

**Improving the Biological Activity of
Pyrrole-Imidazole Polyamides**

Thesis by
David Church Montgomery

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



California Institute of Technology
Pasadena, California
2013
(Defended March 11, 2013)

For my family

Thank you for always supporting me

Acknowledgements

As my time at Caltech comes to a close, I am left to reflect on all of the incredible people I have had the pleasure to know that have greatly contributed to my experience in this unique environment. Perhaps only as I now move forward can I look back and truly appreciate the tremendous caliber of scientists and staff that have surrounded me for the past five years and who have allowed me to be able to play a small role in the world-class research taking place around me. To the entire Caltech community, thank you.

To my research advisor, **Peter Dervan**, I am eternally grateful for you giving me the opportunity to work in your laboratory. I have learned a great deal from working with you, which I know will serve me well in the future. Your mentorship and support have been greatly appreciated, and your thoughtfulness and intelligence foster an excellent learning environment for the laboratory. To my committee, **Bob Grubbs**, **Judy Campbell**, and **Mark Davis**, thank you for supporting me and providing guidance over the years. I have always been struck by what an incredible wealth of knowledge, experience and intellectual analysis is present at my committee meetings, and you have all been very kind and professional throughout. In addition, I would like to thank **Linda Hsieh-Wilson**, who I worked with as a Ch7 TA. Your positive attitude toward teaching was really appreciated, and it was a pleasure to work for you.

Ben, what a blessing it's been having you as a baymate for the last five years. It's hard to quantify exactly what your friendship has meant throughout graduate school, but suffice it to say you've made my time at Caltech immeasurably better. Whether it was

talking research over coffee, talking anything and/or everything else over coffee, working late at night and blasting music, playing sports, going to concerts, etc., it's been a blast. You've been a great friend through the good times and the tougher times both in and outside of lab.

Jordan, I cannot thank you enough for all you have done for me over the last few years. You really helped me persevere through some tough times and get on the right track. You've provided an enormous amount of guidance, mentorship, support and advice. I find your practicality, determination and thoughtfulness inspiring. You have been a great friend outside of lab as well, and your ability to mix true professionalism with an easy-going, friendly and fun attitude is unparalleled.

Katy, your positive energy and sense of humor were infectious in the lab, and I always appreciated your guidance and advice. Perhaps the best compliment is that the lab really felt like it was missing something after you left. **JJ**, you have been a great friend and coworker, and I really appreciate your level-headedness and positive attitude, which make you always very pleasant to be around. **Jim**, you always seem to be willing to selflessly help when others ask for it. I have always appreciated your generosity, advice, and sense of humor. **Fei**, thank you so much for all of your help with various experiments and techniques over the years. You were very generous on many occasions and it didn't go unappreciated. **Mandy**, thanks for your assistance at various times over the years, you were a great collaborator and I appreciated your professionalism. **Dave C.**, thank you for

providing a great deal of assistance early on in my graduate career, you were always very thoughtful and generous.

To the rest of the **Dervan Lab**, it's been a pleasure working with all of you during my time at Caltech: Thomas, Jerzy, John, Justin, Dan H., Julie, Christian, Claire, Carey, Michelle, Mareike, Nick, Josh, Dan G., Sam, Jevgenij, Emil, Jamie, Tanja, Alissa, and Tim. What an incredible and talented group of people I've been able to work with over the years! I know I have learned a great deal from all of you, so thank you for providing a professional, comfortable, and intellectual environment at all times.

To my **friends outside of lab**, I'm so glad I was able to get to know all of you over the years, I couldn't have asked for a better supporting cast. In particular, my longtime roommates **Ian and Justin**, I'm so glad we ended up living together, thanks for choosing me and being so welcoming when I arrived and just being great friends throughout. I'd also like to give shout outs to Chris, Kristina, Keith, Chithra, Young-In, Myles, Jeanluc, Andrew, Thang, Allen, Matt W., Shannon, Eric L., Matthew V.W., Ted, Chethana and Eric, in no particular order, you all are the best! To my **non-Caltech friends**, thank you so much for being a constant source of support and being so welcoming over the years. You are all truly great friends, and I'm not sure I would have made it through grad school without you. In particular, thanks to Matt, Shin, Nick, Alyssa, Josie, Erin, Alex, Michelle, Adam, John K., Lam, Aaron, Mike. I'm blessed to have such incredible friends.

The **Caltech Staff** has always been extremely friendly and helpful to me. I'd like to thank Vicky Brennan and Lynne Martinez in particular, who have both done so much for

the Dervan lab over the years, always with a smile. I would also like to thank Joe Drew, Ron Koen, Paul Carroad, Steve Gould, Anne Penny and Agnes Tong for all that they do. Your hard work and friendly attitudes are appreciated, and your value to the department is immense. In addition, the Caltech facilities are all very well run under the leadership of Mona Shahgholi, David Vander Velde, Scott Ross, Shelley Diamond and Scott Virgil. Thanks for all that you do.

To **Megha**, thank you so much for helping me survive grad school, you were a constant source of support. You are so smart, caring, funny, and amazing in so many other ways that I couldn't attempt to list them all. I'm blessed to have you in my life. To the whole Mehrotra-Somani family and Sunil and Pravina in particular, I've always enjoyed for your warmth and kindness, and occasional hilarious shenanigans. You've always made me feel welcome and part of the family, thank you so much!

Finally, truly none of this would be possible without my family. To my **Mom, Dad, Lauren, Tom, and Susie**, and my entire extended family your constant support has helped me in so many ways get to where I am today. I love you and I can't ever thank you enough.

Abstract

DNA is nature's blueprint, holding within it the genetic code that defines the structure and function of an organism. A complex network of DNA-binding proteins called transcription factors can largely control the flow of information from DNA, so modulating the function of transcription factors is a promising approach for treating many diseases. Pyrrole-imidazole (Py-Im) polyamides are a class of DNA-binding oligomers, which can be synthetically programmed to bind a target sequence of DNA. Due to their unique shape complementarity and a series of favorable hydrogen bonding interactions that occur upon DNA-binding, Py-Im polyamides can bind to the minor groove of DNA with affinities comparable to transcription factors. Previous studies have demonstrated that these cell-permeable small molecules can enter cell nuclei and disrupt the transcription factor-DNA interface, thereby repressing transcription. As the use of Py-Im polyamides has significant potential as a type of modular therapeutic platform, the need for polyamides with extremely favorable biological properties and high potency will be essential. Described herein, a variety of studies have been performed aimed at improving the biological activity of Py-Im polyamides. To improve the biological potency and cellular uptake of these compounds, we have developed a next-generation class of polyamides bearing aryl-turn moieties, a simple structural modification that allows significant improvements in cellular uptake. This strategy was also applied to a panel of high-affinity cyclic Py-Im polyamides, again demonstrating the remarkable effect minor structural changes can have on biological activity. The solubility properties of Py-Im polyamides and use of formulating reagents with their treatment have also been examined. Finally, we describe the study of Py-Im polyamides as a potential artificial transcription factor.

Table of Contents

Acknowledgements	iv
Abstract.....	viii
Table of Contents	ix
List of Figures, Schemes, & Tables	xii
List of Abbreviations.....	xviii
List of Symbols and Nomenclature.....	xxi
Chapter 1: Introduction	23
DNA Structure and Function	24
The Central Dogma of Molecular Biology.....	25
Transcription Factors and Cellular Signaling.....	26
DNA-Binding Small Molecules	28
Recognition of the DNA Minor Groove.....	29
Pyrrole-Imidazole Polyamides.....	31
Development of Py-Im Polyamides with Therapeutic Potential.....	34
Regulation of Gene Expression with Py-Im Polyamides	35
Py-Im Polyamide-Mediated Gene Activation	39
Scope of This Work	40
References	41

Chapter 2: Enhancing the Cellular Uptake of Pyrrole-Imidazole Polyamides Through Next-Generation Aryl Turns	44
Abstract	45
Introduction	46
Results and Discussion.....	50
Conclusions	67
Materials and Methods.....	71
References	83
Chapter 3: Synthesis and Biological Activity of Cyclic Pyrrole-Imidazole Polyamide Libraries	87
Abstract	88
Introduction	89
Results and Discussion.....	91
Conclusions	108
Materials and Methods.....	109
References	123

Chapter 4: Characterization and Solubilization of Pyrrole–Imidazole Polyamide Aggregates.....	126
Abstract	127
Introduction	128
Results and Discussion.....	130
Conclusions	141
Materials and Methods.....	144
References	152
Chapter 5: Synthesis and Evaluation of Polyamide-Isoxazolidine Conjugates as Artificial Transcription Factors	156
Abstract	157
Introduction	158
Results and Discussion.....	167
Conclusions	191
Section 5B: Synthesis of WM Conjugates	192
Materials and Methods.....	195
References	209

List of Figures, Schemes & Tables**Chapter 1**

Figure 1.1: The 3-D structure of helical dsDNA	24
Figure 1.2: Structures of DNA-binding proteins	26
Figure 1.3: Structure of the interferon- β enhanceosome	27
Figure 1.4: Structures of actinomycin D and distamycin A	28
Figure 1.5: The electronic environment of the minor groove of DNA	29
Figure 1.6: Recognition of the DNA minor groove by distamycin.....	30
Figure 1.7: Recognition of the DNA minor groove by Py-Im polyamides.....	32
Figure 1.8: Structure of the Im-Py pairing adjacent to a G•C base pair	33
Figure 1.9: Polyamides can enter cell nuclei and access chromatin	34
Figure 1.10: The two routes of modulating transcription.....	35
Figure 1.11: The allosteric effect of polyamide-DNA binding	36
Figure 1.12: Transcription factor targets.....	38
Figure 1.13: A polyamide based transcriptional activator.....	39

Chapter 2

Figure 2.1: Current polyamide structural motifs	47
Figure 2.2: Aryl turns allow increased uptake	49
Figure 2.3: Cytotoxicity and DNA-binding of β -substituted polyamides.....	51
Table 2.1: Cytotoxicity of β -amino and β -aryl polyamides	52
Figure 2.4: Inhibition of gene expression by β -aryl polyamides.....	53
Figure 2.5: Inhibition of gene expression by β -aryl polyamides (cont.)	54
Figure 2.6: Time courses	55
Figure 2.7: Time course analysis of polyamide-mediated inhibition	56
Figure 2.8: Relative cytotoxicity of fluorescent analogues	57
Figure 2.9: Quantitative fluorescence analysis of β -turn substitution.....	58
Figure 2.10: Percentage of cells labeled as a function of time	59
Figure 2.11: Analysis of β -aryl and β -amino polyamide efflux.....	61
Figure 2.12: Confocal microscopy of β -aryl polyamide uptake	62
Figure 2.13: Time course of uptake of polyamide-FITC conjugates	63
Figure 2.14: Influence of polyamide dosage on cellular concentration	64
Figure 2.15. Cytotoxicity of various β -aryl polyamide cores.....	65
Figure 2.16. Inhibition of hypoxia-induced expression.....	66
Scheme S2.1: Monomers, dimers, and scheme for solid-phase synthesis	74
Scheme S2.2: Scheme for solution-phase derivatization of polyamides	75
Scheme S2.3: Complete structures of polyamides 21-26	75
Scheme S2.4: Scheme and structures for fluorescent polyamides	76

Chapter 3

Figure 3.1: Strategy for synthesis of differentially protected cycles.....	90
Table 3.1 Standard deprotection & microwave-assisted coupling times	91
Scheme 3.1: Microwave-assisted synthesis of cyclic polyamides 1-7.....	93
Table 3.2: MALDI-TOF data and yields for polyamides 1– 8	94
Scheme 3.2: Preparation of cyclic polyamides 9-11	95
Scheme 3.3: Preparation of FITC-labeled cyclic polyamides	96
Scheme 3.4: Preparation of cyclic polyamide 8	97
Scheme 3.5: Preparation of hairpin polyamide 17.....	98
Figure 3.2: Models of substituted turn along the DNA minor groove	100
Table 3.3: <i>T_m</i> Values for polyamide library.....	101
Table 3.4: SRB cytotoxicity data on compounds 1-3, 9, and 15.....	103
Table 3.5: SRB cytotoxicity data on compounds 4-7, 10-11, and 16	104
Figure 3.3: Confocal microscopy of cyclic conjugates, A549	106
Figure 3.4: Confocal microscopy of cyclic conjugates, T47D.....	107
Figure S3.1: Confocal microscopy with bright field images, A549.....	120
Figure S3.2: Confocal microscopy with bright field images, T47D	121
Figure S3.3: Cytotoxicity of Cbz-conjugated hairpin	122

Chapter 4

Figure 4.1: The use of solubilizing agents	129
Figure 4.2: Hairpin polyamide library	130
Figure 4.3: Cyclic polyamide library	131
Table 4.1: Estimated radii of polyamide aggregate particles	133
Figure 4.4: Soluble concentration of select polyamides.....	134
Figure 4.5: Soluble concentration of polyamides with cyclodextrin.....	136
Figure 4.6: HPLC traces of mouse plasma	138
Figure 4.7: FITC-labeled mouse injections	139
Figure 4.8: Timecourse of plasma addition to cell culture	140
Figure S4.1. Concentration vs. peak area plot from HPLC traces	148
Figure S4.2. Structures of carbohydrate solubilizing agents	149
Table S4.1: Estimated radii of polyamide aggregate particles	150
Table S4.2. Mass spectrometry for unpublished compounds.....	151

Chapter 5

Figure 5.1: Structures of DNA-binding proteins	159
Figure 5.2: Modes of polyamide-based gene regulation	160
Figure 5.3: A polyamide attached to a peptide activation domain	161
Figure 5.4: Synthetic activation domains.....	161
Figure 5.5: Luciferase assay for active and inactive isoxazolidines	162
Figure 5.6: General structure of a polyamide-isoxazolidine conjugate.....	163
Figure 5.7: Targeting the pG5luc promoter region.....	164
Figure 5.8: Isoxazolidine compounds sent from the Mapp laboratory	165
Figure 5.9: Proposed library of polyamide-isoxazolidine conjugates.....	166
Figure 5.10: Solid phase synthesis route.....	167
Figure 5.11: Cleavage and coupling of protected isophthalic acid	168
Figure 5.12: Synthesis of carboxylic-acid derivatized polyamide	169
Figure 5.13: Retrosynthesis of polyamide 7	170
Figure 5.14: Synthesis of t-butyl protected tail.....	171
Figure 5.15: Synthesis of tetramers.....	172
Figure 5.16: Solution-phase synthesis of full polyamide	173
Figure 5.17: Synthesis of acid-functionalized polyamide	175
Figure 5.18: Isoxazolidine coupling and <i>tert</i> -butyl deprotection	176
Figure 5.19: Initial synthesized stock of proposed conjugate library.....	177
Figure 5.20: Second batch of isoxazolidines from Mapp laboratory	178
Figure 5.21: Routes for linker attachment (long linker variant)	178
Figure 5.22: Synthesis of polyamide-isoxazolidine with longer linker.....	179

Figure 5.23: Compounds for initial luciferase assay by Mapp laboratory	180
Figure 5.24: Initial luciferase results from Mapp laboratory.....	180
Figure 5.25: Luciferase assay for longer linker conjugate	181
Figure 5.26: LNAR+ experimental timeline.....	183
Figure 5.27: Compounds used in LNAR+ assay	183
Figure 5.28: LNAR+ assay results.....	184
Figure 5.29: LNAR+ assay raw data.....	185
Figure 5.30: LNAR+ assay results for 43	186-7
Figure 5.31: Polyamide downregulation of GILZ transcription.....	188
Figure 5.32: qPCR data for polyamide-isoxazolidine compounds	189
Figure 5.33: WM target compounds	192
Figure 5.34: Synthesis of WM conjugate	194

