<tr>
<td>DMC2-239</td>
<td>1178.5</td>
<td>pos</td>
<td>590.1</td>
<td>372</td>
<td>20</td>
<td>2</td>
</tr>
</tbody>
</table>

*Ionization classification:
1 = Highly ionizable
2 = Intermediately ionizable
3 = Poorly ionizable

The full scan mass spectrum, the product ion spectrum, and a sample chromatogram are shown in Appendix B.
4.2 In vitro ADME-Tox Summary

4.2.1 Caco-2 permeability summary

<table>
<thead>
<tr>
<th>Client ID</th>
<th>test conc (µM)</th>
<th>Assay duration (hr)</th>
<th>mean A&gt;B (P_{app}^a) (10^{-6} cm s^{-1})</th>
<th>mean A&gt;B (P_{app}^b) (10^{-6} cm s^{-1})</th>
<th>Asymmetry ratio(^b)</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warfarin</td>
<td>50</td>
<td>2</td>
<td>35.4</td>
<td>7.9</td>
<td>0.2</td>
<td>high permeability control</td>
</tr>
<tr>
<td>Ranitidine</td>
<td>50</td>
<td>2</td>
<td>1.4</td>
<td>2.4</td>
<td>1.7</td>
<td>low permeability control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10</td>
<td>2</td>
<td>ND</td>
<td>0.11</td>
<td>UD</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10</td>
<td>2</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Apparent permeability
\(^b\)\(P_{app}(B>A)/P_{app}(A>B)\)
ND = no compound detected in receiver solution

4.2.2 Cytotoxicity summary

<table>
<thead>
<tr>
<th>Client ID</th>
<th>Cell line</th>
<th>IC50 (µM)</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chlorpromazine</td>
<td>HepG2</td>
<td>13</td>
<td>Highly cytotoxic control</td>
</tr>
<tr>
<td>Propranolol</td>
<td>HepG2</td>
<td>80</td>
<td>Low cytotoxic control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>HepG2</td>
<td>&gt;100</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>HepG2</td>
<td>&gt;100</td>
<td></td>
</tr>
</tbody>
</table>

4.2.3 Fluorescent Cyp IC50 summary

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Controls</td>
<td>0.2 α-naphthoflavone</td>
<td>2.3 ketoconazole</td>
<td>1.1 sulphonezole</td>
<td>5.6 tranyltetacyprone</td>
<td>0.05 quinidine</td>
<td>1.26 ketoconazole</td>
<td>1.26 ketoconazole</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>47.6</td>
<td>&gt;50</td>
</tr>
<tr>
<td>DMC2-239</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>&gt;50</td>
<td>37.7</td>
<td>&gt;50</td>
</tr>
</tbody>
</table>

4.2.4 hERG FastPatch summary

<table>
<thead>
<tr>
<th>Client ID</th>
<th>IC50 (µM)</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-4031</td>
<td>99% at 0.5</td>
<td>positive control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>&gt;100</td>
<td>*</td>
</tr>
<tr>
<td>DMC2-239</td>
<td>&gt;100</td>
<td>*</td>
</tr>
</tbody>
</table>

*The solubility limit for this experiment, as determined by vehicle controls, was 17.3 x 10⁵ M (horizontal black line). Based on the data obtained, there may be solubility issues for both test articles at 30 and 100 µM in our physiological saline solution (HB-PS, 0.3%DMSO). Precipitation of DH-V-88 at 100 µM was visible to the naked eye.
### 4.2.5 Microsomal intrinsic clearance summary

<table>
<thead>
<tr>
<th>Client ID</th>
<th>Test conc (µM)</th>
<th>Test species</th>
<th>NADPH-dependent CL&lt;sub&gt;int&lt;/sub&gt; (µl min&lt;sup&gt;-1&lt;/sup&gt; mg&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>NADPH-dependent T&lt;sub&gt;1/2&lt;/sub&gt; (min)</th>
<th>NADPH-free CL&lt;sub&gt;int&lt;/sub&gt; (µl min&lt;sup&gt;-1&lt;/sup&gt; mg&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>NADPH-free T&lt;sub&gt;1/2&lt;/sub&gt; (min)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verapamil</td>
<td>1 Human</td>
<td>411.3</td>
<td>5.6</td>
<td>0.6</td>
<td>&gt;180</td>
<td></td>
<td>metabolized control</td>
</tr>
<tr>
<td>Verapamil</td>
<td>1 Rat</td>
<td>2276</td>
<td>1</td>
<td>0.0</td>
<td>&gt;180</td>
<td></td>
<td>metabolized control</td>
</tr>
<tr>
<td>Warfarin</td>
<td>1 Human</td>
<td>0.0</td>
<td>&gt;180</td>
<td>0.0</td>
<td>&gt;180</td>
<td></td>
<td>non-metabolized control</td>
</tr>
<tr>
<td>Warfarin</td>
<td>1 Rat</td>
<td>0.0</td>
<td>&gt;180</td>
<td>0.0</td>
<td>&gt;180</td>
<td></td>
<td>non-metabolized control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>5 Human</td>
<td>0.0</td>
<td>&gt;180</td>
<td>0.0</td>
<td>&gt;180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH-V-88</td>
<td>5 Rat</td>
<td>0.0</td>
<td>&gt;180</td>
<td>0.0</td>
<td>&gt;180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>5 Human</td>
<td>0.0</td>
<td>&gt;180</td>
<td>0.0</td>
<td>&gt;180</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>5 Rat</td>
<td>0.0</td>
<td>&gt;180</td>
<td>0.0</td>
<td>&gt;180</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Microsomal Intrinsic Clearance

### 4.2.6 Plasma half-life summary

<table>
<thead>
<tr>
<th>Compound</th>
<th>Test conc (µM)</th>
<th>Medium</th>
<th>T1/2 (min)</th>
<th>Fraction remaining, max time (%)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Propantheline</td>
<td>10.0</td>
<td>Human</td>
<td>35.5</td>
<td>5.8%</td>
<td>control</td>
</tr>
<tr>
<td>Propantheline</td>
<td>10.0</td>
<td>Rat</td>
<td>149.0</td>
<td>51.6%</td>
<td>control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10.0</td>
<td>Human</td>
<td>&gt;120</td>
<td>95.6%</td>
<td></td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10.0</td>
<td>Rat</td>
<td>&gt;120</td>
<td>94.0%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10.0</td>
<td>Human</td>
<td>&gt;120</td>
<td>124.5%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10.0</td>
<td>Rat</td>
<td>&gt;120</td>
<td>120.3%</td>
<td></td>
</tr>
</tbody>
</table>

*Half-life

### 4.2.7 Plasma protein binding summary

<table>
<thead>
<tr>
<th>Client ID</th>
<th>Test conc (µM)</th>
<th>Assay duration</th>
<th>Species</th>
<th>Mean free fraction (%)</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warfarin</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>0.73%</td>
<td>high binding control</td>
</tr>
<tr>
<td>Warfarin</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>5.47%</td>
<td>high binding control</td>
</tr>
<tr>
<td>Atenolol</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>76.2%</td>
<td>low binding control</td>
</tr>
<tr>
<td>Atenolol</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>84.7%</td>
<td>low binding control</td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>0.0015%</td>
<td></td>
</tr>
<tr>
<td>DH-V-88</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>0.0016%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10</td>
<td>4 hr</td>
<td>Human</td>
<td>0.0000%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10</td>
<td>4 hr</td>
<td>Rat</td>
<td>0.0040%</td>
<td></td>
</tr>
</tbody>
</table>
4.3 In vitro ADME-Tox Individual Data

4.3.1 Caco-2 permeability individual data

<table>
<thead>
<tr>
<th>Client ID</th>
<th>test conc. (μM)</th>
<th>direction</th>
<th>value</th>
<th>1st</th>
<th>2nd</th>
<th>mean</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH-V-88</td>
<td>10</td>
<td>A→B</td>
<td>dQ/dt&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A→B</td>
<td>C&lt;sub&gt;50&lt;/sub&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.1</td>
<td>1.0</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>B→A</td>
<td>dQ/dt&lt;sup&gt;a&lt;/sup&gt;</td>
<td>8.8E-08</td>
<td>1.9E-08</td>
<td>5.3E-08</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>B→A</td>
<td>C&lt;sub&gt;50&lt;/sub&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1.4</td>
<td>1.4</td>
<td>1.4</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>10</td>
<td>A→B</td>
<td>dQ/dt&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>A→B</td>
<td>C&lt;sub&gt;50&lt;/sub&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.05</td>
<td>0.02</td>
<td>3.5E-02</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>B→A</td>
<td>dQ/dt&lt;sup&gt;a&lt;/sup&gt;</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>B→A</td>
<td>C&lt;sub&gt;50&lt;/sub&gt;&lt;sup&gt;b&lt;/sup&gt;</td>
<td>0.1</td>
<td>0.1</td>
<td>8.4E-02</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>rate of test agent permeation, area units/sec
<sup>b</sup>initial concentration (area units/cm²)

4.3.2 Cytotoxicity individual data

![Growth Inhibition of HepG2 Cells by DH-V-88](image)

![Growth Inhibition of HepG2 Cells by DMC2-239](image)
### 4.3.3 hERG FastPatch individual data

<table>
<thead>
<tr>
<th>Test Article ID</th>
<th>IC50 (pM)</th>
<th>Conc. (pM)</th>
<th>Mean % hERG Inhibition</th>
<th>Standard Deviation</th>
<th>Standard Error</th>
<th>n</th>
<th>Individual Data Points (% Inhibition)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DHM-88</td>
<td>&gt;100</td>
<td>0.03</td>
<td>4.6</td>
<td>3.8</td>
<td>2.7</td>
<td>2</td>
<td>7.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>5.6</td>
<td>3.1</td>
<td>2.2</td>
<td>2</td>
<td>7.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3</td>
<td>1.9</td>
<td>4.0</td>
<td>2.9</td>
<td>2</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>9.3</td>
<td>3.1</td>
<td>2.2</td>
<td>2</td>
<td>11.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>2.0</td>
<td>4.8</td>
<td>3.4</td>
<td>2</td>
<td>5.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>3.1</td>
<td>2.3</td>
<td>1.6</td>
<td>2</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>1.2</td>
<td>0.7</td>
<td>0.5</td>
<td>2</td>
<td>0.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>0.9</td>
<td>1.4</td>
<td>1.0</td>
<td>2</td>
<td>1.9</td>
</tr>
<tr>
<td>DMG2b</td>
<td>&gt;100</td>
<td>0.03</td>
<td>-1.2</td>
<td>0.1</td>
<td>0.1</td>
<td>2</td>
<td>-1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>2.1</td>
<td>0.4</td>
<td>0.3</td>
<td>2</td>
<td>-1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3</td>
<td>-4.1</td>
<td>1.2</td>
<td>0.9</td>
<td>2</td>
<td>-4.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>-3.7</td>
<td>3.3</td>
<td>2.4</td>
<td>2</td>
<td>-6.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>-0.7</td>
<td>1.6</td>
<td>1.1</td>
<td>2</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>-2.2</td>
<td>2.8</td>
<td>2.0</td>
<td>2</td>
<td>-0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>5.5</td>
<td>1.4</td>
<td>1.0</td>
<td>2</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>9.0</td>
<td>2.1</td>
<td>1.5</td>
<td>2</td>
<td>7.5</td>
</tr>
</tbody>
</table>
TurboSol Table 1

<table>
<thead>
<tr>
<th>Compound</th>
<th>Figure Legend</th>
<th>Conc. (µM)</th>
<th>n</th>
<th>Average LSU (x1000)</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle</td>
<td>A</td>
<td>0</td>
<td>11</td>
<td>6.1</td>
<td>6.2</td>
</tr>
<tr>
<td>88</td>
<td>B</td>
<td>0.03</td>
<td>3</td>
<td>4.1</td>
<td>0.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>3</td>
<td>4.3</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3</td>
<td>3</td>
<td>4.0</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>6.0</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>5.8</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>3</td>
<td>14.1</td>
<td>4.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>3</td>
<td>131.9</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>3</td>
<td>495.9</td>
<td>7.1</td>
</tr>
<tr>
<td>219</td>
<td>C</td>
<td>0.03</td>
<td>3</td>
<td>3.9</td>
<td>1.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1</td>
<td>3</td>
<td>5.0</td>
<td>1.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.3</td>
<td>3</td>
<td>3.5</td>
<td>0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1</td>
<td>3</td>
<td>3.9</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>5.6</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>3</td>
<td>11.6</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>30</td>
<td>3</td>
<td>70.3</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>100</td>
<td>3</td>
<td>335.9</td>
<td>46.3</td>
</tr>
<tr>
<td>Transmittance Standard</td>
<td>TS</td>
<td>60%</td>
<td>3</td>
<td>495.3</td>
<td>7.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>80%</td>
<td>3</td>
<td>259.0</td>
<td>9.3</td>
</tr>
</tbody>
</table>

TS: Transmittance standard. % indicates percent light transmitted.
LSU: Light scatter unit
4.3.4 Fluorescent cytochrome P450 inhibition individual data

**Graphs:**
- Inhibition of 1A2/EC by DH-V-88
- Inhibition of 1A2/EC by DMC2-239
- Inhibition of 2C8/DBF by DH-V-88
- Inhibition of 2C8/DBF by DMC2-239

**Tables:**

<table>
<thead>
<tr>
<th>Test Agent: DH-V-88</th>
<th>Assay: 1A2/EC</th>
<th>IC50</th>
<th>min</th>
<th>% inhibition</th>
<th>slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&gt;50</td>
<td>0%</td>
<td>max 100%</td>
<td></td>
</tr>
<tr>
<td>[TA]</td>
<td>%um1</td>
<td>%um2</td>
<td>mean</td>
<td>cal</td>
<td></td>
</tr>
<tr>
<td>50.0</td>
<td>44%</td>
<td>37%</td>
<td>41%</td>
<td>38%</td>
<td></td>
</tr>
<tr>
<td>16.67</td>
<td>18%</td>
<td>21%</td>
<td>20%</td>
<td>21%</td>
<td></td>
</tr>
<tr>
<td>5.56</td>
<td>-8%</td>
<td>10%</td>
<td>2%</td>
<td>13%</td>
<td></td>
</tr>
<tr>
<td>1.852</td>
<td>6%</td>
<td>12%</td>
<td>9%</td>
<td>10%</td>
<td></td>
</tr>
<tr>
<td>0.617</td>
<td>8%</td>
<td>6%</td>
<td>7%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>0.2058</td>
<td>10%</td>
<td>18%</td>
<td>14%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>0.0686</td>
<td>8%</td>
<td>-1%</td>
<td>3%</td>
<td>8%</td>
<td></td>
</tr>
<tr>
<td>0.02286</td>
<td>17%</td>
<td>19%</td>
<td>18%</td>
<td>8%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Agent: DMC2-239</th>
<th>Assay: 1A2/EC</th>
<th>IC50</th>
<th>min</th>
<th>% inhibition</th>
<th>slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&gt;50</td>
<td>0%</td>
<td>max 100%</td>
<td></td>
</tr>
<tr>
<td>[TA]</td>
<td>%um1</td>
<td>%um2</td>
<td>mean</td>
<td>cal</td>
<td></td>
</tr>
<tr>
<td>50.0</td>
<td>32%</td>
<td>30%</td>
<td>31%</td>
<td>30%</td>
<td></td>
</tr>
<tr>
<td>16.67</td>
<td>29%</td>
<td>18%</td>
<td>23%</td>
<td>19%</td>
<td></td>
</tr>
<tr>
<td>5.56</td>
<td>2%</td>
<td>-5%</td>
<td>-2%</td>
<td>14%</td>
<td></td>
</tr>
<tr>
<td>1.852</td>
<td>11%</td>
<td>7%</td>
<td>9%</td>
<td>12%</td>
<td></td>
</tr>
<tr>
<td>0.617</td>
<td>16%</td>
<td>10%</td>
<td>13%</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>0.2058</td>
<td>14%</td>
<td>12%</td>
<td>13%</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>0.0686</td>
<td>12%</td>
<td>8%</td>
<td>10%</td>
<td>11%</td>
<td></td>
</tr>
<tr>
<td>0.02286</td>
<td>20%</td>
<td>22%</td>
<td>21%</td>
<td>11%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Agent: DH-V-88</th>
<th>Assay: 2C8/DBF</th>
<th>IC50</th>
<th>min</th>
<th>% inhibition</th>
<th>slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&gt;50</td>
<td>0%</td>
<td>max 100%</td>
<td></td>
</tr>
<tr>
<td>[TA]</td>
<td>%um1</td>
<td>%um2</td>
<td>mean</td>
<td>cal</td>
<td></td>
</tr>
<tr>
<td>50.0</td>
<td>14%</td>
<td>24%</td>
<td>19%</td>
<td>17%</td>
<td></td>
</tr>
<tr>
<td>16.67</td>
<td>-3%</td>
<td>3%</td>
<td>0%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>5.56</td>
<td>-4%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>1.852</td>
<td>-12%</td>
<td>2%</td>
<td>-5%</td>
<td>-2%</td>
<td></td>
</tr>
<tr>
<td>0.617</td>
<td>-2%</td>
<td>-14%</td>
<td>-8%</td>
<td>-3%</td>
<td></td>
</tr>
<tr>
<td>0.2058</td>
<td>-5%</td>
<td>2%</td>
<td>-1%</td>
<td>-3%</td>
<td></td>
</tr>
<tr>
<td>0.0686</td>
<td>-3%</td>
<td>2%</td>
<td>-1%</td>
<td>-3%</td>
<td></td>
</tr>
<tr>
<td>0.02286</td>
<td>0%</td>
<td>5%</td>
<td>3%</td>
<td>-3%</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Test Agent: DMC2-239</th>
<th>Assay: 2C8/DBF</th>
<th>IC50</th>
<th>min</th>
<th>% inhibition</th>
<th>slope</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>&gt;50</td>
<td>0%</td>
<td>max 100%</td>
<td></td>
</tr>
<tr>
<td>[TA]</td>
<td>%um1</td>
<td>%um2</td>
<td>mean</td>
<td>cal</td>
<td></td>
</tr>
<tr>
<td>50.0</td>
<td>27%</td>
<td>14%</td>
<td>20%</td>
<td>21%</td>
<td></td>
</tr>
<tr>
<td>16.67</td>
<td>8%</td>
<td>15%</td>
<td>11%</td>
<td>7%</td>
<td></td>
</tr>
<tr>
<td>5.56</td>
<td>-7%</td>
<td>-2%</td>
<td>-4%</td>
<td>2%</td>
<td></td>
</tr>
<tr>
<td>1.852</td>
<td>-4%</td>
<td>-13%</td>
<td>-9%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>0.617</td>
<td>11%</td>
<td>0%</td>
<td>6%</td>
<td>-1%</td>
<td></td>
</tr>
<tr>
<td>0.2058</td>
<td>8%</td>
<td>3%</td>
<td>5%</td>
<td>-1%</td>
<td></td>
</tr>
<tr>
<td>0.0686</td>
<td>-6%</td>
<td>-5%</td>
<td>-5%</td>
<td>-1%</td>
<td></td>
</tr>
<tr>
<td>0.02286</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
<td>-1%</td>
<td></td>
</tr>
</tbody>
</table>
4.3.5 Microsomal intrinsic clearance individual data
### Rat Microsome Stability of DH-V-88

<table>
<thead>
<tr>
<th>Time</th>
<th>CLint</th>
<th>% Remaining Comp.</th>
<th>T1/2</th>
<th>&gt;180</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>%RE1</td>
<td>%RE2</td>
<td>mean</td>
<td>calc</td>
</tr>
<tr>
<td>10</td>
<td>99%</td>
<td>103%</td>
<td>101%</td>
<td>100%</td>
</tr>
<tr>
<td>20</td>
<td>95%</td>
<td>100%</td>
<td>97%</td>
<td>100%</td>
</tr>
<tr>
<td>40</td>
<td>96%</td>
<td>104%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>60</td>
<td>104%</td>
<td>106%</td>
<td>105%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Rat Microsome Stability of DH-V-88, NADPH-free

<table>
<thead>
<tr>
<th>Time</th>
<th>CLint</th>
<th>% Remaining Comp.</th>
<th>T1/2</th>
<th>&gt;180</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>%RE1</td>
<td>%RE2</td>
<td>mean</td>
<td>calc</td>
</tr>
<tr>
<td>10</td>
<td>97%</td>
<td>100%</td>
<td>97%</td>
<td>100%</td>
</tr>
<tr>
<td>20</td>
<td>97%</td>
<td>105%</td>
<td>101%</td>
<td>100%</td>
</tr>
<tr>
<td>40</td>
<td>110%</td>
<td>101%</td>
<td>106%</td>
<td>100%</td>
</tr>
<tr>
<td>60</td>
<td>106%</td>
<td>94%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Human Microsome Stability of DMC2-239

<table>
<thead>
<tr>
<th>Time</th>
<th>CLint</th>
<th>% Remaining Comp.</th>
<th>T1/2</th>
<th>&gt;180</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>%RE1</td>
<td>%RE2</td>
<td>mean</td>
<td>calc</td>
</tr>
<tr>
<td>10</td>
<td>85%</td>
<td>91%</td>
<td>88%</td>
<td>100%</td>
</tr>
<tr>
<td>20</td>
<td>97%</td>
<td>100%</td>
<td>101%</td>
<td>100%</td>
</tr>
<tr>
<td>40</td>
<td>110%</td>
<td>101%</td>
<td>106%</td>
<td>100%</td>
</tr>
<tr>
<td>60</td>
<td>106%</td>
<td>94%</td>
<td>100%</td>
<td>100%</td>
</tr>
</tbody>
</table>

### Human Microsome Stability of DMC2-239, NADPH-free

<table>
<thead>
<tr>
<th>Time</th>
<th>CLint</th>
<th>% Remaining Comp.</th>
<th>T1/2</th>
<th>&gt;180</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>%RE1</td>
<td>%RE2</td>
<td>mean</td>
<td>calc</td>
</tr>
<tr>
<td>10</td>
<td>87%</td>
<td>90%</td>
<td>89%</td>
<td>100%</td>
</tr>
<tr>
<td>20</td>
<td>97%</td>
<td>103%</td>
<td>97%</td>
<td>100%</td>
</tr>
<tr>
<td>40</td>
<td>100%</td>
<td>116%</td>
<td>103%</td>
<td>100%</td>
</tr>
<tr>
<td>60</td>
<td>119%</td>
<td>122%</td>
<td>121%</td>
<td>100%</td>
</tr>
</tbody>
</table>
Rat Microsome Stability of DMC2-239

% remaining comp.

incubation time (min)

0 20 40 60

120%

100%

80%

60%

40%

20%

0%

Rat Microsome Stability of DMC2-239, NADPH-free

% remaining comp.

incubation time (min)

0 20 40 60

140%

120%

100%

80%

60%

40%

20%

0%

Test Agent: DMC2-239  [TA] (uM): 5.0
Species: Rat

<table>
<thead>
<tr>
<th></th>
<th>W NADPH</th>
<th></th>
<th>W/O NADPH</th>
</tr>
</thead>
<tbody>
<tr>
<td>time</td>
<td>CLint 0</td>
<td>T1/2</td>
<td>&gt;1h0</td>
</tr>
<tr>
<td>0.25</td>
<td>100%</td>
<td>95%</td>
<td>98% 100%</td>
</tr>
<tr>
<td>10</td>
<td>100%</td>
<td>98%</td>
<td>99% 100%</td>
</tr>
<tr>
<td>20</td>
<td>101%</td>
<td>100%</td>
<td>101% 100%</td>
</tr>
<tr>
<td>40</td>
<td>98%</td>
<td>97%</td>
<td>98% 100%</td>
</tr>
<tr>
<td>60</td>
<td>107%</td>
<td>102%</td>
<td>105% 100%</td>
</tr>
</tbody>
</table>
4.3.6 Plasma half-life individual data

Human Plasma Stability of DH-V-88

<table>
<thead>
<tr>
<th>Time (min)</th>
<th>%RE1</th>
<th>%RE2</th>
<th>Average</th>
<th>Calc</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>102%</td>
<td>102%</td>
<td>102%</td>
<td>100%</td>
</tr>
<tr>
<td>15</td>
<td>96%</td>
<td>97%</td>
<td>96%</td>
<td>100%</td>
</tr>
<tr>
<td>30</td>
<td>97%</td>
<td>99%</td>
<td>98%</td>
<td>99%</td>
</tr>
<tr>
<td>60</td>
<td>100%</td>
<td>101%</td>
<td>101%</td>
<td>98%</td>
</tr>
<tr>
<td>120</td>
<td>94%</td>
<td>98%</td>
<td>96%</td>
<td>96%</td>
</tr>
</tbody>
</table>
Rat Plasma Stability of DH-V-88

<table>
<thead>
<tr>
<th>time (min)</th>
<th>%RE1</th>
<th>%RE2</th>
<th>avg</th>
<th>calc</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>98%</td>
<td>99%</td>
<td>98%</td>
<td>100%</td>
</tr>
<tr>
<td>15</td>
<td>103%</td>
<td>101%</td>
<td>102%</td>
<td>99%</td>
</tr>
<tr>
<td>30</td>
<td>99%</td>
<td>96%</td>
<td>97%</td>
<td>98%</td>
</tr>
<tr>
<td>60</td>
<td>96%</td>
<td>96%</td>
<td>96%</td>
<td>97%</td>
</tr>
<tr>
<td>120</td>
<td>97%</td>
<td>91%</td>
<td>94%</td>
<td>94%</td>
</tr>
</tbody>
</table>

TA: DH-V-88  Test conc (µM): 10
T1/2 (min) >120  %rem, last time: 94.0%
Test medium: Rat Plasma
Human Plasma Stability of DMC2-239