List of Abbreviations

A	adenine
Ac	acetyl
Ac ₂ O	acetic anhydride
A ₅₉₅	absorbance maximum
AR	androgen receptor
ARE	androgen response element
A·T	adenine Watson-Crick hydrogen bonded to thymine
ATCC	American Type Culture Collection
β	beta-amino alanine
Boc	<i>tert</i> -butyloxycarbonyl
Boc-Im-OH	(4-[(<i>tert</i> -Butoxycarbonyl)amino]-1-methylimidazole-2-carboxylic acid)
Boc ₂ O	di- <i>tert</i> -butyl dicarbonate
Boc-Py-OBt	[(1,2,3-Benzotriazol-1-yl 4-[(<i>tert</i> -Butoxycarbonyl)amino]-1-methylpyrrole-2-carboxylate)
bp	base pair
BSA	bovine serum albumin
°C degrees	Celsius
C	cytosine
C·G	cytosine Watson-Crick hydrogen bonded to guanine
calc'd	calculated
Cbz	carbobenzyloxy
ChIP	chromatin immunoprecipitation
cm	centimeter
Da	Dalton
dATP	2'-deoxyadenosine triphosphate
DABA	diaminobutyric acid
DCM	dichloromethane
dex	dexamethasone
DFO	deferoxamine
DHT	dihydrotestosterone
DIEA	N,N-diisopropylethylamine
DMF	N,N-dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
Dp	N,N-dimethylaminopropylamine
DPPA	diphenylphosphoryl azide
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme-linked immunosorbant assay
EMSA	electrophoretic mobility shift assay
ESI	electrospray ionization
Et ₂ O	diethyl ether

List of Abbreviations

Ex	excitation
FBS	fetal bovine serum
FITC	fluorescein isothiocyanate
Fmoc	fluorenylmethyloxycarbonyl
γ -DABA	γ -2,4-diaminobutyric acid
G	guanine
G-C	guanine Watson-Crick hydrogen bonded to cytosine
GABA	γ -aminobutyric acid
HIF-1 α	hypoxia inducible factor 1 α
Hr	hour(s)
Hp	3-hydroxypyrrrole
HPLC	high-performance liquid chromatography
HRE	hypoxic response element
Im	N-methylimidazole
IPA	isophthalic acid
K _a	association constant
K _d	dissociation constant
λ	wavelength
LN ₂	liquid nitrogen
<i>m/z</i>	mass to charge ratio
μ	micro (1 x 10 ⁻⁶)
M	molar
m	milli (1 x 10 ⁻³)
Max	Myc associated protein X
max	maximum
MALDI	matrix-assisted LASER desorption/ionization
min	minute(s)
mol	mole(s)
mRNA	messenger ribonucleic acid
MS	mass spectrometry
N	A, T, G, or C
n	nano (1 x 10 ⁻⁹)
n-BuLi	n-butyl lithium
NF- κ B	nuclear factor- κ B
OBt	hydroxytriazole ester
p	pico (1 x 10 ⁻¹²)
PCR	polymerase chain reaction
PSA	prostate-specific antigen
Py-Im	pyrrole-imidazole
qPCR	quantitative polymerase chain reaction
RT	room temperature
RT-PCR	reverse transcriptase polymerase chain reaction

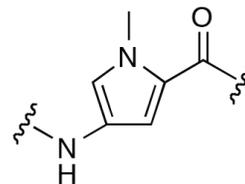
List of Abbreviations

PAGE	polyacrylamide gel electrophoresis
PBS	phosphate-buffered saline
Py	N-methylpyrrole
PyBOP	(benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
R	guanine or adenine
RCF	relative centrifugal force
RIPA	radio immunoprecipitation assay
RNA	ribonucleic acid
RNAi	ribonucleic acid interference
RT	reverse transcription
siRNA	small interfering ribonucleic acid
Smad	Sma and Mad-related protein
T	thymine
T·A	thymine Watson-Crick hydrogen bonded to adenine
t-BuOH	<i>tert</i> -butanol
TF	transcription factor
TFA	trifluoroacetic acid
TFO	triplex-forming oligonucleotides
THF	tetrahydrofuran
T _m	midpoint of transition temperature
TOF	time-of-flight
TFRE	transcription factor response element
tri/triamine	3,3'-diamino-N-methyldipropylamine
U	uracil
UV	ultraviolet
VEGF	vascular endothelial growth factor
Vis	visible
W	adenine or thymine

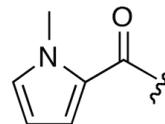
List of Symbols and Nomenclature



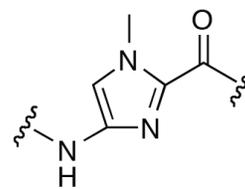
-Py-



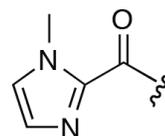
Py-



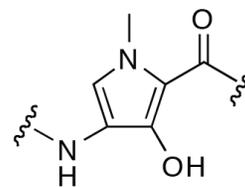
-Im-



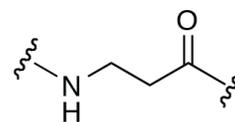
Im-



-Hp-



-β-



List of Symbols and Nomenclature

	-GABA-	
	-(R)-2,4-DABA-	
	-(R)-3,4-DABA-	
	-Dp	
	-tri-	
	-IPA	

Chapter 1

Introduction

DNA Structure and Function

Deoxyribonucleic acid (DNA) is a biological macromolecule that contains genetic information encoding for the entire structure and function of all known organisms. Various biological characteristics are accounted for by individual genes made up of specific sequences of DNA that together provide a blueprint for life, the full set of which is called a genome. DNA's primary structure consists of a backbone of alternating 2-deoxyribose sugar units and phosphate residues, which are joined by phosphodiester bonds at the 3' and 5' positions of the sugar. Attached to each sugar at the 1' position is one of four nucleobases: adenine (A), cytosine (C), guanine (G), or thymine (T). As it typically exists in nature, double-stranded DNA (dsDNA) contains two antiparallel strands of DNA in a double helix, forming a series of specific hydrogen bonding interactions between complementary nucleobases: A and T are specific for each other, as are C and G.

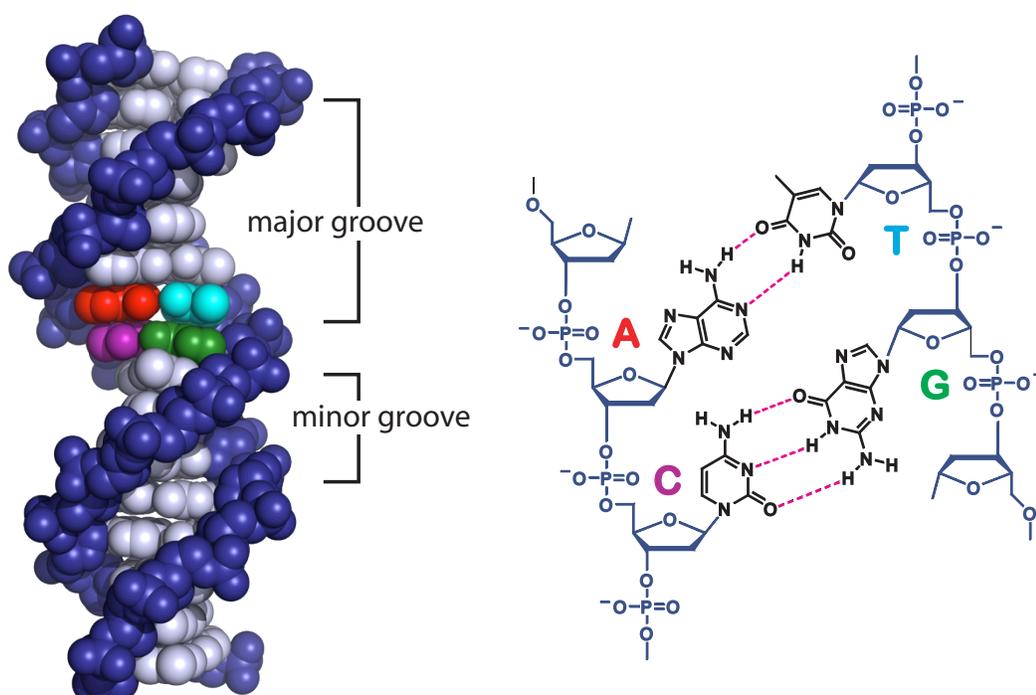


Figure 1.1: The 3-D structure of helical dsDNA¹ showing the backbone in blue and base pairs in gray, along with the minor and major grooves indicated and a set of base pairs highlighted (left), and the specific pairing of nucleobases attached to the sugar-phosphate DNA backbone (right). The two hydrogen bonding interactions between A and T and the three interactions between C and G are indicated in pink.

The Central Dogma of Molecular Biology

The term “central dogma” was first coined in 1958 by Walter Crick who later described it as dealing with “the detailed residue-by-residue transfer of sequential information [and] such information cannot be transferred from protein to either protein or nucleic acid.”² Proteins are a type of biomacromolecule that consist of chains of linked amino acids that (depending on the sequence) can perform a variety of functions within organisms. Proven over many years, the basic concept behind the central dogma is that DNA serves as the genetic code, storing information in the form of genes that eventually lead to the production of specific proteins, but proteins can never transmit sequence information to lead directly back to the corresponding DNA. Encompassed in this idea are the processes of transcription and translation, which together provide the information transfer from DNA to protein. Transcription is the process that copies DNA sequence information into corresponding ribonucleic acid (RNA) chains. RNA is structurally similar to DNA, except for the presence of a 2' hydroxyl group on the sugar moiety. There are a variety of types of RNA that can perform different functions, but in particular messenger RNA (mRNA) can allow for the production of proteins. In this process, called translation, sequence information is transferred from mRNA into a corresponding protein with a specific sequence of amino acids. Because proteins serve so many functions and misregulated protein expression is a factor in a variety of diseases, strategies to disrupt the pathway from DNA to RNA to protein for specific genes or subsets of genes may be of significant therapeutic value.

Transcription Factors and Cellular Signaling

Transcription factors (TFs) are a type of protein that can control the levels of transcription of a given gene, and often sets including multiple genes. This process is typically accomplished by non-covalent binding to a specific DNA sequence adjacent to a target gene, and either blocking (repression) or promoting (activation) the recruitment of RNA polymerase, the main enzyme responsible for transcription. Natural transcription factors vary in structure but typically consist of proteins bearing multiple domains that each serve specific functions such as a DNA-binding domain (DBD), a transcriptional activation domain (TAD), or a signal sensing domain (SSD). Transcription factors can bind to either the wider major groove or narrower minor groove of DNA, but in most cases, specific DNA-binding occurs via interaction between the DBD and 4-8 DNA nucleotides in a sequence specific fashion. These binding events can involve hydrogen bonding, electrostatic interactions, or van der Waals interactions with the base pairs or phosphate backbone of DNA. Proteins that bind to DNA, such as leucine zippers and zinc fingers, can bind as monomers, dimers, or in a combinatorial fashion where multiple DNA-binding proteins are required to regulate a certain gene.

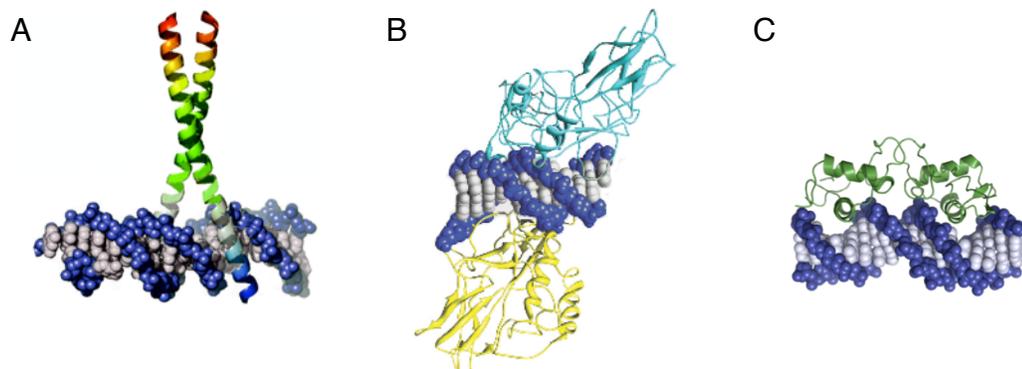


Figure 1.2: Structures of DNA-binding proteins; A) GCN4, a homodimeric leucine zipper; B) NF- κ B, a heterodimer of p50/p65; c) Androgen Receptor, which includes a ligand binding domain.

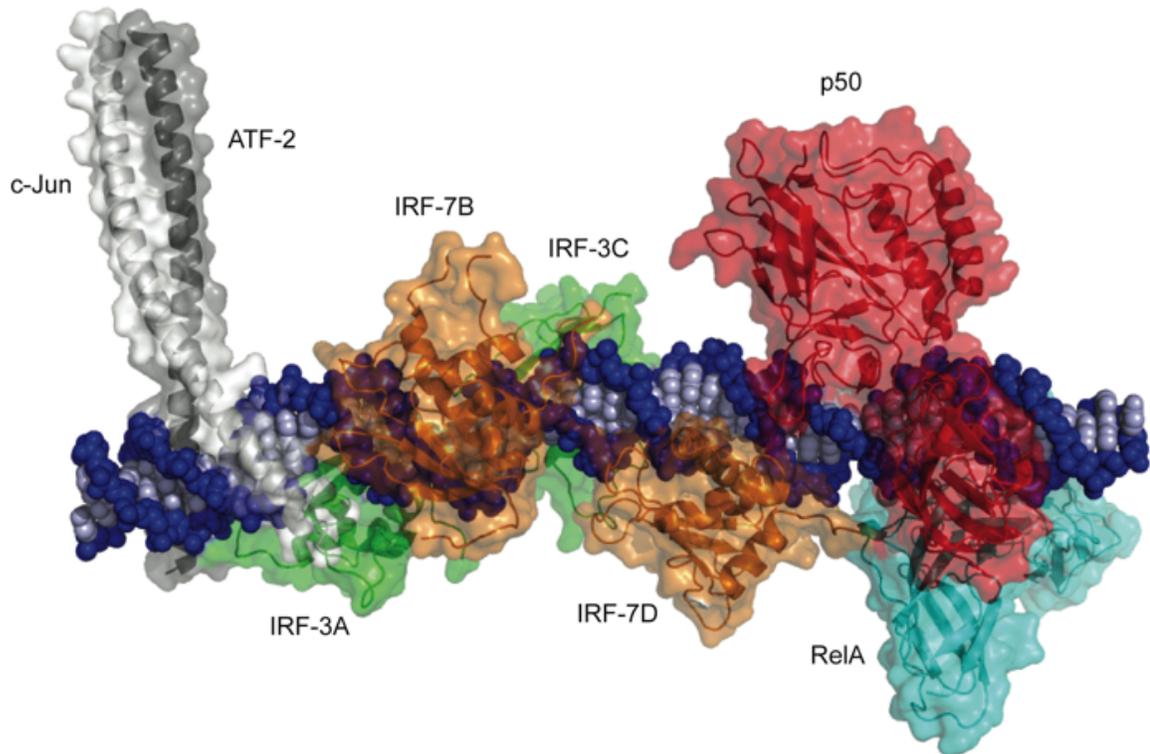


Figure 1.3: Structure of the interferon- β enhanceosome³, an example of combinatorial binding of a variety of proteins. This cooperative assembly occurs without any protein-protein contacts.

Nature has evolved complex networks of protein interactions called signaling pathways that can respond to external stimuli or allow for communication between cells. These pathways control processes crucial for the proper functioning of an organism, such as differentiation, stress response, cell growth, neurological function, metabolism, etc. Transcription factors are often major players in cellular signaling, so modulation of transcription can often cause major downstream (later in signaling pathway) effects or can counteract upstream (earlier in signaling pathway) effects. Therefore, disruption of transcription factor function is a potentially powerful strategy for developing therapeutics with many possible applications.