<table>
<thead>
<tr>
<th>Test agent (µM)</th>
<th>10</th>
</tr>
</thead>
</table>

**TA:** DMC2-239  
**Test medium:** Human Plasma  
**T1/2 (min):** >120  
**%rem, last time:** 124.5%

<table>
<thead>
<tr>
<th>time (min)</th>
<th>%RE1</th>
<th>%RE2</th>
<th>avg</th>
<th>calc</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>76%</td>
<td>73%</td>
<td>74%</td>
<td>100%</td>
</tr>
<tr>
<td>15</td>
<td>74%</td>
<td>72%</td>
<td>73%</td>
<td>100%</td>
</tr>
<tr>
<td>30</td>
<td>93%</td>
<td>86%</td>
<td>89%</td>
<td>100%</td>
</tr>
<tr>
<td>60</td>
<td>114%</td>
<td>116%</td>
<td>115%</td>
<td>100%</td>
</tr>
<tr>
<td>120</td>
<td>129%</td>
<td>120%</td>
<td>124%</td>
<td>100%</td>
</tr>
</tbody>
</table>
4.3.7 Plasma protein binding individual data

<table>
<thead>
<tr>
<th>Client ID</th>
<th>test species</th>
<th>test conc (μM)</th>
<th>1st</th>
<th>2nd</th>
<th>mean</th>
<th>comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH-V-88</td>
<td>Human</td>
<td>10</td>
<td>0.0000%</td>
<td>0.0030%</td>
<td>0.0015%</td>
<td></td>
</tr>
<tr>
<td>DH-V-88</td>
<td>Rat</td>
<td>10</td>
<td>0.00114%</td>
<td>0.0020%</td>
<td>0.0016%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>Human</td>
<td>10</td>
<td>0.00000%</td>
<td>0.00000%</td>
<td>0.0000%</td>
<td></td>
</tr>
<tr>
<td>DMC2-239</td>
<td>Rat</td>
<td>10</td>
<td>0.00000%</td>
<td>0.0079%</td>
<td>0.0040%</td>
<td></td>
</tr>
</tbody>
</table>

*peak area(analyte) / peak area (internal standard)
5 References


6 Storage and Retention of Records

All documents generated in this study (raw data, the study plan, a copy of this report, etc.) will be stored for three years from the date of this document. Only authorized Apredica employees will have access to the archives.

The original final report will be provided to the sponsor and will be kept by the sponsor under its sole responsibility.
7 Appendices

7.1 Appendix A. Standard Aredica Methods

Caco-2 monolayer permeability
Caco-2 cells grown in tissue culture flasks are trypsinized, suspended in medium, and the suspensions were applied to wells of a collagen-coated BioCoat Cell Environment in 24-well format (BD Biosciences) at 24,500 cells per well. The cells are allowed to grow and differentiate for three weeks, feeding at 2-day intervals.

For Apical to Basolateral (A→B) permeability, the test agent is added to the apical (A) side and amount of permeation is determined on the basolateral (B) side; for Basolateral to Apical (B→A) permeability, the test agent is added to the B side and the amount of permeation is determine on the A side. The A-side buffer contains 100 μM Lucifer yellow dye, in Transport Buffer (1.98 g/L glucose in 10 mM HEPES, 1x Hank’s Balanced Salt Solution) pH 6.5, and the B-side buffer is Transport Buffer, pH 7.4. CaCo-2 cells are incubated with these buffers for 2 h., and the receiver side buffer is removed for analysis by LC/MS/MS.

To verify the CaCo-2 cell monolayers are properly formed, aliquots of the cell buffers are analyzed by fluorescence to determine the transport of the impermeable dye Lucifer Yellow.

Data are expressed as permeability ($P_{\text{app}}$): 
$$ P_{\text{app}} = \frac{\frac{dQ}{dt}}{C_0 A} $$
where $\frac{dQ}{dt}$ is the rate of permeation, $C_0$ is the initial concentration of test agent, and $A$ is the area of the monolayer.

In bidirectional permeability studies, the asymmetry index (AI) is also calculated:
$$ AI = \frac{P_{\text{app}}(B \rightarrow A)}{P_{\text{app}}(A \rightarrow B)} $$

An AI > 1 indicated a potential substrate for PGP or other active transporters.

Cytotoxicity
HepG2 human hepatocellular carcinoma cells (originally obtained from ATCC, Manassas, VA) are seeded in 96-well plates at 10,000 cells per well, and grown for 24 hr in Eagle’s Modified Essential Medium supplemented with 2 mM glutamine, nonessential amino acids, 2 mM pyruvate, 10% fetal bovine serum, 100 U/mL penicillin, and 100 mg/L streptomycin at 37°C in 5% CO₂. NIH/3T3 mouse fibroblasts (originally obtained from ATCC) are seeded in 96-well plates at 10,000 cells per well, and grown for 24 hr in Dulbecco’s Modified Essential Medium supplemented with 10% bovine calf serum, 100 U/mL penicillin, and 100 mg/L streptomycin at 37°C in 5% CO₂. HaCaT human keratinocytes are seeded in 96-well plates at 10,000 cells per well, and grown for 24 hr in Dulbecco’s Modified Essential Medium supplemented with 10% fetal bovine serum, 100 U/mL penicillin, and 100 mg/L streptomycin at 37°C in 5% CO₂.

Test article is aseptically diluted in DMSO to 200x the highest concentration, then 100-fold in the growth medium, and serial dilutions are made in 1% DMSO in growth medium. At the start of the assay, the growth medium is removed from the plates and replaced with fresh medium, and an equal volume from each test agent dilution is added. Cells are incubated with test article for 48 h, and the wells are examined microscopically to look for abnormalities.
For Neutral red staining, medium is removed, the cells are washed with PBS, and fresh medium containing 25 μg/mL neutral red (Sigma) is added. After four hours incubation, the cells are washed with PBS, and the cellular dye is solubilized with 1% acetic acid in 50% ethanol. Cellular neutral red is measured by its absorbance at 540 nm. Cytotoxicity is assessed by determining the IC₅₀ (the concentration that causes 50% reduction in uptake of neutral red after 48 hrs exposure to compound).

For MTT staining, 20 μL 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide MTT, 5 mg/mL in PBS, Sigma), is added to each well. After two hours incubation, the medium is removed, and the cellular dye is solubilized with DMSO. Cellular-converted MTT is measured by its absorbance at 540 nm. Cytotoxicity is assessed by determining the IC₅₀ (the concentration that causes 50% reduction in uptake of MTT after 48 hrs exposure to compound).

**Fluorescent cytochrome P450 IC₅₀ determination**

Cytochrome P450 inhibition is measured using fluorogenic substrates. Test agents and substrates are dissolved in acetonitrile for this assay, as DMSO significantly inhibits some cytochrome P450s. Assays were performed at 37 °C using commercially available recombinant human cytochrome P450 expressed in insect cells. Enzyme concentrations and reactions times are optimized for each batch of enzyme to ensure a linear production of product over the course of the reaction. Percent remaining activity is calculated by comparing product formation of wells treated with test agent against wells treated with vehicle, after subtraction of background fluorescence. Percent inhibition is 100% - percent remaining activity. IC₅₀ is calculated using a four-point logistic curve model. The individual reaction conditions are summarized in the following Table.

<table>
<thead>
<tr>
<th>Cytochrome</th>
<th>Substrate</th>
<th>Assay Buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1A2</td>
<td>5 μM CEC</td>
<td>Buffer B</td>
</tr>
<tr>
<td>2A6</td>
<td>3 μM coumarin</td>
<td>Buffer D</td>
</tr>
<tr>
<td>2B6</td>
<td>15 μM MFC</td>
<td>Buffer B</td>
</tr>
<tr>
<td>2C8</td>
<td>1 μM DBF</td>
<td>Buffer E</td>
</tr>
<tr>
<td>2C9</td>
<td>1 μM DBF</td>
<td>Buffer A</td>
</tr>
<tr>
<td>2C19</td>
<td>2 μM DBF</td>
<td>Buffer B</td>
</tr>
<tr>
<td>2D6</td>
<td>1.5 μM AMMC</td>
<td>Buffer C</td>
</tr>
<tr>
<td>2E1</td>
<td>100 μM MFC</td>
<td>Buffer B</td>
</tr>
<tr>
<td>3A4</td>
<td>50 μM BFC</td>
<td>Buffer B</td>
</tr>
<tr>
<td>3A4</td>
<td>1 μM DBF</td>
<td>Buffer B</td>
</tr>
<tr>
<td>3A5</td>
<td>50 μM BFC</td>
<td>Buffer B</td>
</tr>
</tbody>
</table>

| Aromatase  | 0.5 μM DBF    | Buffer B     |

Buffer A: 1.3 mM NADP, 3.3 mM glucose-6-phosphate, 0.4 U/mL glucose-6-phosphate dehydrogenase, 25 mM potassium phosphate, 3.3 mM MgCl₂, pH 7.4.
Buffer B: 1.3 mM NADP, 3.3 mM glucose-6-phosphate, 0.4 U/mL glucose-6-phosphate dehydrogenase, 100 mM potassium phosphate, 3.3 mM MgCl₂, pH 7.4.
Buffer C: 8.2 μM NADP, 0.41 mM glucose-6-phosphate, 0.4 U/mL glucose-6-phosphate dehydrogenase, 100 mM potassium phosphate, 0.41 mM MgCl₂, pH 7.4.
Buffer D: 0.066 mM NADP, 3.3 mM glucose-6-phosphate, 0.4 U/mL glucose-6-phosphate dehydrogenase, 100 mM Tris hydrochloride, 3.3 mM MgCl₂, pH 7.5.
Buffer E: 1.3 mM NADP, 3.3 mM glucose-6-phosphate, 0.4 U/mL glucose-6-phosphate dehydrogenase, 50 mM potassium phosphate, 3.3 mM MgCl₂, pH 7.4.

Abbreviations: AMMC: 3-[2-(N,N-Diethyl-N-methylammonium)ethyl]-7-methoxy-4-methylcoumarin; BFC: 7-Benzoyl-4-(trifluoromethyl)coumarin; CEC: 3-Cyano-7-ethoxycoumarin; DBF: dibenzylfluorescein; MFC: 7-Methoxy-4-(trifluoromethyl)coumarin

**hERG FastPatch**

**Cell culture.** HEK293 cells were stably transfected with hERG cDNA. Stable
transfectants have been selected by coexpression with the G418-resistance gene incorporated into the expression plasmid. Selection pressure was maintained by including G418 in the culture medium. Cells were cultured in Dulbecco’s Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F-12) supplemented with 10% fetal bovine serum, 100 U/mL penicillin G sodium, 100 μg/mL streptomycin sulfate and 500 μg/mL G418. Before testing, cells in culture dishes were washed twice with Hank’s Balanced Salt Solution, treated with trypsin and re-suspended in the culture media (1-1.5 x 10^5 cells in 20 mL). Cells in suspension were allowed to recover for 1-3 hours in a tissue culture incubator set at 37°C in a humidified 95% air, 5% CO2 atmosphere. Immediately before use in the PatchXpress® system, the cells were washed in HB-PS to remove the culture medium and re-suspended in 150 μL of HB-PS.

**Test Article Treatment Groups.** Two concentrations were applied at five (5) minute intervals via disposable polyethylene micropipette tips to cells expressing hERG (n ≥ 2, where n = the number of cells/concentration). Each solution exchange, performed in quadruplicate, consisted of aspiration and replacement of 45 μL of the total 50 μL volume of the extracellular well of the Sealchip16. Duration of exposure to each test article concentration was five (5) minutes.

**Test Article Application Schedule.**

<table>
<thead>
<tr>
<th>Solution</th>
<th>Procedure</th>
<th>Exposure time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vehicle control</td>
<td>four 45 μL exchanges</td>
<td>10 min</td>
</tr>
<tr>
<td>Test article concentration 1</td>
<td>four 45 μL exchanges</td>
<td>5 min</td>
</tr>
<tr>
<td>Test article concentration 2</td>
<td>four 45 μL exchanges</td>
<td>5 min</td>
</tr>
</tbody>
</table>

**Positive Control Treatment Group.** Vehicle was applied to cells expressing hERG (n ≥ 2, where n = the number cells), for a 10-minute exposure interval. Each solution exchange, performed in quadruplicate, consisted of aspiration and replacement of 45 μL of the total 50 μL volume of the extracellular well of the Sealchip16. Application of vehicle, the positive control was applied in the same manner, to verify sensitivity to hERG blockade.

**Automated Patch Clamp Electrophysiological Procedures.** Intracellular solution for whole-cell recordings consisted of (composition in mM): potassium aspartate, 130; MgCl2, 5; EGTA, 5; ATP, 4; HEPES, 10; pH adjusted to 7.2 with KOH. This solution was prepared in batches, aliquoted, stored frozen, and a fresh aliquot thawed each day. In preparation for a recording session, intracellular solution was loaded into the intracellular compartments of the Sealchip16 planar electrode. Cell suspension was pipetted into the extracellular compartments of the Sealchip16 planar electrode. After establishment of a whole-cell configuration, membrane currents were recorded using dual-channel patch clamp amplifiers in the PatchXpress® system. Before digitization, the current records were low-pass filtered at one-fifth of the sampling frequency.

**Voltage-Clamp Procedures.** Onset and block of hERG current was measured using a stimulus voltage pattern (Figure 1, lower panel) consisting of a 500 ms prepulse to −40 mV (leakage subtraction), a 2-second activating pulse to +40 mV, followed by a 2-second test pulse to −40 mV. The pulse pattern was repeated continuously at 10 s intervals, from a holding potential of −80 mV. Peak tail current (Figure 1, upper panel) was measured during the −40 mV test pulse. Leakage current was calculated from the current amplitude evoked by the prepulse and subtracted from the total membrane current record.

**Microsom al intrinsic clearance**

The test agent is incubated in duplicate with microsomes at 37 °C. The reaction contains microsomal protein in 100 mM potassium phosphate, 2 mM NADPH, 3 mM MgCl2, pH 7.4. A control is run for each test agent omitting NADPH to detect NADPH-free degradation. The indicated times, an aliquot is removed from each experimental and
control reaction and mixed with an equal volume of ice-cold Stop Solution (0.3% acetic acid in acetonitrile containing haloperidol, diclofenac, or other internal standard). Stopped reactions are incubated at least ten minutes at -20 °C, and an additional volume of water is added. The samples are centrifuged to remove precipitated protein, and the supernatants are analyzed by LC/MS/MS to quantify the remaining parent. Data are converted to % remaining by dividing by the time zero concentration value. Data are fit to a first-order decay model to determine half-life. Intrinsic clearance is calculated from the half-life and the protein concentrations: 

\[ CL_{int} = \frac{\ln(2)}{T/2} \]

[microsomal protein]}. 

**Plasma protein binding**

Test agent is added to plasma. This mixture is dialyzed in a RED Device (Pierce) per the manufacturers’ instructions along against PBS and incubated in a rocker. After the end of the incubation, aliquots from both plasma and PBS sides are collected, an equal amount of PBS is added to the plasma sample, and an equal volume of plasma is added to the PBS sample. Methanol (three volumes) with haloperidol IS are added to precipitate the proteins and release the agents. After centrifugation, the supernatant was transferred to a new plate and analyzed by LC/MS/MS.

**Plasma half-life**

The test agent is incubated in duplicate with plasma at 37 °C. The reaction contains plasma and 2% DMSO. At the indicated times, an aliquot is removed from each experimental reaction and mixed with three volumes of ice-cold Stop Solution (methanol containing propranolol, diclofenac, or other internal standard). Stopped reactions are incubated at least ten minutes at -20 °C. The samples are centrifuged to remove precipitated protein, and the supernatants are analyzed by LC/MS/MS to quantitate the remaining parent. Data are converted to % remaining by dividing by the time zero concentration value. Data are fit to a first-order decay model to determine half-life.
Appendix C: Programmable Oligomers for Minor Groove DNA Recognition

The text of this chapter was taken in part from a manuscript coauthored with Raymond M. Doss, Michael M. Marques, Shane Foister, and Peter B. Dervan* (Caltech)

Abstract

The four Watson-Crick base pairs of DNA can be distinguished in the minor groove by pairing side-by-side three five-membered aromatic carboxamides, imidazole (Im), pyrrole (Py), and hydroxypyrrole (Hp), four different ways. On the basis of the paradigm of unsymmetrical paired edges of aromatic rings for minor groove recognition, a second generation set of heterocycle pairs, imidazopyridine/pyrrole (Ip/Py) and hydroxybenzimidazole/pyrrole (Hz/Py), revealed that recognition elements not based on analogues of distamycin could be realized. A new set of end-cap heterocycle dimers, oxazole-hydroxybenzimidazole (No-Hz) and chlorothiophene-hydroxybenzimidazole (Ct-Hz), paired with Py-Py are shown to bind contiguous base pairs of DNA in the minor groove, specifically 5’-GT-3’ and 5’-TT-3’, with high affinity and selectivity. Utilizing this technology, we have developed a new class of oligomers for sequence-specific DNA minor groove recognition no longer based on the N-methyl pyrrole carboxamides of distamycin.
C.1 Introduction

Aberrant gene expression is the cause of many diseases, and the ability to reprogram transcriptional pathways using cell-permeable small molecules may, one day, have an impact on human medicine.\textsuperscript{1} DNA-binding polyamides, which are based on the architecture of the natural products netropsin and distamycin A,\textsuperscript{2,3a,b} are capable of distinguishing all four Watson-Crick base pairs in the DNA minor groove and have been the subject of intense study along with many other classes of minor groove binders.\textsuperscript{3c-f,4} Sequence-specific recognition of the minor groove of DNA by polyamides arises from the pairing of three different antiparallel five-membered heterocyclic amino acids, pyrrole (Py), imidazole (Im), and hydroxypyrrole (Hp).\textsuperscript{4-5} The direct readout, or information face, on the inside of the crescent-shaped polyamide may be programmed by the incremental change of atoms on the corners of the ring pairs presented to the DNA minor groove floor. Stabilizing and, importantly, destabilizing interactions with the different edges of the four Watson-Crick bases are modulated by shape complementarity and specific hydrogen bonds.\textsuperscript{4-6,7l} For example, the Im/Py pair distinguishes G•C from C•G, T•A, and A•T. Im presents a lone pair of electrons to the DNA minor groove and can accept a hydrogen bond from the exocyclic amine of guanine.\textsuperscript{5} Additionally, the Hp/Py pair distinguishes T•A from A•T, G•C, and C•G.\textsuperscript{4-6} Hp projects an exocyclic OH group toward the minor groove floor that is sterically accommodated in the cleft of the T•A base pair, preferring to lie over T, not A.\textsuperscript{5} These pairing rules have proven useful for programmed recognition of a broad repertoire of DNA sequences; however, the hydroxypyrrole ring system has proven to be unstable over time and in the presence of acid, further prompting our search for new T•A/A•T recognition elements. In addition, sequence-dependent changes in the microstructure of DNA (intrinsic minor groove width, minor groove flexibility, and inherent DNA curvature)\textsuperscript{7a-k} combined with structural and conformational changes among polyamides make the targeting of certain sequences less than optimal, leading us to explore whether other novel heterocyclic recognition elements could be discovered for use in DNA groove recognition within the unsymmetrical pairing paradigm.\textsuperscript{7l,8-10} Furthermore, from a medicinal chemistry point of view, a broader tool kit of sequence-specific recognition elements for DNA beyond polyamides would be useful as our artificial transcription factor program moves from cell culture\textsuperscript{11} to small animal studies.

We recently reported that the benzimidazole ring can be an effective platform for the development of modular paired recognition elements for the minor groove of DNA.\textsuperscript{9,10} The benzimidazole 6-5 bicyclic ring structure, though having slightly different curvature from the classic five-membered pyrrole-carboxamides, presents an “inside edge” with a similar atomic readout to
the DNA minor groove floor, effectively mimicking Py, Im, and Hp. We demonstrated that the imidazopyridine/pyrrole pair Ip/Py distinguishes G•C from C•G and the hydroxybenzimidazole/pyrrole pair Hz/Py distinguishes T•A from A•T, providing a solution to the unanticipated hydroxypyrrole instability limitation.\textsuperscript{9,10} The question arises whether this second generation solution to DNA recognition can be elaborated further, deleting incrementally almost all carboxamide linkages in the backbone of the hairpin motif.\textsuperscript{12}

We report here a new set of heterocycle dimer pairs,\textsuperscript{12} which represents a step from single base-pair recognition toward a two letter approach to molecular recognition of the minor groove of DNA (Figure C.1). We move from single letters to syllables. New heterocycles were designed by combining the T-specific hydroxybenzimidazole (Hz) with oxazole (No) rings and chlorothiophene (Ct) caps\textsuperscript{8} to afford the recognition elements No-Hz and Ct-Hz, respectively (Figure C.1). Quantitative DNase I footprinting titrations were used to determine DNA binding affinities of hairpin oligomers containing the No-Hz and Ct-Hz dimers paired with Py-Py dimer for each of the four Watson-Crick bases (Figure C.2). When positioned at the termini of hairpin polyamides, the No-Hz/Py-Py and Ct-Hz/Py-Py dimer pairs are found to target 5’-GT-3’ and 5’-TT-3’ sequences, respectively, with high affinity and good specificity. With the development of dimer pairs capable of recognizing a 5’-GT-3’ sequence of DNA, we could address the question whether a hairpin oligomer comprised of four dimer units will bind the site 5’-GTAC-3’, a sequence formally containing all four Watson-Crick base pairs. Such a molecule represents our first programmable oligomer, which demonstrates excellent DNA binding properties without containing a single pyrroleor imidazole-carboxamide based on the natural product distamycin, a design benchmark for biomimetic chemistry and the field of DNA recognition.

C.2 Experimental

C.2.1 Polyamide Synthesis

\textbf{Figure C.1} Structures of dimers. a) imidazole-hydroxybenzimidazole (Im-Hz), b) oxazole-hydroxybenzimidazole (No-Hz), and c) chlorothiophenehydroxybenzimidazole (Ct-Hz) dimer caps in comparison with their respective five membered ring systems. Hydrogen-bonding surfaces to the DNA minor-groove floor are bolded.
Hairpin polyamides were synthesized manually from Boc-β-PAM resin in a stepwise fashion using Boc-protected monomeric and dimeric amino acids according to established solid-phase protocols.13 Base Resin 1 (BR1) (H$_2$N-Py-Py-γ-Im-Py-Py-β-Pam) was synthesized in gram quantities using the following amino acid building blocks: Boc-Py-OBt (4), Boc-Im-OH (5), and Boc-γ-OH (6) (Scheme C.1). The base resins were then split into smaller batches for coupling to the final dimeric caps. Boc-protected amino acid monomers for Boc-Py-OBt (4) and Boc-Im-OH (5) were synthesized according to previously reported procedures.8,13,14 Dimeric cap synthesis for No-HzOMe-OH (7) and Ct-HzOMe-OH (8) are detailed in Section C.7 Supplemental Information. Couplings were achieved using preactivated monomers (Boc-Py-OBt) or HBTU activation in a DIEA and DMF mixture. Coupling times ran from 3 to 24 h at 25-40°C. Deprotection of the growing polyamide was accomplished using 80% TFA/DCM. Polyamides were cleaved from the resin by treatment with neat 3-(dimethylamino)-1-propylamine (Dp) at 80°C for 2 h and purified by preparatory reverse phase HPLC. Deprotection of the methoxy-protected polyamides was done using a mix of thiophenoxide in DMF at 80°C to provide the free hydroxy derivatives after a second HPLC purification: Im-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (1), No-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (2), Ct-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (3). (See Section C.7 Supplemental Information for full experimental details.)

**Figure C.2** Postulated hydrogen-bonding models for the 1:1 polyamide-DNA complexes with their matched sequence and their ball-and-stick representations. a) Im-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (1), b) No-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (2), and c) Ct-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (3).
C.3 Results

C.3.1 DNA Affinity and Sequence Specificity of Dimer Caps

Quantitative DNase I footprinting titrations were carried out for polyamides 1-3. All polyamides were footprinted on the 285-base-pair PCR product of plasmid pCW15. In all cases, the DNA-sequence specificity at the cap position (in bold) was determined by varying a single DNA base pair within the sequence, 5’-TXTACA-3’, to all four Watson-Crick base pairs (X = A, T, G, C) and comparing the relative affinities of the resulting complexes. The variable base-pair position was designed to be adjacent to the Hz ring, which has been shown to specify for T when paired across from Py, so as to be able to determine the binding properties of each compound to the following two base-pair sequences: AT, TT, GT, and CT.

The sequence specificity of the Im-Hz and Ct-Hz dimers for 5’-TXTACA-3’ were evaluated in polyamides 1 and 3, respectively. As expected, polyamide 3 bound its designed match site 5’-TTTACA-3’ (Ka = 2.4 x 10^9 M^-1) (Figure C.3, Table C.1) with both the Ct and the Hz halves of the dimer preferring to rest over the less bulky T in the asymmetric cleft of a T•A base pair. Placing the Ct ring adjacent to the Hz resulted in a 10-fold specificity for T > A using the Ct-Hz system. Polyamide 1, which contains the Im-Hz dimer, did not bind its designed match site 5’-GTACA-3’ with any appreciable level of specificity exhibiting affinities of Ka = 1.6 x 10^8 and 4.0 x 10^8 M^-1 for the GT and AT sites, respectively.

Oxazole cap (polyamide 2) was incorporated into the dimer cap system, and the affinity for its designed match site, 5’-TGTACA-3’ , was examined. Polyamide 2 successfully targeted its designed match site with an appreciable level of specificity (25-fold) and a match site affinity of Ka = 6.8 x 10^9 M^-1 (Figure C.3, Table C.1). With the development of the chlorothiophene and oxazole dimer caps, the range of targetable sequences by polyamides has been expanded (Table C.1).