DNA-Binding Small Molecules

In addition to DNA-binding proteins, there are also many known small molecules capable of binding to DNA in a sequence specific fashion, most of which have been isolated as natural products. There are two major modes of non-covalent small molecule DNA-binding: intercalation and groove binding. Intercalation involves planar aromatic molecules which lodge between DNA base pairs, taking advantage of ring-stacking stabilization interactions with the adjacent nucleobases. In either major or minor groove binding, a variety of structures are possible, all of which take advantage of the unique electronic and steric environment that exists within the groove and vary with different sequences. An example of each binding mode are shown in Figure 1.4: actinomycin D, which preferentially binds to DNA by intercalation at 5'-GC-3' sequences;⁴ and distamycin A, which targets AT tracts of DNA by binding to the minor groove.⁵

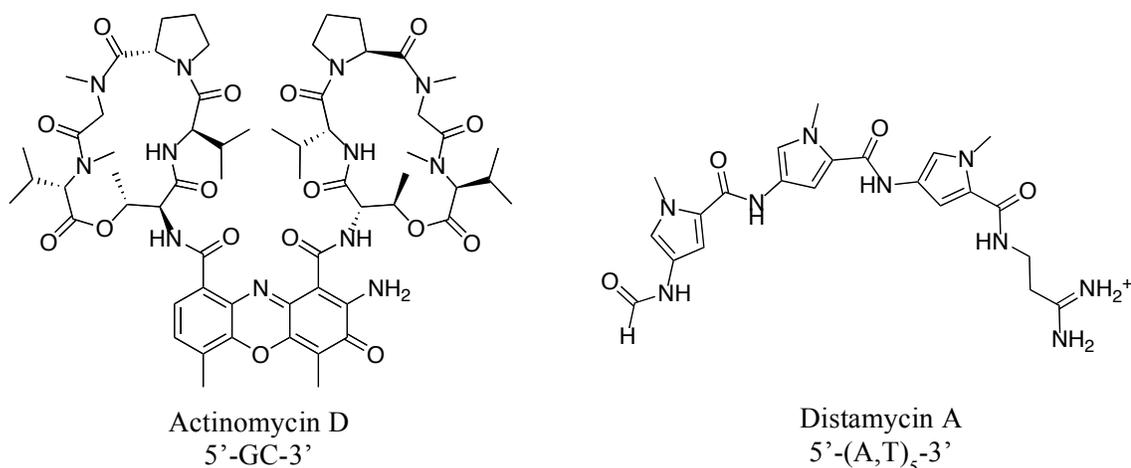


Figure 1.4: Structures of actinomycin D, an intercalator that targets 5'-GC-3', and distamycin A, a minor groove binder that targets 5'-(A,T)₅-3' sequences.

Recognition of the DNA Minor Groove

In the exposed portion of the minor groove, the plane of the base paired rings is perpendicular to the binding molecule, such that one side of the base pairing interaction is accessible (while the opposite side is accessible in the major groove). This imparts a unique electronic environment depending on the sequence of nucleotides, and a specific pattern can therefore be mapped to describe the electronic environment displayed by the minor groove of a specific sequence, as in Figure 1.5B.

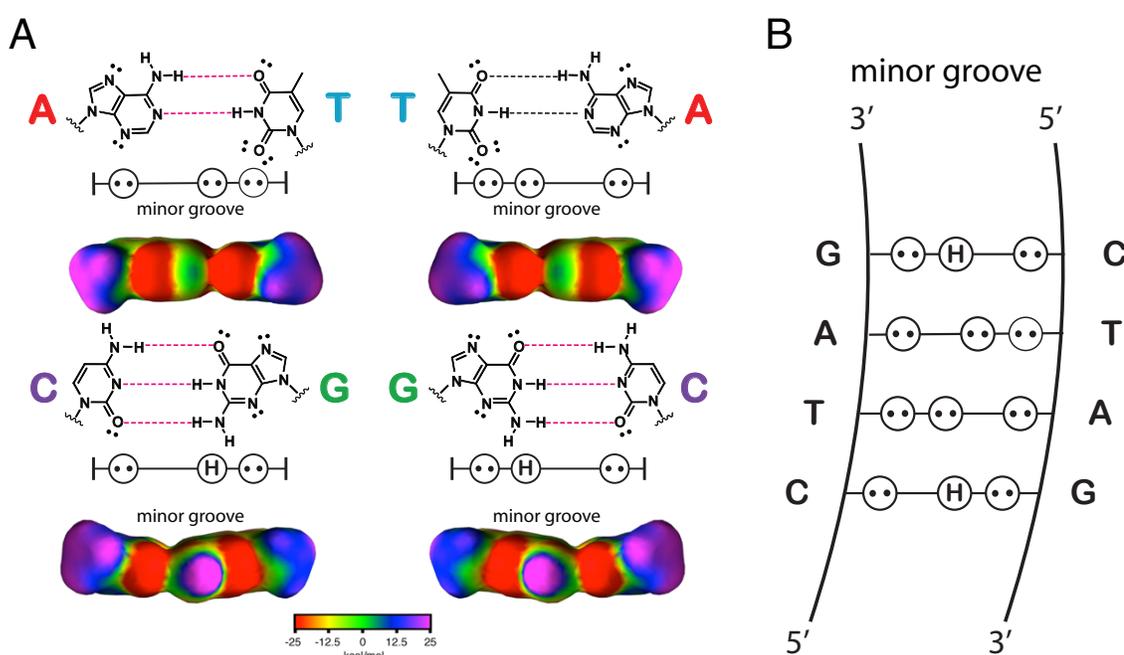


Figure 1.5: The electronic environment of the minor groove of DNA. A) For A•T base pairs, A displays one lone pair while T displays two lone pairs, whereas in C•G base pairs C displays one lone pair while G displays one lone pair and one N-H bond from the exocyclic amine. The unique electronic environments for each are clear from the corresponding electron density maps. B) The different patterns of lone pairs and N-H bonds displayed by a certain sequence can be mapped to provide the unique pattern of electron density for that sequence.

In the case of distamycin, the x-ray crystal structure of its interaction with DNA has been studied in detail.^{6,7} However, it was found that distamycin binds DNA in two different stoichiometries: 1:1 binding with DNA or 2:1 distamycin:DNA binding. Interestingly, crystal structures revealed hydrogen bond contacts between the amide N-H bonds of

distamycin and electron lone pairs of exposed lone pairs from base pairs exposed in the minor groove. In the 2:1 binding mode, the distamycin molecules bind in an antiparallel fashion, and each of the amide N-H bonds forms a favorable hydrogen bond with a lone pair on the same side of the minor groove.

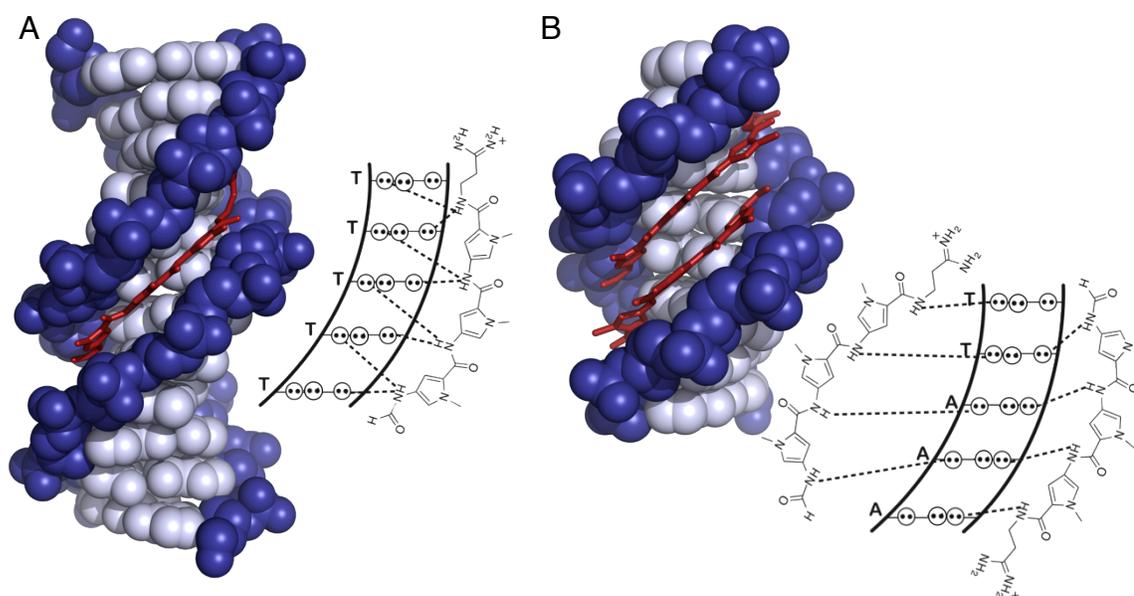


Figure 1.6: Recognition of the DNA minor groove by distamycin in A) a 1:1 complex and B) a 2:1 antiparallel complex.

This structural information from these crystal structures seem to indicate a favorable shape complementarity between distamycin and DNA, allowing for two molecules to bind the minor groove of DNA, each with monomeric pyrrole heterocyclic amino acids units oriented across from each other such that hydrogen bond interactions occur with a single base pair of DNA matched to a single pyrrole monomer unit on each molecule of distamycin.

Pyrrole-Imidazole Polyamides

Based on structure of distamycin, a new class of oligomeric heterocyclic DNA-binding small molecules called pyrrole-imidazole (Py-Im) polyamides have been developed that can be programmed to bind to specific sequences of DNA. In particular, by replacing a pyrrole subunit with an imidazole heterocycle, sequences including G•C or C•G base pairs can be targeted, and incorporation of a hydroxypyrrole heterocycle can distinguish between T•A and A•T base pairs.^{8,9} A set of pairing rules has been developed that allows for programmable targeting specific DNA sequences: N-methylpyrrole/N-methylimidazole (Py/Im) targets C•G, Im/Py recognizes G•C, 3-hydroxy-N-methylpyrrole/N-methylpyrrole (Hp/Py) and Py/Hp target T•A and A•T respectively.¹⁰ Finally, Py/Py pairs are degenerate for T•A and A•T.

In addition to the expanded range of sequences that can be targeted, a γ -aminobutyric acid (GABA) “turn” unit has been developed, to form a “hairpin” polyamide that provides a favorable entropic effect.¹¹ This structural motif has been shown to provide a 100-3600-fold increase in binding affinity for a target DNA sequence relative to the corresponding 2:1 motif.^{12,13} Additionally, a GABA turn unit can allow for the targeting of non-palindromic DNA sequences by incorporating unsymmetrical ring pairings. Further development of a chiral amino substituted GABA turn allows for binding in a single orientation, where a polyamide N- to C- terminus aligns with DNA in a 5' to 3' direction, known as “forward” binding.^{14,15}

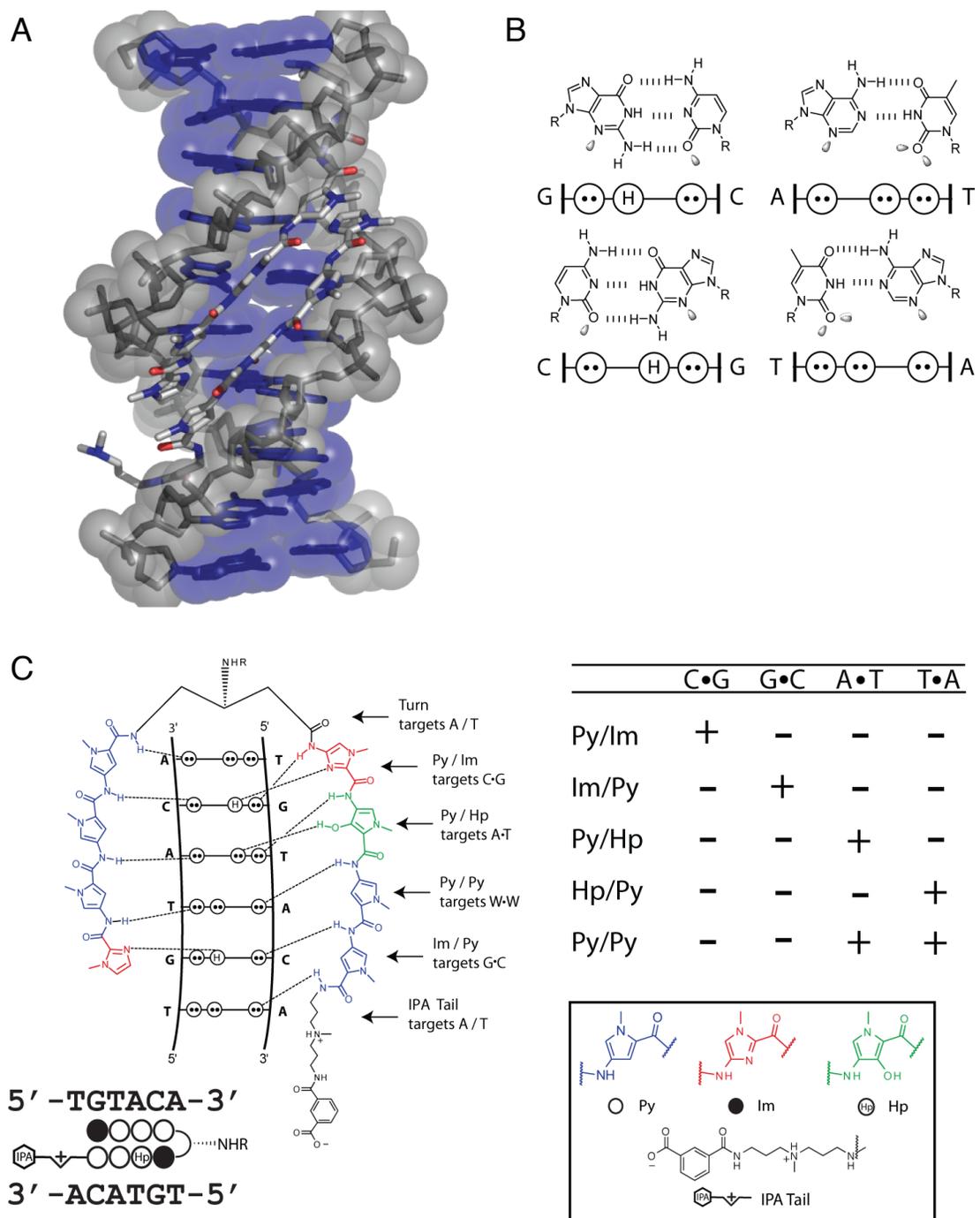


Figure 1.7: Recognition of the DNA minor groove by Py-Im polyamides. A) The structure of a Py-Im polyamide bound to DNA, showing clear ring stacking and shape complementarity. B) Electronic environment created by each base pair in the minor groove. C) Pairing rules for targeting all four base pairs with Py-Im polyamides. Ball-and-stick notation (alongside abbreviations) is also defined here, which is used as standard nomenclature for Py-Im polyamides.

As seen Figures 1.7 and 1.8, the Im subunit is specific for G nucleotides due to a hydrogen bonding interaction between the lone pair of the heterocyclic nitrogen of the Im and the N-H bond of the exocyclic amine of guanine. Hp monomers are specific for T nucleotides based on an interaction between the O-H bond of Hp (that is sterically accommodated by the cleft of the T•A base pair) and the extra lone pair displayed in the minor groove by thymine. Finally, the hairpin turn and the “tail” region at the C-terminus both target either the A•T or the T•A base pairs adjacent to the ends of the core of heterocycles, meaning that 8-heterocycle Py-Im polyamides targets six DNA base pairs.