C.3.2 Design of a Programmable Oligomer for 5’-GTAC-3’

The synthesis of oligomer 9 containing four dimer units was achieved via the stepwise addition of Boc-amino acid dimers in the same manner as previously described polyamide syntheses. This “third generation” oligomer’s binding properties were assessed in the same context as previously reported for first and second generation hairpin polyamides targeting the sequence, 5’-GTAC-3’, containing the four Watson-Crick base pairs. Footprinting of the oligomer on the previously characterized plasmid DEH10 showed a binding affinity of Ka = 2.3 x 10^10 M^-1 for the
match site 5'-GTAC-3' and affinities of $K_a = 3.5 \times 10^9$ and $9.8 \times 10^8 \text{ M}^{-1}$ for the mismatch sites 5'-GAAC-3' and 5'-GATC-3', respectively (Figure C.4). Such a result demonstrates that a compound consisting exclusively of 6-5 fused ring systems and minimal carboxamide linkages is able to maintain good levels of specificity and excellent binding affinity.

C.4 Discussion
Recent advances in hairpin polyamide designs have traditionally focused on developing new modes of single base-pair recognition by heterocyclic ring pairings. Previous studies, however, have highlighted the fact that the microstructure of DNA depends on the sequence in question. In addition, structural and conformational changes among polyamides are thought to have an impact on their binding affinity and specificity.

Figure C.3: Quantitative DNase I footprinting experiments in the hairpin motif for polyamides 1, 2, and 3, respectively, on the 278 bp, 5'-end-labeled PCR product of plasmid CW15: (lane 1) intact DNA; (lane 2) A reaction; (lane 3) DNase I standard; (lanes 4-14) 1 pM, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, 100 nM polyamide, respectively. Each footprinting gel is accompanied by the following: (top) Chemical structure of the pairing of interest; (bottom) binding isotherms for the four designed sites. θ values were obtained according to published methods. A binding model for the hairpin motif is shown centered at the top as a ball-and-stick model with the polyamide bound to its target DNA sequence. Imidazoles and pyrroles are shown as filled and nonfilled circles, respectively; β-alanine is shown as a diamond; the γ-aminobutyric acid turn residue is shown as a semicircle connecting the two subunits.

Recent advances in hairpin polyamide designs have traditionally focused on developing new modes of single base-pair recognition by heterocyclic ring pairings. Previous studies, however, have highlighted the fact that the microstructure of DNA depends on the sequence in question. In addition, structural and conformational changes among polyamides are thought to have an impact...
on DNA affinity and sequence specificity. Thus, we have taken a more global view of molecular recognition, where our efforts have expanded from designing modules that distinguish the four DNA base-pairs (i.e., pairing rules) to designing those that target short, discrete DNA sequences.

Upon incorporation and evaluation of the G specific Im ring into the Im-Hz dimer cap (polyamide 1), we were surprised to find that it failed to demonstrate any preference for its designed site in addition to displaying a significantly decreased affinity. The shortcomings of the Im-Hz dimer prompted a search for a ring system that was capable of specifying for G > C within the X-Hz context. The oxazole (No) cap (Figure C.1) was considered because of its structural resemblance to Im (both rings present a nitrogen atom capable of hydrogen bonding to the minor groove). When the No-Hz dimer was incorporated into polyamide 2, it was found to be specific for its designed sequence of 5'-TGTACA-3' with a 25-fold preference for G > C and an affinity of $K_a = 6.8 \times 10^9$ M$^{-1}$ at its match site (Figure C.3). The No-Hz dimer presents the same functionality to the minor groove as the Im-Hz dimer but with an enhanced ability to target a 5'-GT-3' site, which could be due to a combination oxazole lone-pair basicity and differential solvation/desolvation effects between oxazole and imidazole.

The Ct-Hz dimer cap represents our first effort to target a short sequence of DNA using sequence-inspired recognition elements. Studies have shown that Hz exhibits specificity for T > A at the N-1 positions relative to the polyamide N-terminus and that Ct polyamides exhibited specificity for T > A at the cap position with excellent polyamide affinities. We hoped that a hybrid dimer would impart specificity for the TT sequence while maintaining a biologically relevant affinity. Polyamide 3 bound its designed match site 5'-TTTACA-3' with an affinity of $K_a = 2.4 \times 10^9$ M$^{-1}$.

<table>
<thead>
<tr>
<th>Polyamide</th>
<th>5'-tATACa-3'</th>
<th>5'-tTTACa-3'</th>
<th>5'-tGTACa-3'</th>
<th>5'-tCTACa-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$9.7(\pm0.7) \times 10^7$</td>
<td>$4.5(\pm0.6) \times 10^8$</td>
<td>$1.7(\pm0.4) \times 10^9$</td>
<td>$\leq 1.0 \times 10^7$</td>
</tr>
<tr>
<td>2</td>
<td>$8.6(\pm0.3) \times 10^8$</td>
<td>$9.5(\pm0.3) \times 10^8$</td>
<td>$6.8(\pm0.4) \times 10^9$</td>
<td>$2.7(\pm0.5) \times 10^8$</td>
</tr>
<tr>
<td>3</td>
<td>$2.1(\pm0.3) \times 10^8$</td>
<td>$2.4(\pm0.2) \times 10^9$</td>
<td>$2.6(\pm0.4) \times 10^9$</td>
<td>$\leq 1.0 \times 10^7$</td>
</tr>
</tbody>
</table>

Table C.1 Affinities of X/Py ring pairs proximal to a hydroxybenzimidazole bicycle $K_a$ (M$^{-1}$).^a,b

---

^a) Values reported are the mean values from at least three DNase I footprinting titration experiments, with the standard deviation given in parentheses.  
^b) 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.
10⁹ M⁻¹ and a specificity of 10-fold for T•A over A•T. This result is attributed to the fact that both the sulfur and hydroxyl groups prefer to lie over the less-bulky thymine base in a T•A base pair and that the –OH of the Hz ring is able to form an energetically favorable hydrogen bond with the O(2) carbonyl of thymine. Combined, these attributes make this dimer the preferred solution for targeting consecutive thymine residues.

As a first step in the design of programmable oligomers devoid of directly linked pyrrole- or imidazole-carboxamides, we incorporated the new No-Hz dimer into a hairpin structure, oligomer 9, consisting only of 6-5 fused ring systems (Figure C.2). The oligomer was designed to target the site 5’-GTAC-3’ and is a third generation molecule from our previously reported hairpin polyamides,

Figure C.4 a) Postulated hydrogen-bonding model and structure of oligomer 9. b) Ball and stick representation of 9 and the 6-base-pair binding site with variable region (W = A or T) shown. c) Quantitative DNase I footprint titration experiment on the 5’-32P-labeled PCR product shown with an illustration and complete sequence of the 285 bp EcoRI/PvuII restriction fragment from plasmid pDEH10. Binding affinities are shown next to their respective binding sites and the match site is designated.
which were shown to code for the four Watson-Crick base pairs in a sequence specific manner.\textsuperscript{16} To evaluate the impact of removing four carboxamide linkages and moving to a system consisting of only 6-5 fused recognition elements, binding properties were evaluated using quantitative DNase I footprinting titrations. Oligomer 9 was found to bind its match site with an impressive affinity of $K_a = 2.3 \times 10^{10} \text{M}^{-1}$ while discriminating against its mismatch sites of 5'-GAAC-3' and 5'-GATC-3' with specificities ($K_a$ match/$K_a$ mismatch) of ~7- and ~23-fold, respectively (Figure C.4). The four 6-5 fused rings of oligomer 9 present an “inside edge” with complimentary shape to the minor groove floor as the carboxamide linkages of traditional Py-Im-Hp polyamides. The complementary bumps and holes fit together between the oligomer and the DNA surface, which is the key to specificity. The dimer recognition elements in oligomer 9 are linked by a single carbon-carbon bond, resulting in fewer degrees of rotational freedom, which may result in a reduced entropic penalty for minor-groove binding. Undoubtedly, much of the favorable energetics for complexation with DNA for all these molecules is a result of differential solvation. The oligomer’s large hydrophobic surface may result in increased favorable van der Waals interactions with the walls of the minor groove.

C.5 Conclusion

Hairpin polyamides containing the No-Hz and Ct-Hz dimer caps at the polyamide N-terminus are able to target 5'-GT-3' and 5'-TT-3' sequences with good affinity and specificity and represent new recognition elements for the minor groove of DNA. The No-Hz and Ct-Hz dimer caps represent attempts to broaden heterocycle designs beyond single base pair interactions. In addition, the development of the No-Hz cap has allowed for the design of a DNA binding molecule, which in a formal sense is no longer a polyamide, hence the term programmable oligomer. We are encouraged by the fact that this oligomer demonstrates excellent affinity for DNA while exhibiting good levels of specificity. We hope to apply these new heterocycles to the targeting of biologically relevant sequences in the context of integrating artificial transcription factors with living biological systems.

C.6 Notes and References


2. (a) Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X. Netropsin, A new antibiotic produced by a streptomyces. J. Am. Chem. Soc. 1951, 73, 341-343. (b) Arcamone, F. N. V.; Penco,


12. Heterocycle pair refers to cofacial stack (noncovalent interaction) whereas dimer refers to two covalently attached heterocycles.


16. Im-Hp-Py-Py-γ-Im-Hp-Py-Py-β-Dp targets 5’-WGTACW-3’ with $K_a = 7.0 \times 10^8 \text{ M}^{-1}$ and Im-Hz-Py-Py-γ-Im-Hz-Py-Py-β-Dp targets 5’-WGTACW-3’ with $K_a = 4.6 \times 10^8 \text{ M}^{-1}$ where $W = A$ or T


C.7 Supplemental Information

C.7.1 General
$N,N$-dimethylformamide (DMF), $N,N$-diisopropylethylamine (DIEA), thiophenol (PhSH), 3-dimethylamino-1-propylamine (Dp), Triethylamine (TEA), and thiourea were purchased from Aldrich. Boc-β-alanine-(4-carbonylaminomethyl)benzyl-ester-copoly(styrenedivinylbenzene) resin (Boc-β-Pam-resin), dicyclohexylcarbodiimide (DCC), hydroxybenzotriazole (HOBt), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 4-dimethylaminopyridine (DMAP), and Boc-β-alanine were purchased from NOVA Biochem. Trifluoroacetic acid (TFA) was purchased from Halocarbon. All other solvents were reagent grade from EMD Chemicals. Oligonucleotide inserts were synthesized by the Biopolymer Synthesis Center at the California Institute of Technology. Precoated silica gel plates 60 F254 for TLC and silica gel 60 (40 μm) for flash chromatography were from Merck. Glycogen (20 mg/mL), dNTPs (PCR nucleotide mix), and all enzymes, unless otherwise stated, were purchased from Boehringer-Mannheim. pUC19 was purchased from New England Biolabs, and deoxyadenosine [$\gamma$-$^{32}$P] triphosphate was provided by ICN. Calf thymus DNA (sonicated, deproteinized) and DNaseI (7500 units/mL, FPLC pure) were from Amersham Pharmacia. AmpliTaq DNA polymerase was from Perkin-Elmer and used with the provided buffers. Tris.HCl, DTT, RNase-free water, and 0.5 M EDTA were from United States Biochemical. Calcium chloride, potassium chloride, and magnesium chloride were purchased from Fluka. Tris-borate-EDTA was from GIBCO, and bromophenol blue was from Acros. All reagents were used without further purification. NMR spectra were recorded on a Varian spectrometer at 300 MHz in DMSO-$d_6$ or CDCl$_3$, with chemical shifts reported in parts per million relative to residual solvent. UV spectra were measured on a Hewlett-Packard model 8452A diode-array spectrophotometer. High-resolution FAB and EI mass spectra were recorded at the Mass Spectroscopy Laboratory at the California Institute of Technology. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) was conducted at the Mass Spectroscopy Laboratory at the California Institute of Technology.

C.7.2 Heterocycle Synthesis

The synthesis of compounds 4 (Boc-Py-OBt), 5 (Boc-Im-OH), 10 (Boc-Py-Bi-OH), 11 (Boc-Im-HzOMe-OH), 19 (Ct-OH), 10, 22 (Im-HzOMe-OH), and 16 (aryl diamine) have previously been reported.$^{8,10a}$

C.7.2.1 Methyl 5-methyl-1,3-oxazole-4-carboxylate (No-OMe) (14)

Compound 14 was prepared following an exact preparation reported by Suzuki et al.$^{18}$ and purified
by normal phase column chromatography. According to the Suzuki procedure a mixture of methyl-
alpha-isocyanoacetate (3.0 g, 30.2 mmol) and DBU (4.5 g, 30.2 mmol) in dry THF (40 mL) at
10°C, was treated with acetic anhydride (3.06 g, 30 mmol) in dry THF (10 mL) dropwise. The
reaction was allowed to warm to room temperature and stirred for 10 h. The solvent was removed
by rotoevaporation and water (100 mL) was added. The mixture was extracted with EtOAc (2 x
100 mL). The organic layers were combined, dried over sodium sulfate, filtered, and concentrated
in vacuo to provided 14 as a crude amber oil. The oil was subjected to column chromatography
using (3:2 Hex/EtOAc) to provide 14 (2.73 g, 64% Yield) as a crystalline white solid. TLC (3:2
Hex/EtOAc) $R_f$ 0.5; $^1$H NMR (DMSO-$d_6$) 8.32 (s, 1H), 3.77 (s, 3H), 2.55 (s, 3H); $^{13}$C (DMSO-$d_6$
161.9, 156.1, 150.3, 126.2, 51.5, 11.5; HR-MS (EI): calculated for C$_6$H$_7$NO$_3$: 141.0426; found:
141.0427.