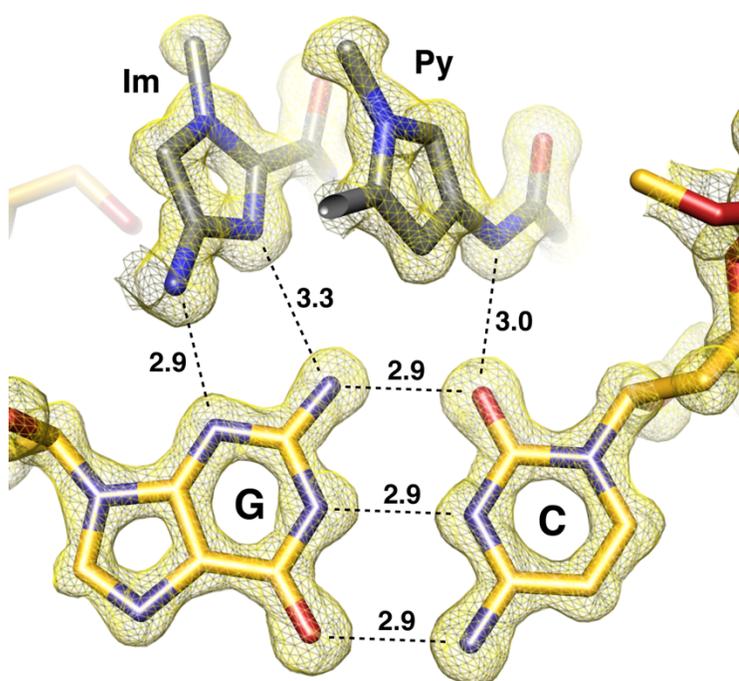


Figure 1.8: Crystal structure showing the orientation and hydrogen bonding interactions of an Im-Py heterocycle pairing adjacent to a G•C base pair.¹⁶ The interactions between amide N-H bonds and DNA lone pairs are clearly visible, as well as between the lone pair of the heterocyclic nitrogen and the N-H bond of the exocyclic amine of guanine. Also indicated are the three hydrogen bonds formed by the G•C base pair.

Development of Py-Im Polyamides with Therapeutic Potential

For Py-Im polyamides to be viable in a biological context, it is critical that they are able to enter live cells, and bind to DNA inside the cell nucleus. Previous cell culture studies have indicated that many fluorophore-labeled Py-Im polyamides are able to enter a variety of live cell lines, and traffic to the nucleus.^{17,18} In addition, crystal structures of Py-Im polyamides bound to nucleosome core particles demonstrated the ability of these compounds to bind to DNA in the context of chromatin, causing structural changes in nucleosomal DNA without significantly perturbing the structure of the histone octamer.¹⁹

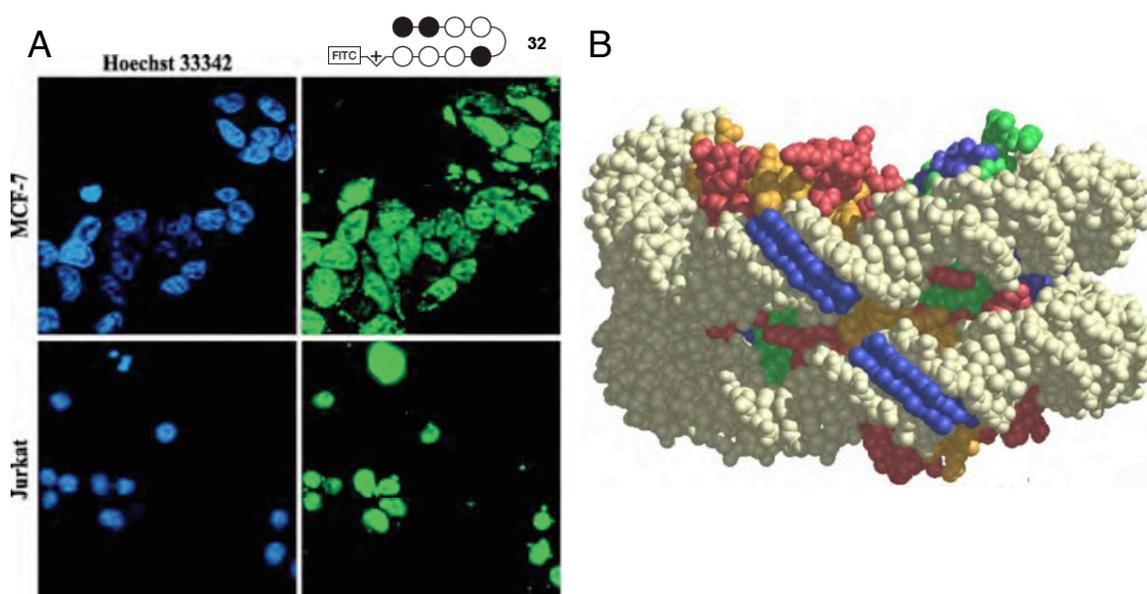


Figure 1.9: A) Py-Im polyamides can often enter live cells. In this experiment, a fluorescently labeled (FITC) polyamide is added to two live cell lines, each incubated with a nuclear stain (Hoechst 33342, left panels). In both cell lines, clear co-localization occurs.²⁰ B) Crystal structure of a nucleosome core particle, showing DNA-bound polyamide in blue.

Py-Im polyamides with isophthalic acid moieties attached at the tail region have also shown enhanced cell permeability properties, leading to a class of cell-permeable, sequence-specific, and high-affinity DNA-binding small molecules.²¹

Regulation of Gene Expression with Py-Im Polyamides

Due to the favorable biological properties of Py-Im polyamides and the power of modulating of gene expression for therapeutic purposes, the potential for using polyamides to repress or activate transcription is an attractive approach. In addition, the modular sequence-specificity of polyamides suggests that polyamides can be tailored to a variety of different biological targets for various therapeutic applications. Regulation of transcription could occur by two main routes: (1) repression of transcription by interfering with the function of existing transcription factors or (2) activation of transcription factor through conjugation of a DNA-binding polyamide to a TAD.

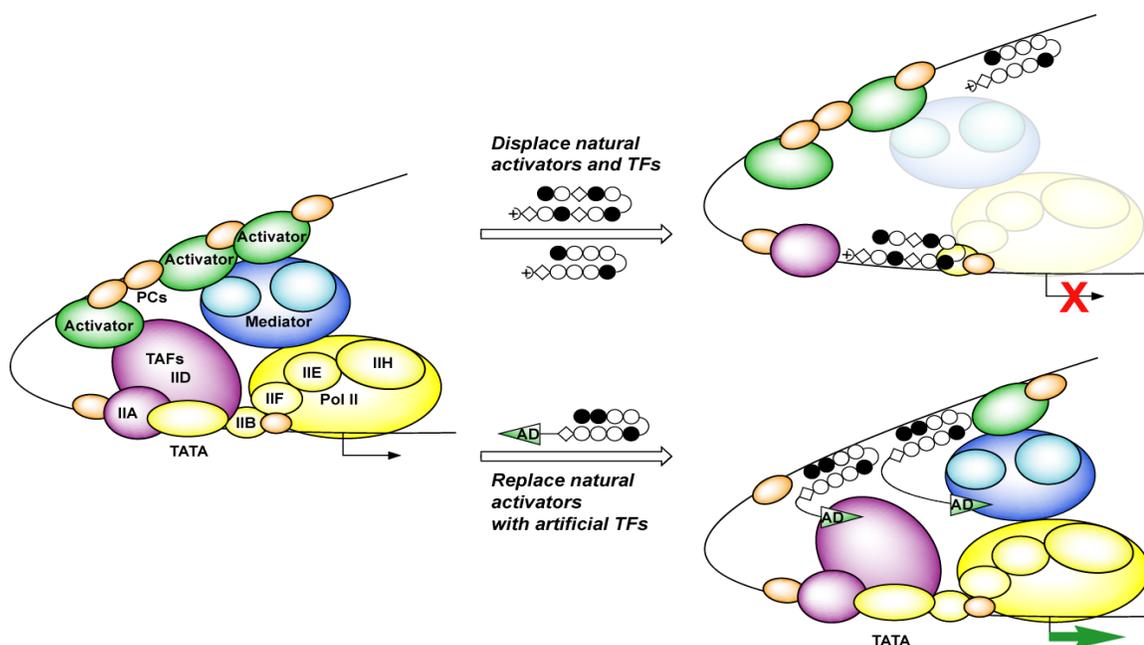


Figure 1.10: The two routes of modulating transcription. In the case of repression (top), a polyamide targeted to the binding site of a transcription factor may be able to interfere with the TF-DNA interaction, limiting the ability of the TF to recruit RNA polymerase and initiate transcription. For activation to occur (bottom), an activation domain could be attached to a polyamide targeted adjacent to a gene of interest, causing an artificial activation of transcription if the polyamide-TAD complex can properly recruit RNA polymerase.

While the strategy of interfering with the proper functioning of a transcription factor is a promising strategy for repression of transcription, there initially remained the question of exactly how a Py-Im polyamide could achieve this goal. Because many

transcription factors bind to the major groove of DNA, it is not immediately obvious how the process of a polyamide disrupting transcription factor function would work. A possible model for this to occur is evident when examining the crystal structure of polyamide-DNA complexes.

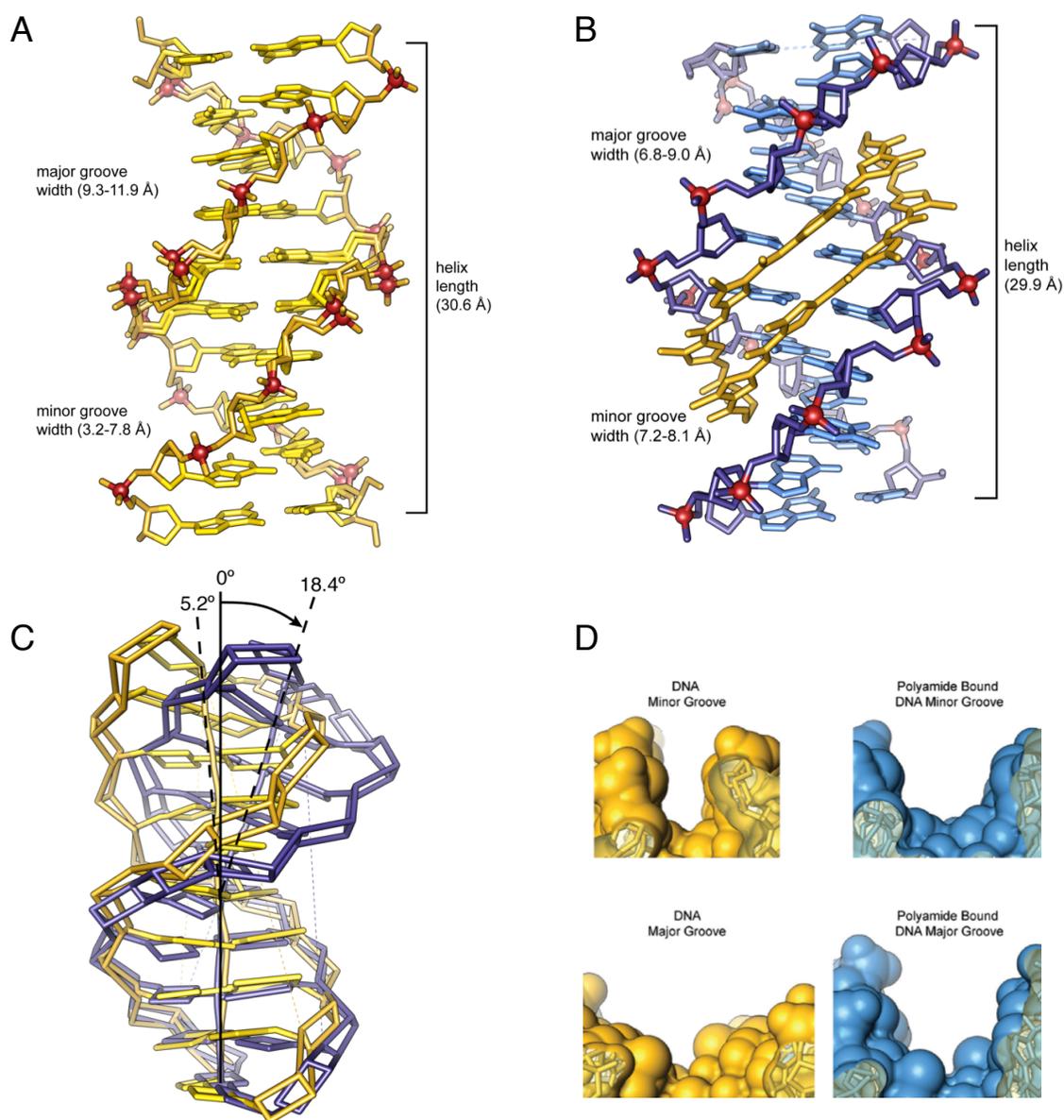


Figure 1.11: The allosteric effect of Py-Im polyamide-DNA binding.¹⁶ A) The structure of DNA with a narrow minor groove and wide major groove. B) When DNA is in a complex with a polyamide, significant widening of the minor groove occurs in order to accommodate the polyamide, causing a concomitant narrowing of the major groove. C) In addition to the effect on groove shape and width, the DNA strand becomes bent along the helical axis upon polyamide binding. D) An alternate view of the size and shape of the major and minor groove before and after polyamide binding.

As seen in Figure 1.11, significant deformation of the structure of DNA in the vicinity of the polyamide occurs upon binding. It is believed that this allosteric effect is the molecular basis for how polyamides binding to the minor groove of DNA can cause a disruption of the transcription factor-DNA interface for major groove-binding transcription factors.

One of the initial transcription factors targeted for transcriptional downregulation by polyamides was hypoxia inducible factor 1 α (HIF-1 α), which activates a set of genes in response to low oxygen environments by binding to the Hypoxic Response Element (HRE) consensus sequence 5'-TACGTG-3'. One of these genes is vascular endothelial growth factor (VEGF), which when activated can lead to tumor vascularization. In an induced system, a polyamide targeted to the HRE (Figure 1.12A) was demonstrated to repress VEGF expression via qRT-PCR.^{22,23} In contrast, treatment with a mismatch polyamide that does not bind the HRE did not significantly affect VEGF expression. On a global scale, microarray analysis indicated that the match compound downregulated a subset of genes activated by induction of HIF-1 α mediated expression. Chromatin immunoprecipitation (ChIP) experiments further demonstrated a decrease in occupancy of HIF-1 α at the VEGF HRE following treatment with the match compound, supporting a sequence-dependent mechanism.

Similar studies were conducted for transcription factors such as: (1) androgen receptor,²⁴ which drives the expression of a set of genes crucial in the development and progression of prostate cancer (Figure 1.12B), (2) glucocorticoid receptor,²⁵ which is a major regulator of inflammatory response (Figure 1.12C), and (3) NF- κ B,²⁶ which plays a

role in inflammatory response, cell differentiation, and senescence, and its misregulation has been implicated in cancer (Figure 1.12D).

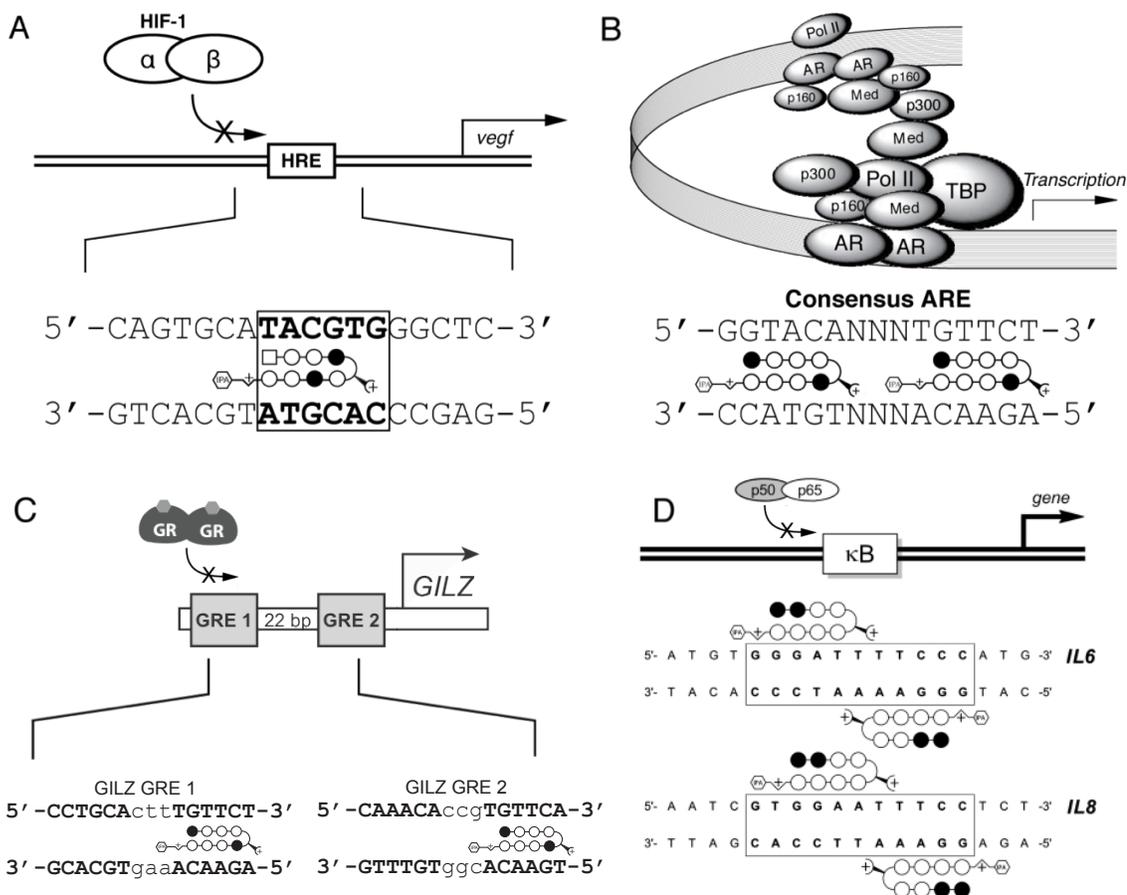


Figure 1.12: Various transcription factors that have been targeted with Py-Im polyamides, for which reduction of TF-mediated transcription has been demonstrated. In each case, the consensus sequence that was targeted is shown, along with the corresponding polyamide. A) Hypoxia inducible factor 1α, B) Androgen receptor, C) Glucocorticoid receptor, D) Nuclear factor κB.

Py-Im Polyamide-mediated Gene Activation

In contrast to repression of TF-mediated transcription, it may also be possible to activate specific genes using a Py-Im polyamide. Through the attachment of a transcriptional activation domain (TAD) capable of recruiting transcriptional machinery (Figure 1.13A), polyamides can be used to activate transcription at target sites. This construct is also referred to as an artificial transcription factor (ATF), as it functions like a transcription factor and contains similar domains, but can be entirely synthetic. These domains consist of a polyamide as the DNA-binding domain, and a synthetic molecule capable of recruiting activating transcription. Several polyamide-TAD conjugates have previously been demonstrated to upregulate transcription *in vitro*,²⁷⁻³⁰ but a lack of efficient cellular uptake—likely due to the large molecular weight of these compounds—prevented *in vivo* applications.^{10,31,32}

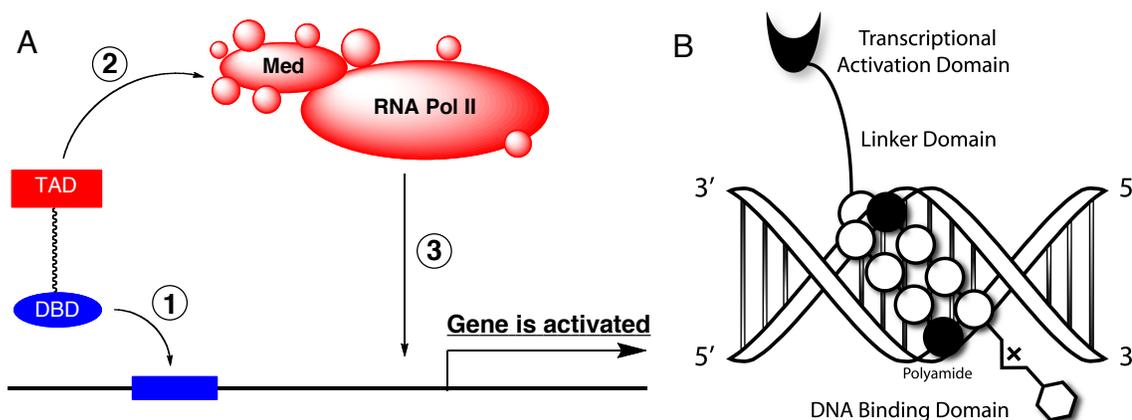


Figure 1.13: The concept of a polyamide based transcriptional activator. A) A small molecule containing DBD and TAD domains can bind to DNA (1), recruit transcriptional machinery (2), and allow for the assembly of the RNA Polymerase and initiation of transcription (3). B) General structure of a polyamide-TAD conjugate, indicating the DBD, linker domain, and TAD.

Scope of this Work

The work presented in this thesis involves various studies aimed at advancing Py-Im polyamides as a potential class of therapeutics. By making improvements that allow for increased biological activity in cancer cell lines, we have developed a more potent class of compounds with greatly increased cellular uptake. In Chapter 2, we present a new class of Py-Im polyamides bearing modifications to the γ -aminobutyric acid turn unit, with unprecedented levels of cytotoxicity in cell culture. We used confocal microscopy and flow cytometry to qualitatively and quantitatively assess levels of nuclear concentration, which strongly suggest that this effect was based on a significant increase in levels of nuclear uptake. We found that a simple modification to include an aryl-turn moiety can potentiate the biological effects of a Py-Im polyamide by up to two orders of magnitude. In Chapter 3, we develop an improved synthetic route that allows a practical method for unsymmetrical modifications to be made at the turn units of cyclic Py-Im polyamides. We then designed a panel of cyclic compounds to study the effects of a variety of substitutions that included aryl-turn technology, and found surprising results that further emphasize the dramatic effect that small structural changes can have on biological activity. In Chapter 4, we examine the aggregation propensity of a diverse library of Py-Im polyamides, finding that the solubility of polyamides does not seem to correlate with biological activity. We also investigate the effects of solubilizing agents, and find that this strategy can improve polyamide solubility and allow for improved injection conditions in mice. Finally, in Chapter 5 we develop a class of Py-Im polyamides conjugated to isoxazolidine moieties, and investigate their potential as artificial transcription factors.

References

1. Drew, H. R.; Wing, R. M.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. Structure of a B-DNA dodecamer - conformation and dynamics. *Proc. Natl. Acad. Sci. USA* **1981**, 78, 2179-2183.
2. Crick, F., Central dogma of molecular biology. *Nature* **1970**, 227, 561.
3. Panne, D.; Maniatis, T.; Harrison, S. C., An atomic model of the interferon-beta enhanceosome. *Cell* **2007**, 129 (6), 1111-1123.
4. Kamitori, S.; Takusagawa, F., Multiple binding modes of anticancer drug actinomycin-D - X-ray, molecular modeling, and spectroscopic studies of d(GAAGCTTC)₂-actinomycin D complexes and its host DNA. *J. Am. Chem. Soc.* **1994**, 116, 4154-4165.
5. Coll, M.; Frederick, C. A.; Wang, A. H. J.; Rich, A., A bifurcated hydrogen-bonded conformation in the d(A·T) base pairs of the DNA dodecamer d(CGCAAATTTGCG) and its complex with distamycin. *Proc. Natl. Acad. Sci. USA* **1987**, 84, 8385-8389.
6. Mitra, S. N.; Wahl, M. C.; Sundaralingam, M., Structure of the side-by-side binding of distamycin to d(GTATATAC)₂. *Acta Crystallogr. Sect. D* **1999**, 55, 602-609.
7. Pelton, J. G.; Wemmer, D. E., Structural characterization of a 2-1 distamycin A·d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. USA* **1989**, 86, 5723-5727.
8. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998**, 391, 468-471.
9. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., A structural basis for recognition of A·T and T·A base pairs in the minor groove of B-DNA. *Science* **1998**, 282, 111-115.
10. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, 13, 284-299.
11. Mrksich, M.; Parks, M. E.; Dervan, P. B., Hairpin peptide motif - a new class of oligopeptides for sequence-specific recognition in the minor-groove of double-helical DNA. *J. Am. Chem. Soc.* **1994**, 116, 7983-7988.
12. Trauger, J. W.; Baird, E. E.; Dervan, P. B., Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **1996**, 382, 559-561.
13. Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Cycle polyamide motif for recognition of the minor groove of DNA. *J. Am. Chem. Soc.* **1999**, 121, 1121-1129.

14. White, S.; Baird, E. E.; Dervan, P. B., Orientation preferences of pyrrole-imidazole polyamides in the minor groove of DNA. *J. Am. Chem. Soc.* **1997**, 119, 8756-8765.
15. Herman, D. M.; Baird, E. E.; Dervan, P. B., Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.* **1998**, 120, 1382-1391.
16. Chenoweth, D. M.; Dervan, P. B., Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. USA* **2009**, 106, 13175-13179.
17. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.* **2004**, 32, 2802-2818.
18. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.* **2002**, 10, 2767-2774.
19. Suto, R. K.; Edayathumangalam, R. S.; White, C. L.; Melander, C.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.* **2003**, 326, 371-380.
20. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA* **2003**, 100, 12063-12068.
21. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res.* **2007**, 35, 363-370.
22. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. USA* **2004**, 101, 16768-16773.
23. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.* **2007**, 2, 561-571.
24. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci. USA* **2007**, 104, 10418-10423.
25. Muzikar, K. A.; Nickols, N. G.; Dervan, P. B., Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, 106, 16598-16603.

26. Raskatov, J. A.; Meier, J. L.; Puckett, J. W.; Yang, F.; Ramakrishnan, P.; Dervan, P. B., Modulation of NF-kappaB-dependent gene transcription using programmable DNA minor groove binders. *Proc. Natl. Acad. Sci. USA* **2012**, 109, 1023-1028.
27. Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. USA* **2000**, 97, 3930-3935.
28. Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., Towards a minimal motif for artificial transcriptional activators. *Chem. Biol.* **2001**, 8, 583-592.
29. Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B., Design of artificial transcriptional activators with rigid poly-L-proline linkers. *J. Am. Chem. Soc.* **2002**, 124, 13067-13071.
30. Kwonj, Y.; Arndt, H. D.; Qian, M.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M., Small molecule transcription factor mimic. *J. Am. Chem. Soc.* **2004**, 126, 15940-15941.
31. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA* **2003**, 100, 12063-12068.
32. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res.* **2007**, 35, 363-370.

Chapter 2

Enhancing the Cellular Uptake of Pyrrole-Imidazole Polyamides Through Next-Generation Aryl Turns

The text of this chapter was taken in part from a manuscript co-authored with Jordan L. Meier and Peter B. Dervan (California Institute of Technology)

(Meier, J. L.; Montgomery, D. C.; Dervan, P. B. “Enhancing the cellular uptake of Py-Im polyamides through next-generation aryl turns” *Nucleic Acids Research* **2012**, 40 (5): 2345-2356)

Abstract

Pyrrole–imidazole (Py–Im) hairpin polyamides are a class of programmable, sequence-specific DNA binding oligomers capable of disrupting protein–DNA interactions and modulating gene expression in living cells. Methods to control the cellular uptake and nuclear localization of these compounds are essential to their application as molecular probes or therapeutic agents. Here, we explore modifications of the hairpin γ -aminobutyric acid turn unit as a means to enhance cellular uptake and biological activity. Remarkably, introduction of a simple aryl group at the turn potentiates the biological effects of a polyamide targeting the sequence 5'-WGWWCW-3' (W = A/T) by up to two orders of magnitude. Confocal microscopy and quantitative flow cytometry analysis suggest this enhanced potency is due to increased nuclear uptake. Finally, we explore the generality of this approach and find that aryl-turn modifications enhance the uptake of all polyamides tested, while having a variable effect on the upper limit of polyamide nuclear accumulation. Overall this provides a step forward for controlling the intracellular concentration of Py–Im polyamides that will prove valuable for future applications in which biological potency is essential.

Introduction

Hairpin pyrrole-imidazole (Py-Im) polyamides are a class of heterocyclic amino acid oligomers that can be programmed to bind a wide repertoire of DNA sequences with high affinity and specificity.^{1,2} Sequence-selective recognition of the minor groove of DNA is achieved through side-by-side stacked ring pairings: Im/Py distinguishes G·C from C·G, while Py/Py is degenerate for T·A and A·T. In recent years our group has focused on the biological evaluation of eight-ring polyamides arranged in a hairpin configuration through a γ -aminobutyric acid linker (Figure 2.1).^{3,4} These compounds are of modest (~1300 Da) molecular weight and recognize 6 base pairs (bp) of DNA, similar to the size of many eukaryotic transcription factor binding sites.⁵ When applied to living cells, hairpin polyamides can disrupt protein-DNA interactions and modulate the expression of genes induced by many transcription factors, including the ligand-activated nuclear receptors glucocorticoid receptor (GR) and androgen receptor (AR).⁶⁻¹⁰ However, one persistent challenge encountered when applying Py-Im polyamides to new biological systems is cellular uptake. Previous studies have shown that the nuclear localization of fluorescently labelled polyamides can be influenced by several variables including molecular weight, modifications to the C-terminal moiety, and composition of Py/Im content.¹¹⁻¹³ In particular, Py-Im polyamides incorporating multiple (>2-3) *N*-methylimidazole subunits show reduced nuclear localization, limiting the ability to target GC rich sequences *in vivo*.¹⁴ Therefore, new solutions for enhanced uptake are important for advancing Py-Im polyamides as probes of transcription factor binding and, potentially, as therapeutic agents.

While a large amount of work has been done on the effect of C-terminal modifications on polyamide nuclear localization,^{13,15} relatively few studies have explored the impact of modifications to the γ -aminobutyric acid (GABA) turn moiety on biological activity. The incorporation of a chiral (R)- α -amino substituent on the GABA turn enforces polyamide binding in a N-terminus \rightarrow C-terminus orientation with respect to the 5' \rightarrow 3' direction of the adjacent DNA strand.¹⁶

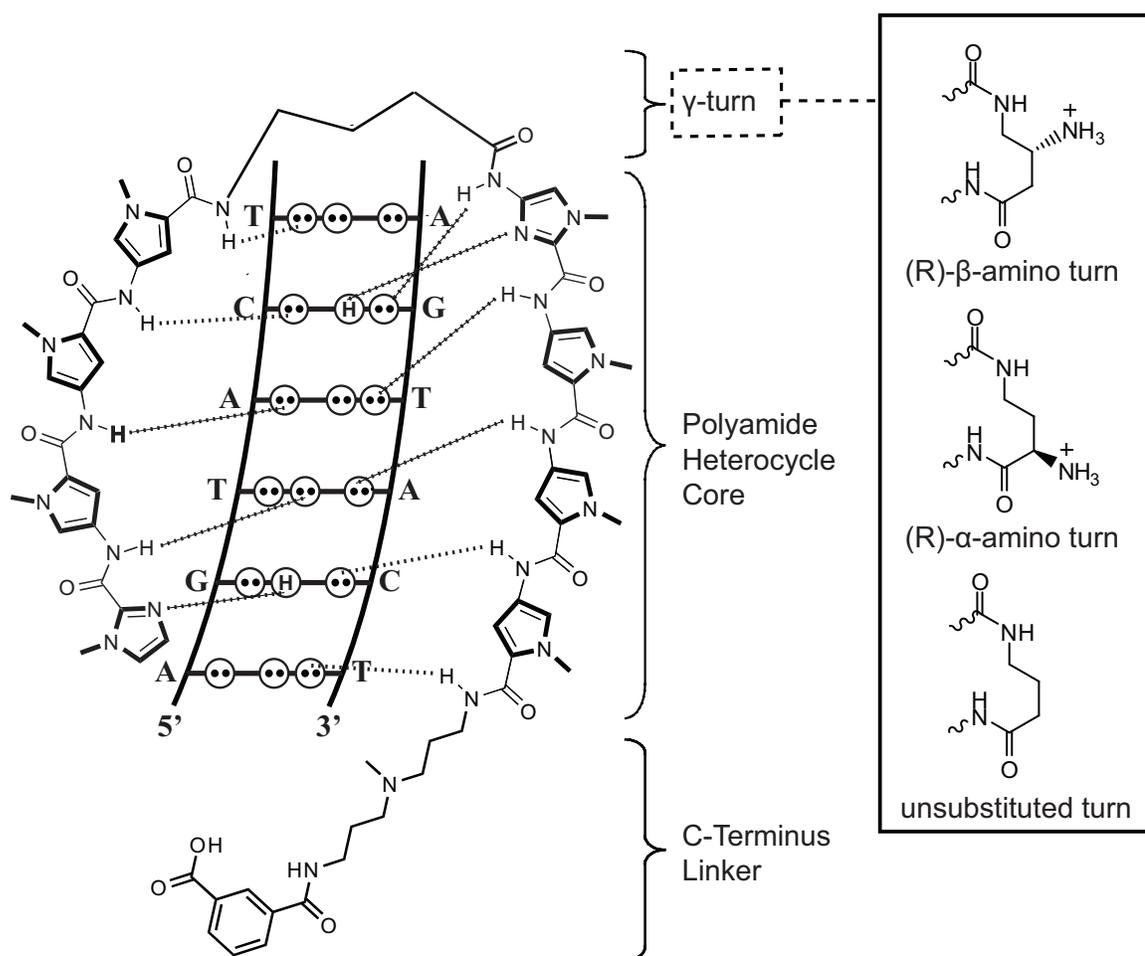


Figure 2.1: Schematic diagram of eight-ring Py-Im polyamides targeting the sequence 5' -WGWWCW-3' (W=A/T). Dashed lines indicate hydrogen bonds between the polyamide and DNA base pairs. The γ -aminobutyric acid turn unit enforces an antiparallel hairpin configuration, and codes for W.¹⁶ Substitution of the chiral turn functionality can have substantial effects on DNA-binding and biological activity.