C.7.2.2 5-Methyl-1,3-oxazole-4-carboxylic acid (No-OH) (15)
A mixture of 14 (1 g, 7.08 mmol), 1N NaOH (10 mL) and MeOH (5 mL) was stirred at 40°C
for 4 h. The MeOH was removed by rotoevaporation and the pH adjusted to pH = 2 with 1N
HCl. The precipitate was extracted with EtOAc (3 x 10 mL), the organics dried over sodium
sulfate and removed by rotoevaporation to provide 15 (738 mg, 82% Yield) as a fibrous white
solid. TLC (3:2 EtOAc/Hex +10% AcOH) $R_f$ 0.4; $^1$H NMR (DMSO-$d_6$) 8.27 (s, 1H), 2.45 (s, 3H);
$^{13}$C (DMSO-$d_6$) 163.0, 155.6, 150.1, 145.6, 11.6; HR-MS (EI): calculated for C$_5$H$_5$NO$_3$: 127.0269; found:
127.0268.

C.7.2.3 Methyl 7-methoxy-2-(5-methyl(1,3-oxazole-4-yl))benzimidazole-6-carboxylate
(NoHz(OMe)OMe) (17)
To a solution of 15 (0.3 g, 2.36 mmol) in DMF (4 mL) was added DIEA (915 mg, 1.23 mL, 7.08
mmol) and HBTU (895 mg, 2.36 mmol). The mixture was stirred at room temperature for 1 h,
followed by the addition of the aryl diamine 16 (463 mg, 2.36 mmol). The reaction was then heated
to 35°C and stirred for an additional 24 h. The reaction was allowed to cool to room temperature and
then poured into a separatory funnel containing water (200 mL). The water was then extracted with
EtOAc (2 x 100 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated in vacuo to provide a crude solid, which was dissolved in glacial acetic acid (5 mL) and heated to 90°C. The reaction was stirred for 12 h, followed by removal of the solvent by rotoevaporation. The resultant residue was subjected to column chromatography (4:1 EtOAc/Hex) to provide 17 as a thin film. Addition of hexanes to the film, followed by rotoevaporation and drying under high vacuum provided 17 as a white solid (379 mg, 56% Yield). TLC (4:1 EtOAc/Hex) Rf 0.75; 1H NMR (DMSO-d6) 8.49 (s, 1H), 7.51 (d, 1H, J = 8.7 Hz), 7.16 (d, 1H, J = 8.7 Hz), 4.34 (s, 3H), 3.77 (s, 3H); 13C (DMSO-d6) 166.7, 151.0, 149.3, 145.5, 139.0, 135.3, 125.6, 124.8, 114.2, 105.3, 60.8, 51.6, 11.3; HR-MS (EI): calculated for C14H13N3O4: 287.0906; found: 287.0914.

C.7.2.4 7-methoxy-2-(5-methyl(1,3-oxazole-4-yl))benzimidazole-6-carboxylic acid (NoHz(OMe) OH) (7)
A mixture of 17 (200 mg, 0.69 mmol), DMSO (0.8 mL), and 1N NaOH (4 mL) was stirred at 23°C for 3 h. The solution was diluted with 8 mL of H2O and the pH was adjusted to pH = 2 with 1N HCl (4.4 mL). The precipitate was filtered and washed with diethyl ether. The crude material was purified by silica gel chromatography (1:1 EtOAc/Hexanes + 2% AcOH to 3:1 EtOAc/Hexanes + 2% AcOH) to provide 7 (156 mg, 82% Yield) as a white solid after drying under high vacuum. TLC (3:1 EtOAc/Hex +2% AcOH) Rf 0.55; 1H NMR (DMSO-d6) 13.12 (s, 1H), 12.37 (s, 1H), 8.51 (d, 1H, J = 8.5 Hz), 7.15 (d, 1H, J = 8.5 Hz), 4.34 (s, 3H), 2.78 (s, 3H); 13C (DMSO-d6) 167.8, 151.1, 150.9, 149.3, 145.5, 139.0, 135.5, 125.8, 125.2, 115.2, 105.4, 60.9, 11.4; HR-MS

Scheme C.3 Synthesis of 7.
(FAB): calculated for $\text{C}_{13}\text{H}_{12}\text{N}_{3}\text{O}_{4}$: 274.0828; found: 274.0833 [M+H]$^+$. 

Scheme C.4 Synthesis of 8.

C.7.2.5 Ct-H2(OMe)OMe (19)

To a solution of 18 (294 mg, 1.80 mmol) in DMF (5 mL) was added DIEA (0.348 mL, 1.98 mmol) and HBTU (650 mg, 1.72 mmol). The mixture was stirred at room temperature for 1 h, followed by the addition of the aryl diamine 16 (300 mg, 1.80 mmol). The reaction was then heated to 35°C and stirred for an additional 24 h. The reaction was allowed to cool to room temperature and then poured into a Falcon tube containing cold water (40 mL) resulting in a cloudy precipitate. The Falcon tube was centrifuged at 14000 rpm for 10 min, the mother liquor decanted, and the precipitate dried under high vacuum. After drying, the crude solid was dissolved in glacial acetic acid (5 mL) and heated to 90 °C. The reaction was stirred for 12 h, followed by removal of the solvent by rotoevaporation. The resultant residue was subjected to column chromatography (3:2 EtOAc/Hex) to provide 19 as an off-white solid (647 mg, 57% Yield). TLC (3:2 EtOAc/Hex) $R_f$ 0.5; $^1$H NMR (DMSO-$d_6$) 12.85 (s, 1H), 7.89 (d, 1H, $J = 5.3$ Hz), 7.56 (d, 1H, $J = 8.5$ Hz), 7.27 (d, 1H, $J = 8.2$), 7.26 (d, 1H, $J = 5.3$), 4.30 (s, 3H), 3.80 (s, 3H); $^{13}$C (DMSO-$d_6$) 166.6, 151.2, 144.0, 139.8, 134.8, 129.7, 129.4, 126.3, 125.6, 123.2, 114.8, 106.1, 61.2, 51.8; HR-MS (EI): calculated for $\text{C}_{14}\text{H}_{11}\text{N}_{2}\text{O}_{3}\text{SCl}$: 322.0179; found: 322.0171.
C.7.2.6 Ct-Hz(OMe)OH (8)

A mixture of 19 (1.8 g, 5.58 mmol), 1N NaOH (36 mL) and MeOH (27 mL) was stirred at 45°C for 24 h. The MeOH was removed by rotoevaporation and the pH adjusted to pH = 5.5 with 1N HCl, resulting in a precipitate. The mixture was spun down in a centrifuge (5 min at 14,000 rpm). The tube was decanted leaving a white solid that was dried under high vacuum to provide 8 (2.6 g, Quant. Yield) as a light yellow solid. TLC (3:2 EtOAc/Hex + 10% AcOH) R_f 0.5; ^1H NMR (DMSO-d_6) 13.00 (s, 1H), 7.89 (d, 1H, J = 5.3 Hz), 7.57 (d, 1H, J = 8.5 Hz), 7.30 (d, 1H, J = 8.5), 7.25 (d, 1H, J = 5.3 Hz), 4.28 (s, 3H); ^13C (DMSO-d_6) 167.7, 151.1, 143.9, 139.7, 134.9, 129.6, 129.3, 126.4, 125.7, 123.2, 115.9, 106.1, 61.2; HR-MS (EI): calculated for C_{13}H_{19}N_2O_3SCl: 308.0022; found: 308.0033.

Scheme C.5 Synthesis of 1.
C.7.3 Polyamide Synthesis

Polyamides were synthesized from Boc-β-alanine-Pam resin (50 mg, 0.59 mmol/g) and purified by preparatory HPLC according to published manual solid phase protocols. The synthesis of batch resin BR1 (H$_2$N-Py-Py-γ-Im-Py-Py-β-Pam) has previously been reported.

C.7.3.1 Im-Hz-Py-Py-γ-Im-Py-Py-β-Dp (1)

22 (Im-HzOMe-OH) (25 mg, 88.5 μmol) was incorporated by activation with HBTU (32 mg, 84 μmol), DIEA (23 mg, 31 μl, 177 μmol) and DMF (250 μl). The mixture was allowed to stand for
15 min at room temperature and then added to the reaction vessel containing base resin BR1 H₂N-Py-Py-γ-Im-Py-Py-Py-β-Pam. Coupling was allowed to proceed for 12 h at room temperature. The resin-bound polyamide was then washed with DCM and treated as described in the deprotection protocol below to provide Im-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (1) (0.9 mg, 2.4 % recovery) as a fine white powder under lyophilization of the appropriate fractions. MALDI-TOF-MS: calculated for C₅₉H₇₀N₂₁O₁₀: 1232.56; found 1232.55 [M+H]⁺.

C.7.3.2 No-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (2)

7 (No-HzOMe-OH) was incorporated as described above for Im-Hz-OH (polyamide 1) to provide No-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (2) (1.5 mg, 4.1 % recovery) as a fine white powder under lyophilization of the appropriate fractions. MALDI-TOF-MS: calculated for C₅₉H₆₉N₂₀O₁₁: 1233.55; found 1233.54 [M+H]⁺.

C.7.3.3 Ct-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (3)

8 (Ct-HzOMe-OH) was incorporated as described above for Im-Hz-OH (polyamide 1) to provide Ct-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (3) (1.1 mg, 2.9 % recovery) as a fine white powder under lyophilization of the appropriate fractions. MALDI-TOF-MS: calculated for C₅₉H₆₇ClN₁₉O₁₀S: 1269.47; found 1269.47 [M+H]⁺.

C.7.4 Deprotection of the O-Methyl-Protected Polyamides

O-Methyl protected polyamides were cleaved from resin, purified, deprotected and subject to further purification using the following general procedure. Upon completion of solid phase synthesis, 3-dimethylamino-1-propylamine (Dp) (500 μL) was added to the synthesis vessel containing the resin (50 mg). The mixture was allowed to stand for 2 h at 85°C with occasional agitation. The resin was then filtered and the solution diluted to 8 mL using 0.1% TFA. The sample was purified by reversed phase HPLC and lyophilized to provide polyamides containing the O-methyl protected hydroxybenzimidazole unit (-HzOMe-) as a dry solid. The polyamides were then dissolved in DMF (200 μL) and added to a suspension of sodium hydride (40 mg, 60% oil dispersion) and thiophenol (200 μL) in DMF (400 μL) that was pre-heated for 5 min at 85°C. The mixture was heated for 2 h at 85°C. The mixture was then cooled to 0 °C and 20% TFA (7.0 mL) was added. The aqueous layer was washed three times with diethyl ether (8 mL) and then diluted to a total volume of 9.5 mL using 0.1% TFA. The mixture was then purified by reverse-phase HPLC to give the deprotected
Oligomer 9 was prepared using Kaiser oxime resin (0.65 meqiv/gram) from Nova Biochem. Oxime resin (0.1 grams), was added to a solid phase synthesis vessel. The resin was then washed with DCM (15 mL) followed by DMF (15 mL). In a separate vessel was added Boc-PyBi-OH (10) (60 mg, 0.168 mmol), HBTU (61 mg, 0.159 mmol), DIEA (43 mg, 60 μL, 0.336 mmol), and DMF (400 μL). The mixture was vortexed and allowed to activate for 20 minutes at room temperature. The
activated mixture was then added to the resin. The reaction vessel was shaken at room temperature overnight. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). The resin was then washed with a solution of 25% TFA in DCM (20 mL), followed by shaking with 25% TFA in DCM at room temperature for 25 min to deprotect the Boc-protected alkyl amine, followed by shaking at room temperature for 25 min to deprotect the Boc-protected amine of pyrrole. In a separate vessel, a mixture of Boc-ImHz-OH (11) was activated using HBTU, DIEA, and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated Boc-ImHz-OBt dimer was then added to the vessel containing (Resin-BiPY-NH₂). The mixture was shaken at room temperature for 2.5 h. The reaction vessel was then drained and washed with DMF (15 mL) and DCM (15 mL). A solution of 50% TFA in DCM was then washed over the resin (20 mL), followed by shaking at room temperature for 25 minutes. In a separate vessel, a mixture of α-Fmoc-γ-Boc-(R)-2,4-diaminobutyric acid, (α-Fmoc-α-Boc-γ-Daba, and (12) was activated using HBTU, DIEA, and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated (R)-Fmoc-α-Boc-γ-Daba-OBt monomer was then added to the vessel containing (Resin-BiPy-HzIm-NH₂). The mixture was shaken for 2 h at room temperature. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). The resin was then washed with a solution of 25% TFA in DCM (20 mL), followed by shaking with 25% TFA in DCM at room temperature for 25 min to deprotect the Boc-protected alkyl amine. In a separate vessel, a mixture of Boc-PyBi-OH (10) was activated using HBTU, DIEA and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated Boc-PyBi-OBt dimer was then added to the vessel containing (Resin-BiPy-HzIm-γDaba-NH₂). The mixture was shaken at room temperature for 2.5 h. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). The resin was then washed with a solution of 25% TFA in DCM (20 mL), followed by shaking with 25% TFA in DCM at room temperature for 25 min to deprotect the Boc-protected alkyl amine, followed by shaking at room temperature with 25% TFA in DCM for 25 min to deprotect the Boc-protected amine of pyrrole. In a separate vessel, a mixture of NoHzOMe-OH (13) was activated using HBTU, DIEA and DMF as described above for (10). The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated NoHzOMe-OBt dimer was then added to the vessel containing (Resin-BiPy-HzIm-γDaba-BiPyNH₂). The mixture was shaken at room temperature for
2.5 h. The reaction vessel was then drained and washed with DMF (15 mL) and DCM (15 mL) to provide (Resin-BiPy-HzIm-γDaba-BiPy-HzNo).

C.7.6 O-Methyl Deprotection

To the synthesis vessel containing (Resin-BiPy-HzIm-γDaba-BiPy-HzNo) was added DCM (800 μL) and BCl₃ (400 μL), (2M in heptanes). The mixture was shaken at room temperature for 2.5 h. The vessel was then drained and washed with DCM (15 mL). Deprotection was determined to be complete by analytical HPLC.

C.7.7 Cleavage From Resin

Following O-Methyl deprotection, DCM (200 μL) and MeNH₂ (1 mL) (2M in THF) was added to the synthesis vessel. The mixture was then shaken at 35°C for 12 h. The filtrate was collected from the synthesis vessel and the organics removed in vacuo. The remaining residue was dissolved in 0.1% TFA and purified using preparatory reverse phase HPLC. Lyophilization of the appropriate fractions provided oligomer 9 as a fine powder solid. MALDI-TOF-MS: calcd for C₅₅H₄₉N₁₉O₈: 1103.40; found 1103.41 [M+H]+.
Appendix D: Peptoid Cell Uptake Studies
Abstract

Peptoids 5-FAM (1) and 6-FAM (2), previously reported by the Kodadek group, were synthesized and their cell uptake properties evaluated in HeLa cells using confocal laser scanning microscopy. A dual imaging laser system was used to image the peptoids and a DNA-binding nuclear stain simultaneously. The results demonstrate that the peptoids 5-FAM (1) and 6-FAM (2) are not cell permeable in HeLa cells.
D.1 Summary

Peptoids are a class of N-alkylated poly-glycine oligomers that have recently been exploited for the rapid synthesis of protein-targeted combinatorial libraries in an effort to discover new bioactive compounds with superior cell permeability and protease sensitivity as compared to traditional peptide scaffolds.\textsuperscript{1-3} Recently, the Kodadek group has demonstrated that these compounds are often cell permeable and can be utilized as molecular recognition domains.\textsuperscript{4} A peptoid was shown to be effective as a transcriptional activation domain for GST-KIX in a HeLa cell reporter gene assay.\textsuperscript{5} The peptoid domain which bound a GST-KIX fusion protein with a $K_d$ of 11.6 $\mu$M was discovered from a library of 50,000 members screened against the murine CREB core KIX domain. A carboxyfluoresceinated version of this peptoid (mixture of 1 and 2, Figure D.1) was used for KIX binding studies in comparison to a scrambled peptoid sequence, which was shown not to bind the GST-KIX domain. In an effort to evaluate the potential use of this peptoid activation domain for use with different polyamide DNA binding scaffolds, the exact peptoid-fluorophore conjugates (1 and 2, Figure D.1) from the Kodadek study were synthesized independently (not as a mixture) as shown in Figure D.2. The cell uptake properties of these peptoid-FAM conjugates were assessed using confocal laser scanning microscopy and dual imaging studies with DNA binding dyes were performed to provide unambiguous location of the cell nuclei. These studies were performed in HeLa cells and show that the peptoid was completely excluded from the cell interior as shown in Figures D.3 and D.4.

![Figure D.1](image)

**Figure D.1** Compound 1, 5-FAM, and 2, 6-FAM.
Figure D.2 Synthesis of compound 1 and 2. For details, see Section D.2 Experimental.
Figure D.3 HeLa cell uptake studies for compound 1 (5-FAM) using 2μM concentration. Dual image with Hoechst 33342. (2-Photon Laser λ = 810 nm, 5% Power, BP480 - 520 nm filter.).
D.2 Experimental

D.2.1 Materials

D.2.1.1 Resin
Knorr Amide MBHA resin, Nova Biochem, 01-64-0459, A33927
Resin loading = 0.78 mmol/g
Amount = 200 mg

D.2.1.2 Stock solution preparation
Stock solution A

Figure D.4 HeLa cell uptake studies for compound 2 (6-FAM) using 2μM concentration. Dual image with Hoechst 33342. (2-Photon Laser λ = 810 nm, 5% Power, BP480 - 520 nm filter.)
2 mL of piperidine in 8 mL of DMF.

Stock solution B
3.7 g of bromoacetic acid (3) in 14 mL of DMF.

Stock solution C
7 mL of \( \text{N,N'} \)-diisopropylcarbodiimide in 7 mL of DMF.

D.2.1.3 Amine solution preparation
Solution M1
300 μL of allylamine (4) in 1.7 mL DMF.
Solution M2
510 μL of (R)-(+)\(-\alpha\)-methylbenzylamine (5) in 1.49 mL DMF.
Solution M3
1.5 g (1.53 mL) of \( \text{N-Boc-1,4-butane diamine} \) (6) in 2.47 mL DMF.
Solution M4
1.5 g (1.53 mL) of \( \text{N-Boc-1,4-butane diamine} \) (6) in 2.47 mL DMF.
Solution M5
500 μL of piperonylamine (7) in 1.5 mL DMF.

D.2.2 General bromoacetic acid addition procedure
Resin was treated with 2 mL of stock solution B followed by 2 mL of stock solution C. Next, the synthesis vessel was capped and microwaved for 15 sec on the low power setting followed by shaking and another 15 sec in the microwave. Next, the resin synthesis vessel was drained and washed with DMF (4 x 5 mL) followed by a final wash with 10 mL of anhydrous DMF.

D.2.3 General amine addition procedure
Resin was treated with the appropriate amine solution and agitated for 15 sec followed by 15 sec in the microwave on the low power setting. The reaction mixture was agitated for another 15 sec and microwaved again for 15 sec on the lower power setting followed by draining of the vessel and washed with DMF (4 x 5mL). Finally, a wash with 10 mL of anhydrous DMF was performed.

D.2.4 Peptoid synthesis procedure (Synthesis of Resin PR-1)
A solid phase synthesis vessel was charged with 200 mg of Knorr Amide MBHA resin and swelled
in DMF for 20 min with agitation. The resin was then washed 4 times with DMF, treated with 20% piperidine in DMF for 20 min (2 mL x 20 min each), washed with DMF (8 x 5 mL), and washed with 10 mL of anhydrous DMF.

1. General bromoacetic acid addition procedure.
2. General amine addition procedure using solution M1.
3. General bromoacetic acid addition procedure.
4. General amine addition procedure using solution M2.
5. General bromoacetic acid addition procedure.
7. General bromoacetic acid addition procedure.
9. General bromoacetic acid addition procedure.
10. General amine addition procedure using solution M5.

After step 10, the resin was washed with DMF (5 x 5 mL), dichloromethane (10 x 5 mL), and then dried under high vacuum. This resin was called **PR-1**. Next, a very small aliquot of resin was cleaved using 95% TFA/H₂O for 1 h at room temperature. The solution was diluted with water, filtered, and checked by HPLC and ESI-MS. HPLC showed a single peak for product and the ESI-MS results are listed below.

- Expected mass of product = 722.41
- Found [M+H]+ = 723.3, [M+Na]+ = 745.4

**D.2.5 Procedure for first mini-PEG coupling (Synthesis of Resin **PR-2**)**

The resin, **PR-1**, from the peptoid synthesis procedure above was swelled in DMF for 30 min and drained. A separate vial was charged with 180 mg of Fmoc-8-amino-3,6-dioxaoctanoic acid (Fmoc-AEEA, **8**, CAS 166108-71-0, Peptides International), 244 mg of PyBOP, 92 μL of DIEA, and 1 mL of DMF. This reaction mixture was stirred at 23 C for 15 min and then added to the swelled resin **PR-1**. The reaction mixture was then put in a shaker at 37 C for 8 h. The resin was washed with DMF (5 x 5 mL), dichloromethane (10 x 5 mL), and then dried under high vacuum to give resin **PR-2** (Yield = 351 mg of dry resin). Next, a very small aliquot of resin was cleaved using 95% TFA/H₂O for 1 h at room temperature. The solution was diluted with water, filtered, and checked by HPLC and ESI-MS. HPLC showed a single peak for product and the ESI-MS results are listed below.

- Expected mass of product = 1089.55
- Found [M+H]+ = 1091.6, [M+Na]+ = 1113.6
D.2.6 Preparation of peptoid 5-FAM (1)

The resin PR-2 (20 mg) from the peptoid synthesis procedure above was swelled in DMF for 20 min and drained. Next, the Fmoc protecting group was removed by addition of 20% piperidine in DMF (2 × 4 mL for 10 min each). The resin was then washed with DMF (10 × 2 mL) then 5 mL of anhydrous DMF. Next, the deprotected resin was treated with a solution of 2 mg of 5-carboxyfluorescein succinimidyl ester (5-FAM-SE, Molecular Probes, C2210) in 300 μL of DMF. The reaction flask was covered with foil and agitated at 23 C for 1 h followed by 37 C for 1 hr. Next, the resin was drained, washed with DMF (10 × 1 mL), dichloromethane (10 × 1 mL), and dried under high vacuum. Next, the FAM labeled peptoid was cleaved from resin using 95% TFA/2.5% H2O/2.5% triisopropylsilane for 1 h. The crude product was purified by preparative reverse phase HPLC eluting from 10% acetonitrile/90% (0.1% TFA-H2O) to 50% acetonitrile/50% (0.1% TFA-H2O) over 70 min. The pure product eluted at minute 47 to give 5-FAM (1.74 μmol). Extinction coefficient for 5-FAM is 75000 at pH = 9.0, λmax = 498 nm.

1 (5-FAM): ESI-MS calculated for [M+H]+: 1226.5, observed [M+H]+: 1226.4

D.2.7 Preparation of peptoid 6-FAM (2)

Peptoid 6-FAM (2) was synthesized using the exact same procedure as for making 5-FAM except 6-carboxyfluorescein succinimidyl ester (6-FAM-SE, Molecular Probes, C-6164) was used. Purification was performed in the same manner and the elution time for the pure product was the same to give 6-FAM (1.13 μmol). Extinction coefficient for 6-FAM is 75000 at pH = 9.0, λmax = 498 nm.

2 (6-FAM): ESI-MS calculated for [M+H]+: 1226.2, observed [M+H]+: 1226.2

D.3 Notes and References


Appendix E: Cell Uptake Studies of 4G Targeting Polyamides

The research in this appendix is an extension of results from a combination of Chapter 7 and 8. Julie A. Poposki (Caltech) is thanked for HeLa cell culture work and assistance with cell uptake experiments.
Abstract

A benzimidazole-containing polyamide targeted to the sequence 5’-WGGGW-3’ was discovered to have fluorescent properties in aqueous solution in the presence and absence of DNA. This set of properties allowed for the unique opportunity to interrogate polyamide cell uptake in the absence of fluorophore conjugation, which alters the molecular architecture of a compound and often influences cell permeability. A 23-member library of the core benzimidazole polyamide with variations at the tail and turn was synthesized and uptake properties were evaluated in HeLa cells using 2-photon confocal laser scanning microscopy. Investigations of cellular localization using organelle specific dyes were also conducted along with the elucidation of an efflux mechanism for certain polyamides in the library.
E.1 Introduction

The core benzimidazole containing polyamide 2 from Chapter 6 (Figure E.1) was discovered to possess a unique set of photophysical properties where the emission maximum was found to be 475 nm and the excitation maximum was \( \lambda = 330 \) nm. Excitation in the range of 380–390 nm was used to minimize photobleaching while still producing a detectible amount of emission at 500 nm. This inherent fluorescent property allowed for the unique opportunity to study cellular localization of this polyamide architecture without attachment of a fluorophore, which can often influence the cellular uptake process. A library of polyamides (Figure E.2) containing the fluorescent core was synthesized, where the functionality at the turn and tail was modified in an effort to alter the overall uptake properties. HeLa cells were used for all studies and the polyamides were imaged using two-photon confocal laser scanning microscopy. After assessing the cellular uptake properties of the library of polyamide tail and turn modifications a study was conducted to determine the location of the polyamides in cells (mitochondria, nuclei, vescicals, lysosomes, etc.). Additionally, a polyamide that was found to remain extracellular was investigated to discriminate between cellular impermeability versus efflux using the the racemic efflux pump inhibitor (±)-Verapamil (2-(3,4-dimethoxyphenyl)-5-[2-(3,4-dimethoxyphenyl)ethyl-methyl-amino]-2-(1-methylethyl) pentanenitrile, CAS 52-53-9). The effect of ethanol fixing on the cellular uptake of polyamides as a function of time was also investigated and the ability of common transfection reagents to influence polyamide uptake was studied.