More recently, we introduced hairpin polyamides bearing (R)- β -amino GABA turns which show increased DNA-binding affinity and, importantly for the purposes of this study, a negligible decrease in DNA-binding affinity upon acylation (Figure 2.1).¹⁶ Structural studies suggest this is due to their unique stereochemical presentation of the chiral β -amino moiety, which is directed up and out of the minor groove floor, thereby providing a chemical handle for introduction of functionality at the turn position.¹⁷ Furthermore, biological evaluation of a β -acetylated polyamide targeted to the consensus androgen response element (ARE) half site 5'-WGWWCW-3' (W = A/T) showed inhibition of prostate specific antigen (PSA) gene expression at 10-fold lower concentrations relative to polyamides bearing unsubstituted β -amino turns, illustrating the ability of turn substitution to potentiate gene regulatory effects.¹⁸

Guided by these results, we explored modifications of the Py-Im polyamide hairpin γ -aminobutyric acid turn unit as a means to enhance uptake and biological activity. Our strategy applied cytotoxicity analysis at an extended time point as an effective assay for identifying biologically active polyamides from small focused libraries. This led to the discovery that for a hairpin polyamide targeting the DNA sequence 5'-WGWWCW-3' (W = A/T), simple conversion of the β -amino turn to a β -benzamide turn led to a 100-fold increase in gene regulatory activity. To better understand the mechanism of this phenomenon, fluorescent analogues of β -amino and β -aryl polyamides were synthesized and analyzed by confocal microscopy and quantitative flow cytometry. Our results suggest the enhanced potency of β -aryl polyamides is due to increased nuclear uptake of these compounds.

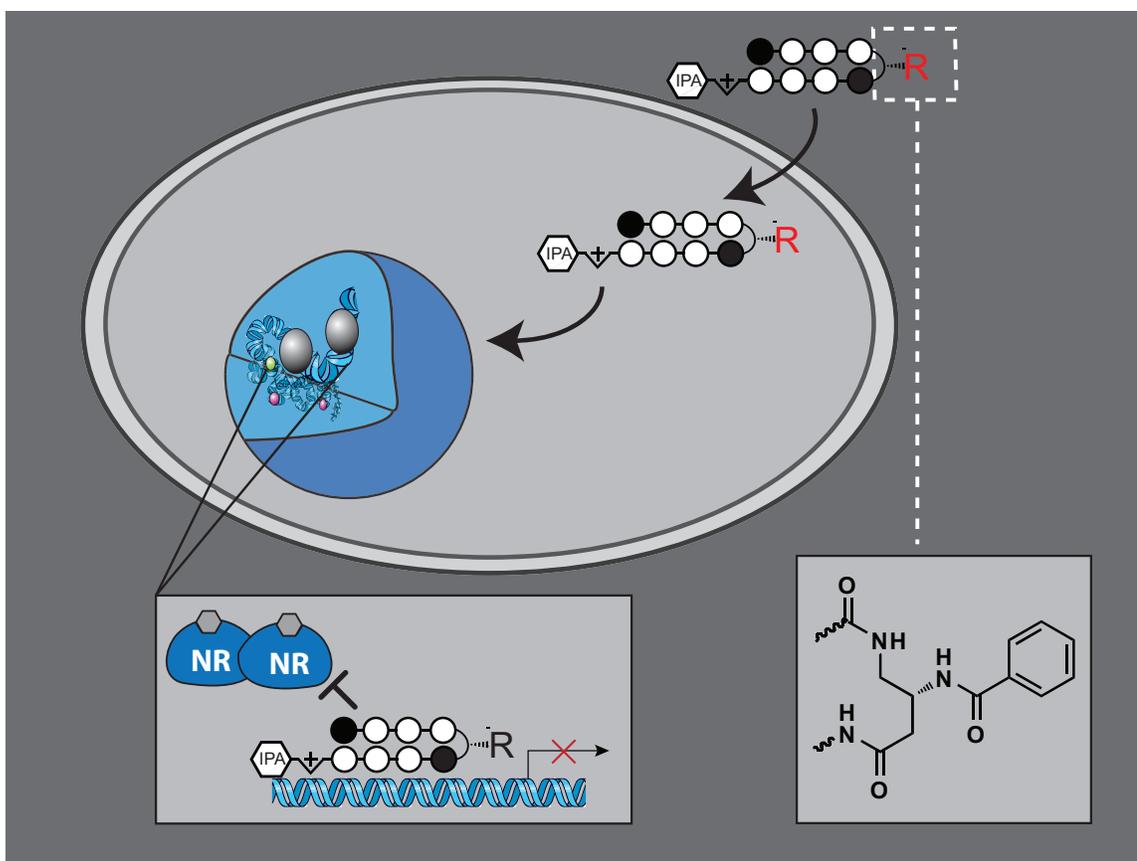


Figure 2.2: Based on a simple structural modification, pyrrole-imidazole polyamides have been developed with increased cellular uptake. This class of compounds can traffic to cell nuclei and access chromatin, imparting gene regulatory effects at lower concentrations.

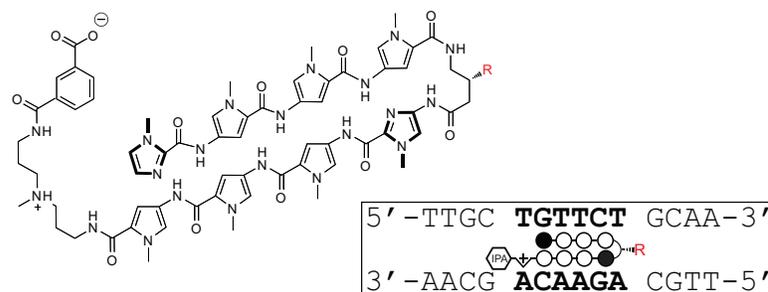
Finally, we explore the generality of this modification and find that aryl-turn substitution capably enhances the uptake of other polyamides cores, but has a variable effect on the upper limits of polyamide nuclear accumulation. Overall this provides a new direction for controlling the intracellular concentration of Py-Im polyamides that will prove essential for future applications in which biological potency is required.

Results and Discussion

Synthesis, DNA-Binding, and Cytotoxicity of β -aryl Polyamides

Following an unanticipated observation that a hairpin polyamide conjugated to an aryl group at the β -amino position showed greater activity in cell culture, we synthesized a small panel of β -aryl substituted polyamides (**4-12**) targeting the sequence 5'-WGWWCW-3' and benchmark their DNA-binding affinities and biological activity against unsubstituted parent (**1**), β -amino (**2**), and β -acetylated (**3**) GABA turns in the A549 human lung carcinoma cell line (Figure 2.3). We used cytotoxicity at 96 hours as a proxy for uptake following the observation that polyamide uptake, gene regulatory activity, and cytotoxicity are often highly correlated.¹³ Thermal denaturation analysis of a DNA duplex of the sequence 5'-TTGCT**GT**TCTGCAA-3' (polyamide match site in bold) shows all polyamides containing a β -amino GABA group (**2-14**) similarly increase the melting temperature by ~ 13 - 15 °C, suggesting no substantial energetic penalty for appendage of the bulky β -aryl groups (Figure 2.3). Cytotoxicity analyses demonstrate that β -amino GABA incorporating polyamide **2** ($IC_{50} = 3.155$ μ M) is considerably more cytotoxic than its unsubstituted counterpart **1** ($IC_{50} > 30$ μ M). While this trend mirrors the relative duplex stabilization of these molecules (ΔT_m **1** = 8.8 °C; ΔT_m **2** = 13.3 °C), simple *N*-acetylation of the β -amino turn (**3**) results in another order of magnitude increase in growth inhibition while not greatly affecting binding affinity (ΔT_m **3** = 13.2 °C). Replacement of the acetyl unit of **3** with a benzoyl functionality (**4**) results in approximately another order of magnitude increase in cytotoxicity ($IC_{50} = 35$ nM), again without concomitant change in the duplex stabilizing ability of this minor groove binder. Within the aryl series (**4-12**) several trends are seen, including increased cytotoxicity of

p-substituted benzoic acids (compare **9** and **10**) and a preference for electron-withdrawing groups at the *p*-position (compare **6** and **7**). Remarkably, significantly increasing the steric bulk of the β -aryl turn, as in polyamides **8** and **11**, does not greatly affect either DNA-binding or cytotoxicity, arguing against interaction of the β -aryl turn with a small pocket of a specific protein partner as a mechanism of cytotoxicity.



R:	IC ₅₀ (nM)	$\Delta T_m / ^\circ\text{C}$	R:	IC ₅₀ (nM)	$\Delta T_m / ^\circ\text{C}$
	> 30,000	8.8 ± 0.2		92 ± 38	14.0 ± 0.8
	3200 ± 900	13.3 ± 0.6		13 ± 5	13.2 ± 0.9
	180 ± 15	13.2 ± 0.1		37 ± 17	13.0 ± 1.1
	35 ± 9	13.0 ± 0.9		200 ± 25	12.2 ± 1.7
	21 ± 5	15.1 ± 0.6		85 ± 13	13.0 ± 0.7
	10 ± 4	12.4 ± 0.6		190 ± 3	13.0 ± 0.5

Figure 2.3: Biological activity and DNA-binding of β -substituted hairpin polyamides. Cytotoxicity analyses were conducted 96 hr after polyamide treatment in the A549 lung carcinoma cell line. IC₅₀ values represent the mean of three biological replicates. ΔT_m denotes the shift in melting temperature following polyamide treatment for the 5'-WGWCW-3' duplex DNA sequence shown.

Since previous studies have noted that polyamide activity can be strongly influenced by cell type,¹² we tested the generality of the increased cytotoxicity of **4** and **9** in LNCaP prostate cancer, MCF-7 breast cancer, and HCT-116 colon cancer cell lines. All three cell lines showed a similar increase in potency for β -aryl compared to β -amino polyamides as was observed in A549 cells (Table 2.1).

Polyamide	cell line			
	A549	LNCaP	HCT-116	MCF-7
2	3155 \pm 895	2550 \pm 91	4680 \pm 954	24400 \pm 9530
4	35 \pm 9	28 \pm 12	106 \pm 18	666 \pm 168
9	37 \pm 17	109 \pm 19	136 \pm 13	470 \pm 77

Table 2.1: Cytotoxicity of β -amino (**2**) and β -aryl (**4**, **9**) polyamides targeting the sequence 5'-WGWWCW-3' towards alternative cancer cell lines. Growth inhibition values represent the mean and standard deviation of three biological replicates.

Suppression of Nuclear-Receptor Mediated Gene Expression by β -Aryl Polyamides

A549 lung epithelial cells have been widely applied as a model for inflammatory gene expression mediated by the transcription factor glucocorticoid receptor (GR).²³⁻²⁵ GR is a member of larger family of nuclear receptors that utilize activation by small molecule ligands in order to affect release from cytoplasmic inhibitory complexes, after which they traffic to the nucleus, multimerize with their cognate protein partners, and activate (or repress) transcription (Figure 2.4A). Hairpin polyamides have been previously shown to inhibit nuclear receptor-DNA interactions in cell culture, making them promising agents for mechanistic studies of nuclear receptor-DNA binding and therapeutic modulation of nuclear receptor activity in diseases such as prostate cancer.

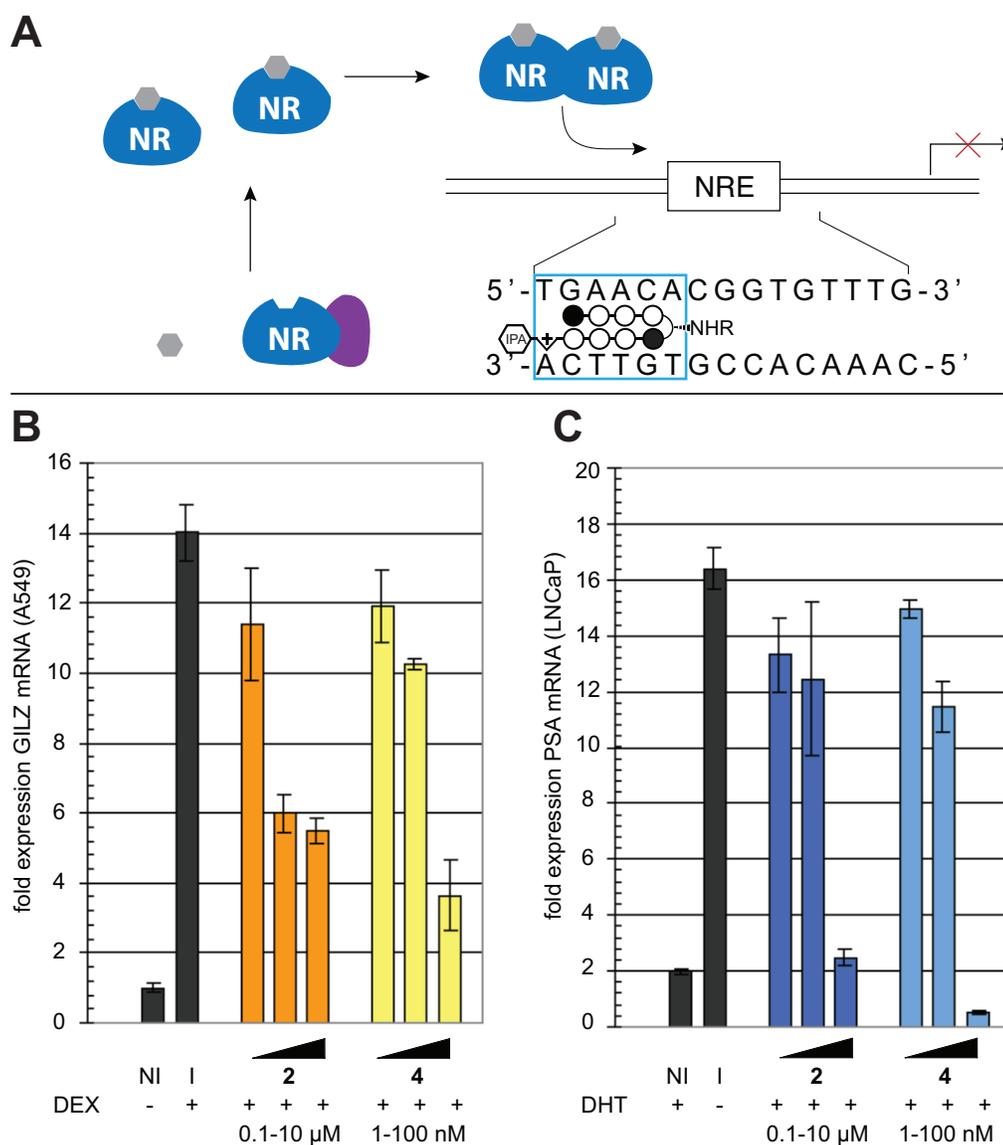


Figure 2.4: Inhibition of nuclear-receptor mediated gene expression by β -aryl polyamides. (a) General schematic of nuclear receptor-mediated gene expression and polyamide inhibition of the protein-DNA interface. (b) Effect of β -turn polyamides **2** and **4** on dexamethasone (DEX)-induced *GILZ* gene expression as measured by qRT-PCR analysis. I = DEX-induced. NI = non-induced. Polyamide **2** concentrations: 100, 1000, 10,000 nM, polyamide **4** concentrations: 1, 10, 100 nM. (c) Effect of β -substituted polyamides **2** and **4** on dihydrotestosterone (DHT)-induced *PSA* gene expression as measured by qRT-PCR analysis. Polyamide **2** concentrations: 100, 1000, 10,000 nM, polyamide **4** concentrations: 1, 10, 100 nM. I = induced. NI = not induced.