E.2 Results and Discussion

Figure E.1 Fluorescent polyamide 2 from Chapter 7 of this thesis. (bottom) Overlaid emission and excitation spectra of polyamide 2.
A focused library of 23 compounds was synthesized containing tail and turn modifications on a common polyamide core (Figure E.2). The polyamide was synthesized on Kaiser Oxime resin by standard solid-phase synthesis methods as shown in Scheme E.1 to provide a base resin loaded with the desired polyamide core used for tail and turn diversification. With the core in hand, diversification commenced with deprotection and cleavage from resin to afford triamine derivatives 3, 9, 20, 22, and 23 (Scheme E.2). Deprotection and resin cleavage with a small set of amine nucleophiles produced compounds 1, 2, 4-8, 10, 11, 13-16, and 19. Acylation prior to deprotection followed by cleavage from resin with amine nucleophiles produced compounds 12, 17, and 18. Deprotection and cleavage from resin with lithium hydroxide produce compound 21.

The cell permeability of this small focused library of compounds was evaluated in HeLa cells using 2-photon laser-scanning confocal microscopy and the results are shown in Figures E.3–E.8. The results are grouped according to their cellular localization profile with Figures E.3–E.6 showing compounds that appear to be cell permeable but non-nuclear,
Scheme E.1. Synthesis of R4. Standard solid-phase synthesis methods from Chapter 7 were used.
producing punctate cytoplasmic staining pattern. The compounds shown in Figure E.7 appear to be primarily extracellular and aggregated or precipitated with the first two compounds coating the exterior cell membranes. Figure E.8 shows compounds that are primarily extracellular although in a few cases there is a small amount of non-nuclear cellular localization observed.

In an effort to determine the location of the compounds inside the cells in Figures E.3–E.6 (punctate cytoplasmic staining) two organelle specific dyes were utilized. MitoTracker Red (CM-H2XRos) was used to assess the possibility of mitochondrial localization and LysoTracker Red (DND-99, a lysosome and trans-golgi stain) to assess localization in acidic lysosomes within the cell. Previous reports by Lown and coworkers demonstrated that a distamycin-fluorophore conjugate

**Scheme E.2.** Synthesis of compound library. Reaction conditions: (i) NMP, Triamine, 80 °C (Heat Block), 15-20 min; (ii) Conjugation to triamine, TFA, DCM; (iii) TFA, DMC; (iv) NMP, Nucleophile, 80 °C (Heat Block), 15-20 min; (v) Ac₂O, NMP.
primarily localizes in the mitochondria of human ovarian adenocarcinoma cells. The mechanism of mitochondrial staining by MitoTracker relies upon passive diffusion of the pro-fluorescent dye into the cell where it is oxidized and sequestered in the mitochondria (Figure E.9). Once in the mitochondria, nucleophilic attack resulting in thiol conjugation by proteins and peptides serves to retain the dye. Polyamide treated HeLa cells were dosed with MitoTracker prior to dual imaging using confocal laser scanning microscopy. The results show that the mitochondria appear as a diffuse cytoplasmic stain (Figure E.10, upper left live cell panel, magenta). The upper right panel shows the polyamide imaged with an orthogonal wavelength to Mitotracker and the bottom right panel shows that the two do not colocalize. It appears that the polyamide is not localized in mitochondria. Interestingly, ethanol fixing (fixed cell panel, bottom right, Figure E.10) of the cells leads to rapid nuclear trafficking of the polyamide however the diffuse mitochondria staining appears unchanged. It appears that unlike the Distamycin conjugates studied by Lown this class of polyamides does not localize in the mitochondria. However, this is not a surprising result given the dramatic differences in molecular structure and cell type.

Next, LysoTracker was used to determine if the polyamides were localized in acidic lysosomes of the cells. Figure E.11 shows the results of this study and it appears that the fluorescence
Figure E.4 Fluorescent polyamide cell uptake studies 5-8. 2-Photon laser imaging setup and polyamide dosing concentration as described in Figure E.2.

Figure E.5 Fluorescent polyamide cell uptake studies 9-11. 2-Photon laser imaging setup and polyamide dosing concentration as described in Figure E.2.
signal of the internalized polyamides colocalizes with the fluorescence signal of LysoTracker, indicating that the polyamides are likely sequestered in acidic vesicles within the cells. This result is consistent with previous studies showing that in all cases 8-ring polyamides with 4 imidazoles show poor nuclear uptake profiles.\(^3\) Weakly basic amines have also been shown to selectively accumulate within low pH cellular compartments such as lysosomes.\(^6\) It should also be noted that the \(N, N\)-dimethylaminopropylamine tail commonly used to terminate polyamides and other classes of minor groove binders is very similar to the functionality present in LysoTracker responsible for targeting it to the lysosomes and acidic vesicles of cells. Additionally, high imidazole content has been correlated with low cell uptake and the four contiguous imidazoles in the polyamide core may be basic enough to target the polyamides to the lysosomes without the influence of any other functionality, however benzimidazole involvement can not be ruled out.\(^3,4\)

The uptake properties of polyamide 23 in Figure E.8 were investigated in an effort to determine their mechanism of exclusion from the cells. Polyamide 23 (Figure E.8) dosed HeLa cells were treated with (±)-Verapamil and imaged using 2-photon laser confocal microscopy. The results in Figure E.12 show that treatment of polyamide dosed cells with (±)-Verapamil causes uptake...
of the polyamide however the polyamide appears to be localized in acidic vescicles (potentially lysosomes) as in Figures E.3–E.6. Figure E.12 also shows that treatment of the cells with ethanol allows the polyamide to readily traffic to the nucleus with significant nuclear localization after 2 h and complete localization after 4 h. Next the effect of a commonly used transfection reagent (Lipofectamine 2000) on polyamide nuclear localization was investigated (Figure E.13). The results show that the transfection reagent has little to no affect on the nuclear localization of this class of polyamide, however increased extracellular aggregation was observed.

Figure E.7 Fluorescent polyamide cell uptake studies 16–19. 2-Photon laser imaging setup and polyamide dosing concentration as described in Figure E.2.
Figure E.8 Fluorescent polyamide cell uptake studies 20–23. 2-Photon laser imaging setup and polyamide dosing concentration as described in Figure E.2.
**Figure E.9** Mechanism of mitochondrial staining using MitoTracker.

**Figure E.10** Results of HeLa cell uptake study using MitoTracker and compound 4. Dual Laser Imaging Setup: Coherent Chameleon 2-photon Laser, ($\lambda = 810$ nm, 5% power), HFT KP 680 dichroic, BP 480-520 nm MitoTracker, HeNe Laser ($\lambda = 543$ nm, 10% power), HFT KP 680 dichroic, BP 565-615 nm
Figure E.11 Results of HeLa cell uptake study using Lysotracker and compound 4. Dual Laser Imaging Setup: Coherent Chameleon 2-photon Laser, $\lambda = 810$ nm, 5% power), HFT KP 680 dichroic, BP 480-520 nm Lysotracker, HeNe Laser $\lambda = 543$ nm, 10% power), HFT KP 680 dichroic, BP 565-615 nm
Figure E.12 Cell uptake with (±)-verapamil. Compound 23 concentration = 2 μM in all experiments and (±)-verapamil concentration = 100 μM. HeLa cells used in all experiments. 2-Photon Laser Imaging Setup: Coherent Chameleon 2-photon Laser, (λ = 810 nm, 5% power), HFT KP 680 dichroic, BP 480-520 nm
Figure E.13 Cell uptake study with Lipofectamine 2000 and compound 23 (concentration = 2μM in all experiments). Lipofectamine 2000 (Lipo2000) concentration = 1μL/150μL media. HeLa cells used in all experiments. 2-Photon Laser Imaging Setup: Coherent Chameleon 2-photon Laser, (λ = 810 nm, 5% power), HFT KP 680 dichroic, BP 480-520 nm

E.3 Conclusion

The results presented in these studies of a focused 23-member library of tail and turn modified polyamides demonstrate that the uptake properties of benzimidazole-containing polyamides, targeted to the sequence 5’-WGCGGW-3’, can be divided into 2 broad categories. The first category includes compounds that are intracellular but non-nuclear, displaying a punctate cytoplasmic staining. The second category of compounds remain extracellular and have overall poor cellular uptake properties. Localization studies of non-nuclear intracellular polyamides indicated lysosomal sequestering as determined by LysoTracker Red DND-99 colocalization studies. This is in stark contrast to observations by Lown and coworkers who found that fluorescently-labeled dystamycin derivatives localized primarily to the mitochondria of human ovarian adenocarcinoma.
cells (SKOV-3). In addition, the mechanism of poor uptake for compound 23 was investigated using the efflux pump inhibitor (±)-verapamil and results from this study indicate that inhibition of efflux pumps causes rapid cellular uptake, however lysosomal sequestration is observed as indicated by colocalization studies with LysoTracker. As documented previously, ethanol fixation of cells causes rapid nuclear uptake due to increased membrane permeability. It has been reported that weakly basic amines selectively accumulate in cellular compartments with low internal pH and can be used to investigate the biosynthesis and pathogenesis of lysosomes. In addition to the use of weakly basic amines on the tail and turns of polyamides, high imidazole content has also been shown to be a negative determinant of nuclear localization. Currently, the cell uptake of high imidazole-containing polyamide sequences remains an unsolved problem.

E.4 Experimental

Synthesis of polyamides 1–23 was performed according to general solid phase synthesis protocols outlined in Chapter 7 and MALDI-TOF-MS data for all compounds is presented in Section E.6. The setup for 2-photon confocal laser scanning microscopy is reported in figure captions in this appendix.

E.5 Notes and References


E.6 Spectra and Supplemental Information

1: MALDI-TOF-MS calculated for [M]: 1155.51, observed [M+H]+: 1156.13
2: MALDI-TOF-MS calculated for [M]: 1348.63, observed [M+H]+: 1348.54
3: MALDI-TOF-MS calculated for [M]: 1240.57, observed [M+H]+: 1241.63
4: MALDI-TOF-MS calculated for [M]: 1181.53, observed [M+H]+: 1182.50
5: MALDI-TOF-MS calculated for [M]: 1112.47, observed [M+H]+: 1113.76
6: MALDI-TOF-MS calculated for [M]: 1084.44, observed [M+H]+: 1085.58
7: MALDI-TOF-MS calculated for [M]: 1114.45, observed [M+H]+: 1115.53
8: MALDI-TOF-MS calculated for [M]: 1128.47, observed [M+H]+: 1129.78
9: MALDI-TOF-MS calculated for [M]: 1362.57, observed [M+H]+: 1363.71
10: MALDI-TOF-MS calculated for [M]: 1183.51, observed [M+H]+: 1184.55
11: MALDI-TOF-MS calculated for [M]: 1197.52, observed [M+H]+: 1198.74
12: MALDI-TOF-MS calculated for [M]: 1197.52, observed [M+H]+: 1198.45
13: MALDI-TOF-MS calculated for [M]: 1141.50, observed [M+H]+: 1142.37
14: MALDI-TOF-MS calculated for [M]: 1153.50, observed [M+H]+: 1154.61
15: MALDI-TOF-MS calculated for [M]: 1178.49, observed [M+H]+: 1179.76
16: MALDI-TOF-MS calculated for [M]: 1201.52, observed [M+H]+: 1202.90
17: MALDI-TOF-MS calculated for [M]: 1223.54, observed [M+H]+: 1224.59
18: MALDI-TOF-MS calculated for [M]: 1240.57, observed [M+H]+: 1241.60
19: MALDI-TOF-MS calculated for [M]: 1198.55, observed [M+H]+: 1199.50
20: MALDI-TOF-MS calculated for [M]: 1587.59, observed [M-FITC]+: 1199.90; ESI-MS calculated for [M]: 1587.6, observed [M+H]+: 1588.1
21: MALDI-TOF-MS calculated for [M]: 1071.41, observed [M+H]+: 1072.51
22: MALDI-TOF-MS calculated for [M]: 1312.58, observed [M+H]+: 1313.53
23: MALDI-TOF-MS calculated for [M]: 1346.57, observed [M+H]+: 1347.78