Polyamides **1-12** target the 5'-WGWWCW-3' sequence found in the consensus glucocorticoid response element (GRE). Therefore, as an initial test of whether the

increased activity of β -aryl polyamides might be extended to gene regulatory studies, we analyzed the effects of β -aryl polyamide **4** and parent polyamide **2** on endogenous GR-mediated transcription in A549 cells. Following 48 hours treatment with **2**, **4**, or vehicle DMSO, A549 cells were induced with dexamethasone (Dex) before isolation of total RNA and analysis by qRT-PCR. As seen in Figure 2.4B, β -aryl polyamide **4** blunts GR-driven transcription of the canonical GR-regulated gene *GILZ* in a dose-dependent fashion.

Remarkably, inhibition of gene expression by **4** occurs at polyamide concentrations 100x lower than parent compound **2**. Similar results are seen for expression of *FKBP5*, another prototypical GR target (Figure 2.5).

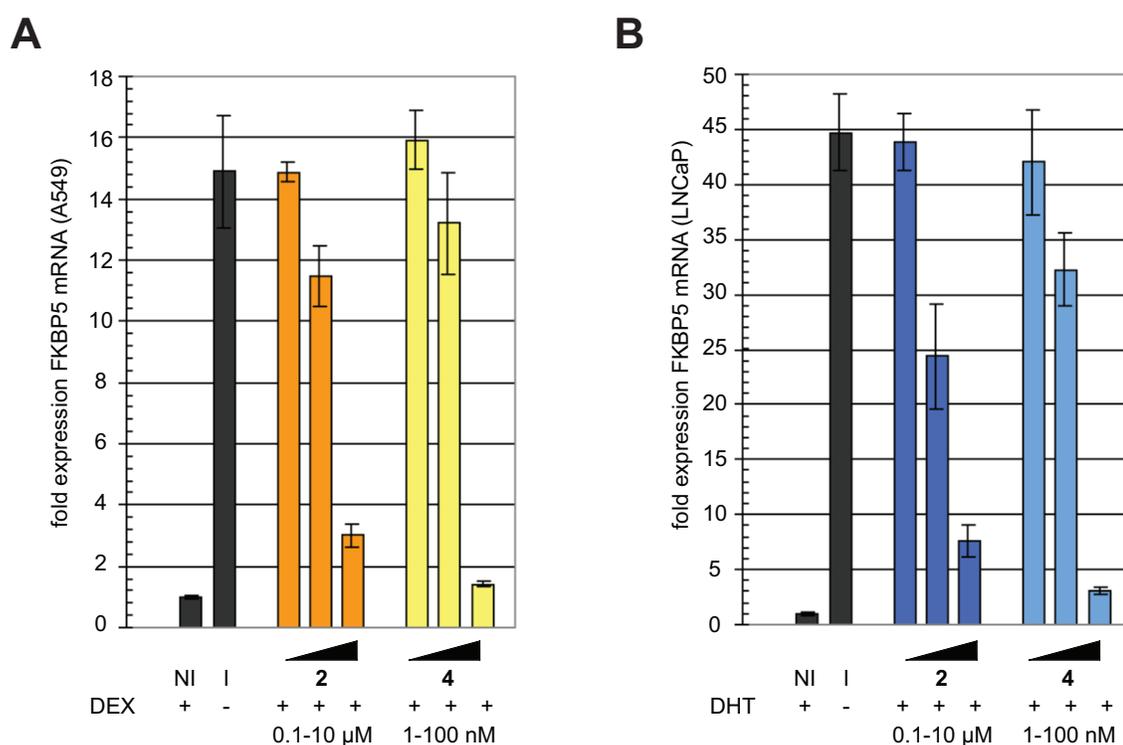


Figure 2.5: Inhibition of nuclear-receptor mediated gene expression by β -aryl polyamides. (a) Effect of β -turn polyamides **2** and **4** on dexamethasone-induced *FKBP5* gene expression as measured by qRT-PCR analysis. I = induced. NI = non-induced. Polyamide **2** concentrations: 100, 1000, 10,000 nM, polyamide **4** concentrations: 1, 10, 100 nM. (b) Effect of β -turn polyamides **2** and **4** on dihydrotestosterone-induced *FKBP5* gene expression as measured by qRT-PCR analysis. Polyamide **2** concentrations: 100, 1000, 10,000 nM, polyamide **4** concentrations: 1, 10, 100 nM. I = induced. NI = not induced.

Notably, these effects are not expected to be due to cytotoxicity, as these experiments utilize a fivefold higher cell plating density than cytotoxicity analyses and are normalized to a housekeeping gene (*GUSB*) that remains stable to polyamide treatment (Figure 2.6).

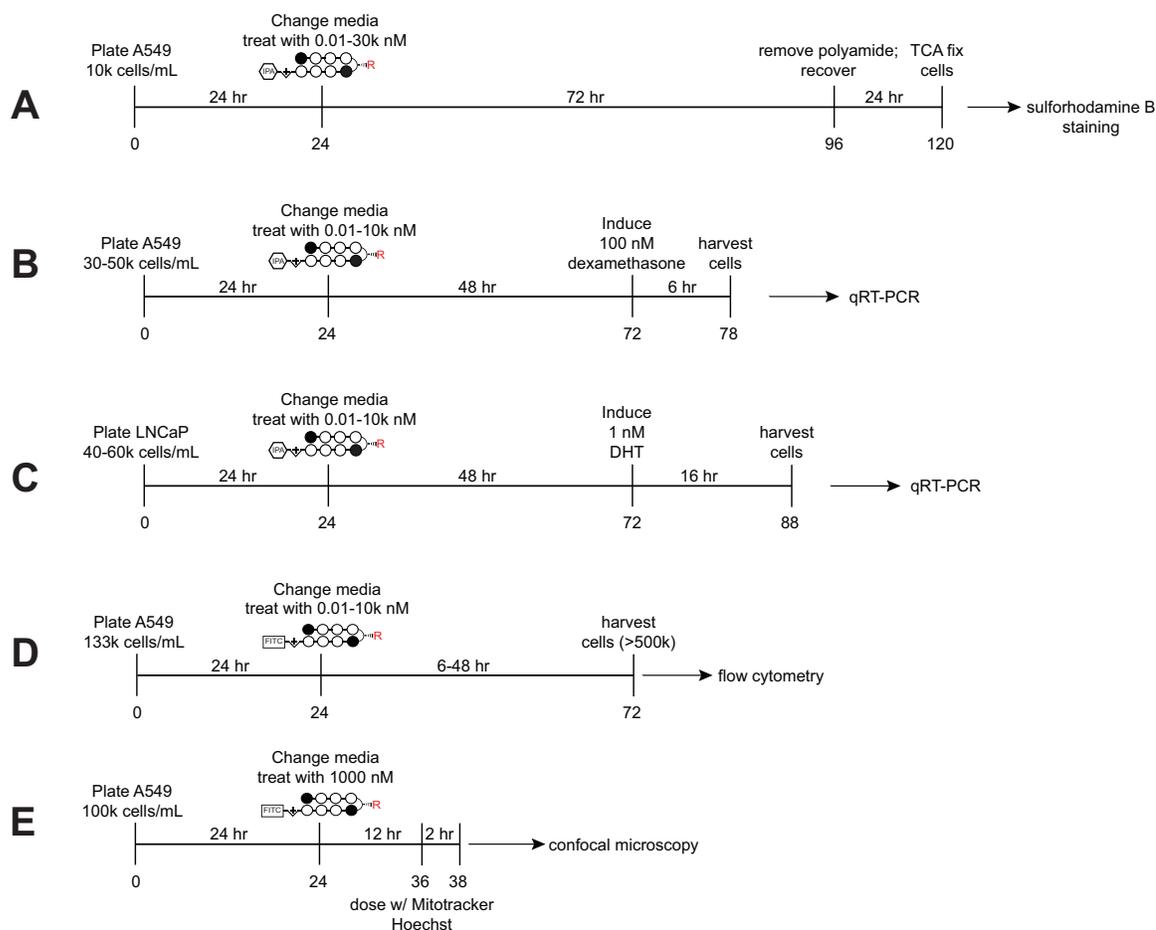


Figure 2.6: Time courses for (a) cytotoxicity analysis, (b) dexamethasone-induced gene expression studies, (c) dihydrotestosterone-induced gene expression studies, (d) flow cytometric analysis of uptake of polyamide-FITC conjugates, (e) confocal microscopy analysis of uptake of polyamide-FITC conjugates.

Time course experiments show substantial inhibition of dexamethasone-induced transcription as early as 12 hours after polyamide treatment, far prior to the onset of cytotoxicity (Figure 2.7).

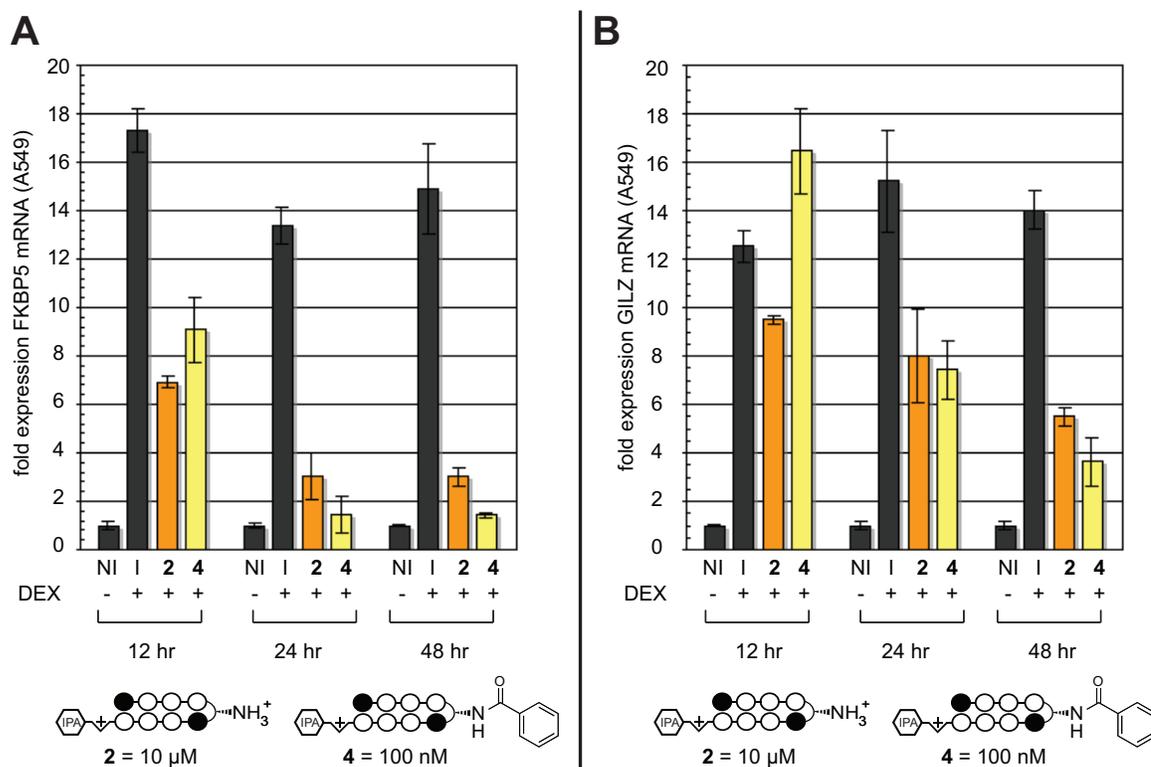


Figure 2.7: Time course analysis of polyamide-mediated inhibition of dexamethasone-induced gene expression. (a) Effect of β -turn polyamides **2** and **4** on dexamethasone (DEX)-induced *GILZ* gene expression as measured by qRT-PCR analysis. I = DEX-induced. NI = non-induced. Polyamide **2** concentration: 10,000 nM, polyamide **4** concentration: 100 nM. These concentrations of **2** and **4** were chosen because they show approximately equivalent inhibition of GR-mediated gene expression at 48 hr. (b) Identical time course analysis of polyamide inhibition of *FKBP5* gene expression over time.

Since the sequence targeted by **4**, 5'-WGWWCW-3', is also found in the androgen response element (ARE), we next tested whether **4** showed similarly enhanced inhibition of androgen receptor (AR) regulated gene expression.⁷ LNCaP prostate cancer cells were exposed to polyamides **2** or **4** for 48 hours prior to induction with the AR-activating ligand 4,5 α -dihydrotestosterone (DHT). Quantitative PCR analysis of reverse-transcribed mRNA shows a drastic decrease in transcription of the known AR target gene prostate specific antigen (*PSA*, also known as *KLK3*) following exposure to **4** (Figure 2.4C). Once again, this inhibition is greater than that observed when parent compound **2** is applied at 100x greater concentrations, reducing PSA mRNA to below basal (non-induced) levels. These results highlight the activity of β -aryl polyamides as potent

antagonists of nuclear receptor-mediated gene expression in living cells.

The Effect of β -Aryl Substitution on Polyamide Uptake by Quantitative Fluorescence Analysis

Hypothetically, the increased biological activity of β -aryl turn polyamides could be attributed to either 1) improved ability to impede protein-DNA interactions, 2) increased uptake, or 3) reduced efflux. The former seemed unlikely, given that our initial structure-activity analysis showed no correlation between steric bulk of the β -aryl turn, which would be expected to affect interaction of groove-binding proteins, and cytotoxicity (Figure 2.3). Therefore, to examine cellular uptake in a systematic and direct fashion we synthesized fluorescent analogues of polyamides **2** and **4** and analyzed their accumulation by confocal microscopy and flow cytometry.^{11,26} Polyamide-FITC conjugates **13** and **14** show similar trends in terms of biological activity compared to parent compounds, although the observed gap in cytotoxicities is decreased from $\sim 100x$ to $\sim 10x$ (Figure 2.8).

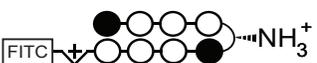
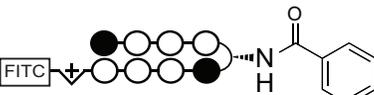
polyamide	IC ₅₀ (nM)
	13 1200 \pm 96
	14 92 \pm 27

Figure 2.8: Relative cytotoxicity of fluorescent analogues **13** and **14** in A549 cells. Cytotoxicities of the parent compounds **2** and **4** are found in Figure 2.3.

Following addition to growth media, polyamides **13** and **14** penetrate the membrane and localize to the nucleus of A549 cells, as verified by co-localization with

the well-known DNA stain Hoechst (Figure 2.9A). However, flow cytometry analysis reveals quantitative differences in the kinetics and degree of uptake. Cells treated with 100 nM β -aryl polyamide **14** demonstrate a rapid increase in fluorescence intensity between 0-12 hours, compared to much slower accumulation of β -amino polyamide **13** dosed under identical conditions (Figure 2.9B).

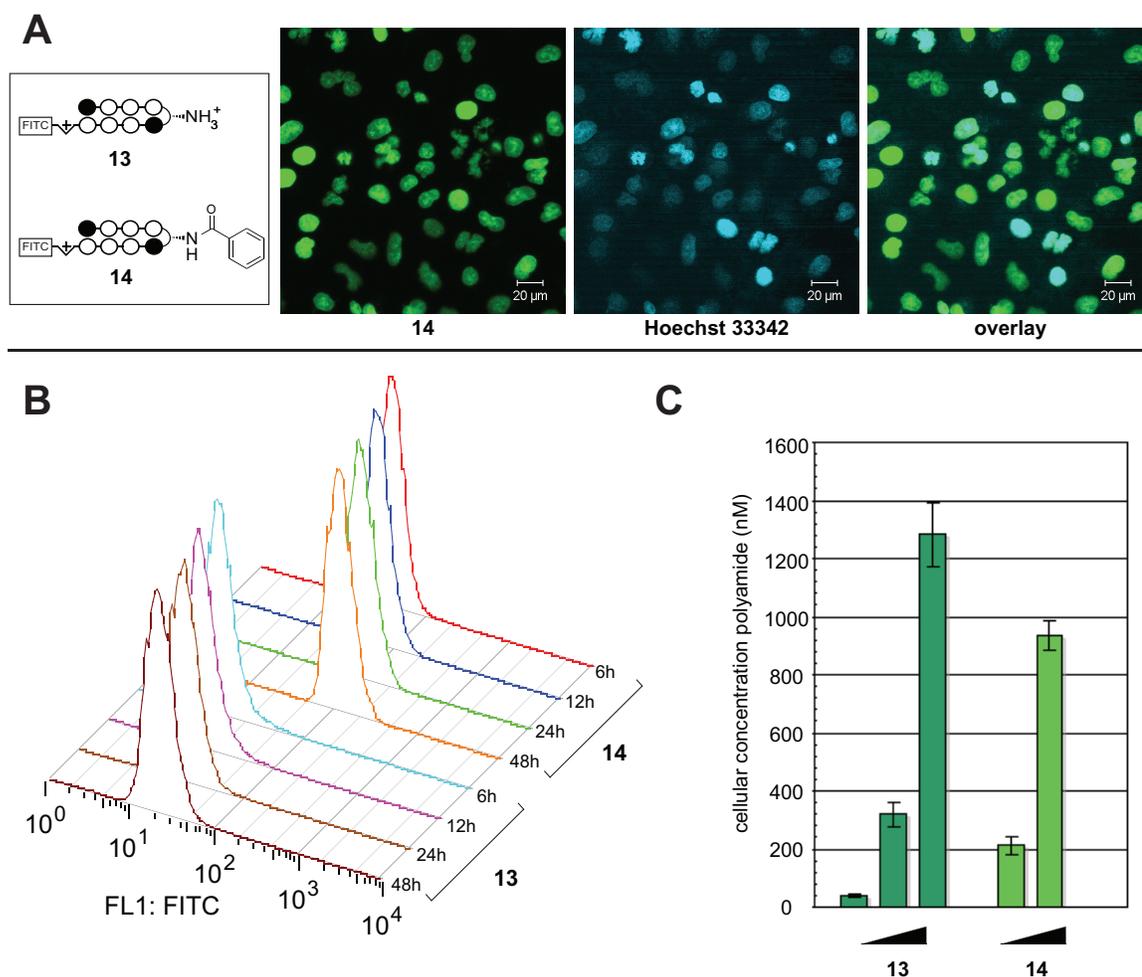


Figure 2.9: Quantitative fluorescence analysis of β -turn substitution on Py-Im polyamide nuclear uptake. (a) Nuclear localization of β -aryl polyamide **13**, as verified by co-localization with Hoechst 33342. (b) Influence of incubation time on fluorescence for A549 cells treated with 100 nM polyamide **13** or **14**. X-axis: relative median fluorescence (FL1: FITC channel); Y-axis: hours of polyamide treatment. (c) Influence of dosage concentration on nuclear accumulation of polyamides. Polyamide **13** concentration: 100 nM, 1000 nM, 10,000 nM. Polyamide **14** concentration: 100 nM, 1000 nM.

Analyzing the overall percentage of fluorescently labelled cells as compared to a DMSO-treated control shows that treatment with 100 nM **14** results in fluorescent

labelling of ~ 88% of A549 cells after 6 hr, while cells treated with 100 nM **13** show labelling of only ~5% of cells over the same time period (Figure 2.10).

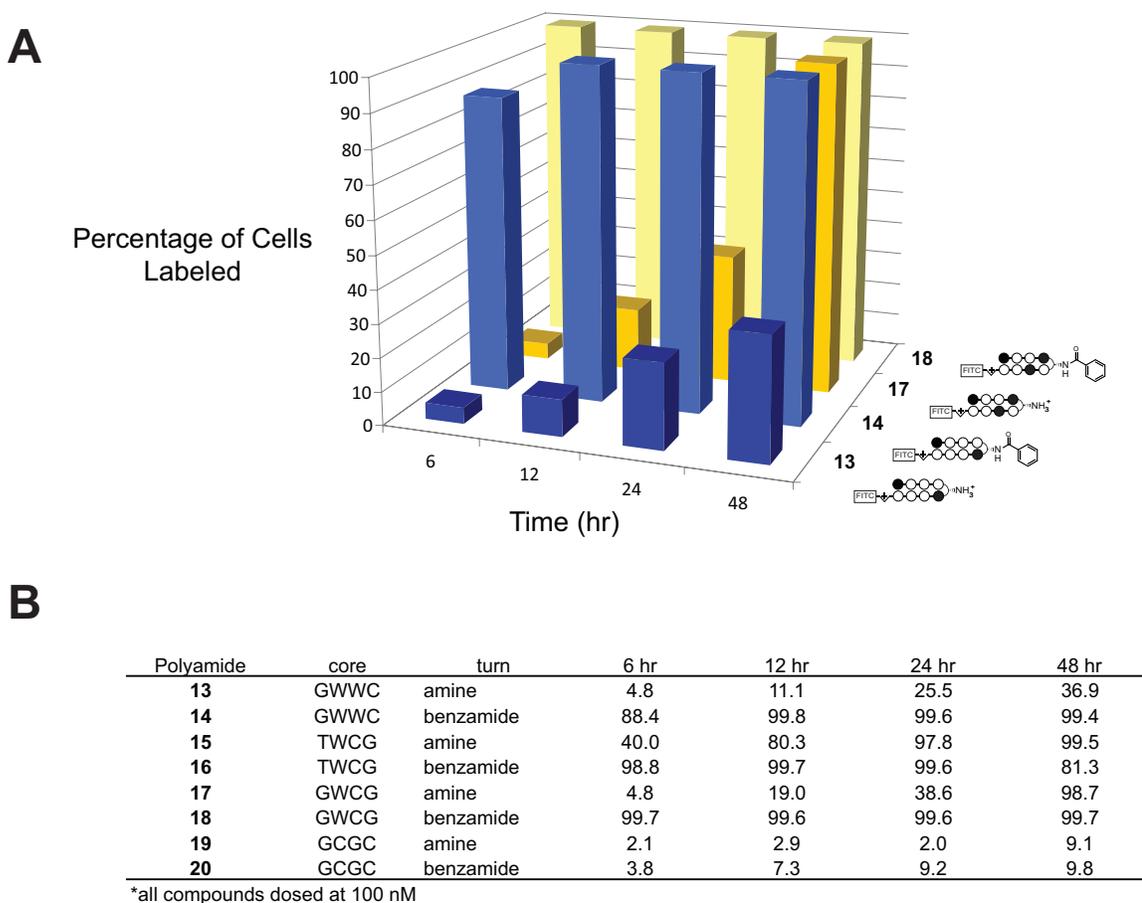


Figure 2.10: Percentage of cells labelled as a function of time exposed to β -turn polyamides β -aryl **13-20** (100 nM). (a) Bar graph depicting relative percentage of cells fluorescently labelled by β -amino polyamides **13** and **17** compared to their β -aryl analogues **14** and **18**. (b) Complete data for all polyamide cores analyzed in this study. Percentage of cells labelled is calculated relative to control cells treated with vehicle DMSO.

In order to gain a more quantitative view of the fluorescence increase, we calculated the nuclear concentrations of fluorescent polyamides **13** and **14** through comparison to fluorescent beads functionalized with known amounts of the FITC fluorophore.^{14,26} Using this methodology, at 48 hours we observe a >4x greater accumulation of **14** than **13** in A549 nuclei when dosed at identical concentrations (100 nM). However, these concentrations can be shifted by increasing polyamide

concentration, as 10x greater dosage concentrations result in a ~3x increase in polyamide concentration values over 48 hours (Figure 2.9C). Notably, this is not due merely due to decreased cell growth, as analysis of cell count and viability prior to flow cytometry revealed no differences between treated samples. Finally, to differentiate uptake and efflux, we compared the effect of verapamil on uptake of β -aryl and β -amino polyamides. Verapamil is an inhibitor of the *p*-glycoprotein transporter that has previously been implicated in cellular efflux of polyamides.²⁷ If β -aryl polyamide **4** is attaining higher concentrations through reduced efflux, verapamil treatment should have little or no effect on intracellular polyamide concentration, whereas if β -aryl polyamide **4** is attaining higher concentrations through enhanced uptake, verapamil will have additive effect on nuclear accumulation. Our results are consistent with the latter mechanism, as we observed similarly higher fluorescent labelling by both **2** and **4** in A549 cells co-treated with a non-toxic (10 μ M) dose of verapamil (Figure 2.11). Overall these findings have two implications: 1) β -aryl turns can significantly increase the rate of polyamide uptake at sub-micromolar concentrations and 2) polyamide cytotoxicity and cellular uptake are well correlated, as the 10x increase in cytotoxicity of **14** relative to **13** is mirrored by its accumulation in cells at 10x lower concentrations.

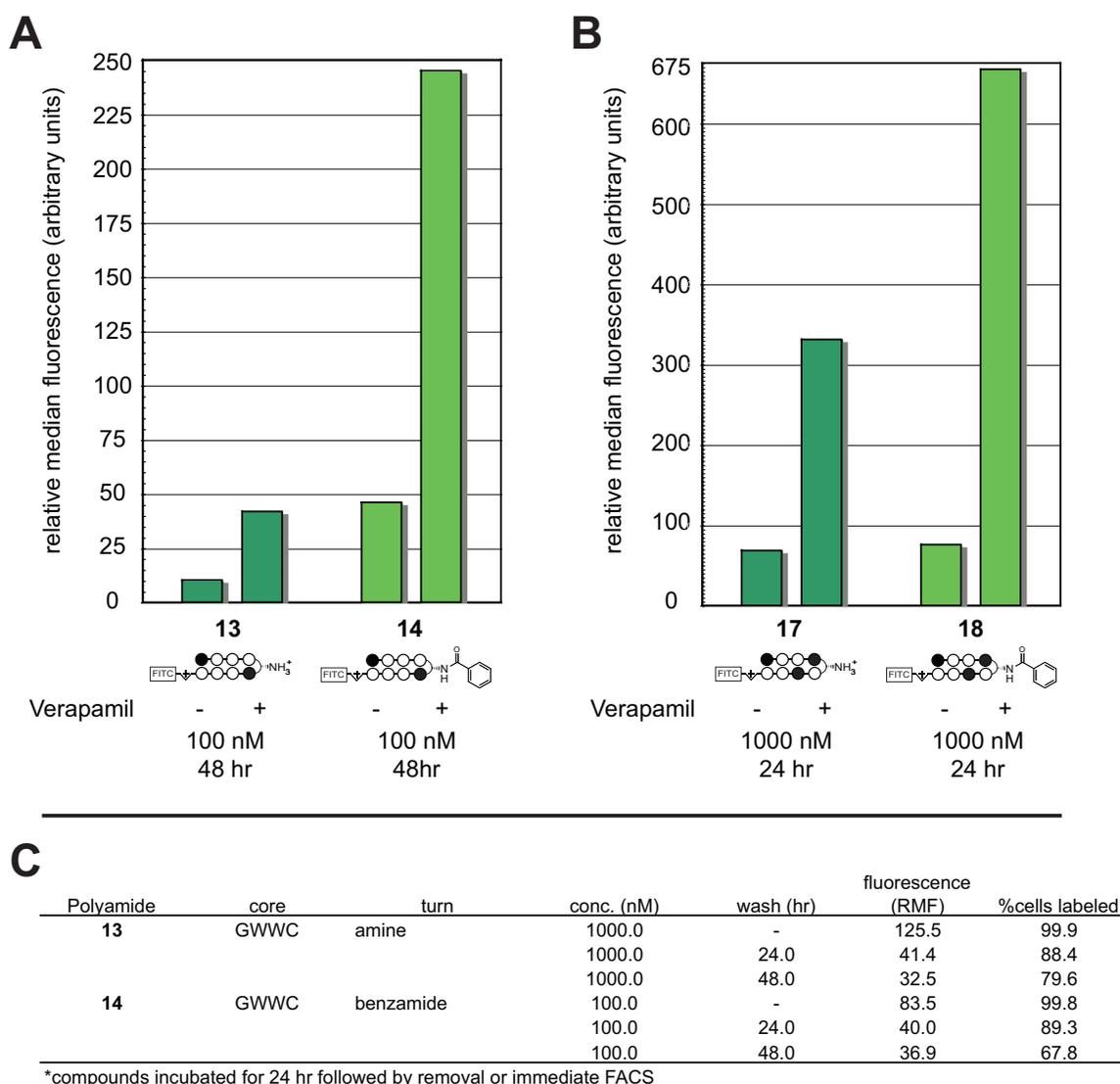


Figure 2.11: Analysis of β -aryl (**14**) and β -amino (**13**) polyamide efflux. (a) Adherent A549 cells were incubated with 100 nM **13** or **14** in the presence or absence of verapamil (10 μ M) for 48 hr and analyzed by flow cytometry. Both polyamides show similar increases in uptake. All samples were analyzed in the same biological experiment. Cellular fluorescence reported as arbitrary units. (a) Adherent A549 cells were incubated with 1000 nM **13** or **14** in the presence or absence of verapamil (10 μ M) for 24 hr and analyzed by flow cytometry. Both polyamides show similar increases in uptake. (c) Washout study of polyamides. A549 cells were treated with **13** or **14** for 24 hr, at which point polyamides were removed and cells were grown in fresh media for 0, 24, or 48 hr followed by FACS analysis. Both compounds show similar profiles, suggesting efflux or dilution by cell growth proceeds similarly for each.

Exploring the Utility of β -Aryl Substitution on Alternative Polyamide Cores

Finally, we examined the ability of β -aryl turns to influence the uptake of polyamide cores targeting alternative sequence motifs. We synthesized fluorescent β -amino and β -benzamide polyamides targeting the sequences 5'-WTWCGW-3' (**15-16**),

5'-WGWCGW-3' (**17-18**), and 5'-WGCGCW-3' (**19-20**). Compound **15** is a high affinity binder of the 5'-ATACGT-3' sequence found within the hypoxia response element (HRE) of the VEGF enhancer,⁶ while compounds **17-20** probe the ability of β -aryl turns to facilitate the uptake of polyamides with increased N-methylimidazole content, a known negative determinant of polyamide nuclear localization.^{12,14} Next, uptake was verified by confocal microscopy (Figure 2.12).

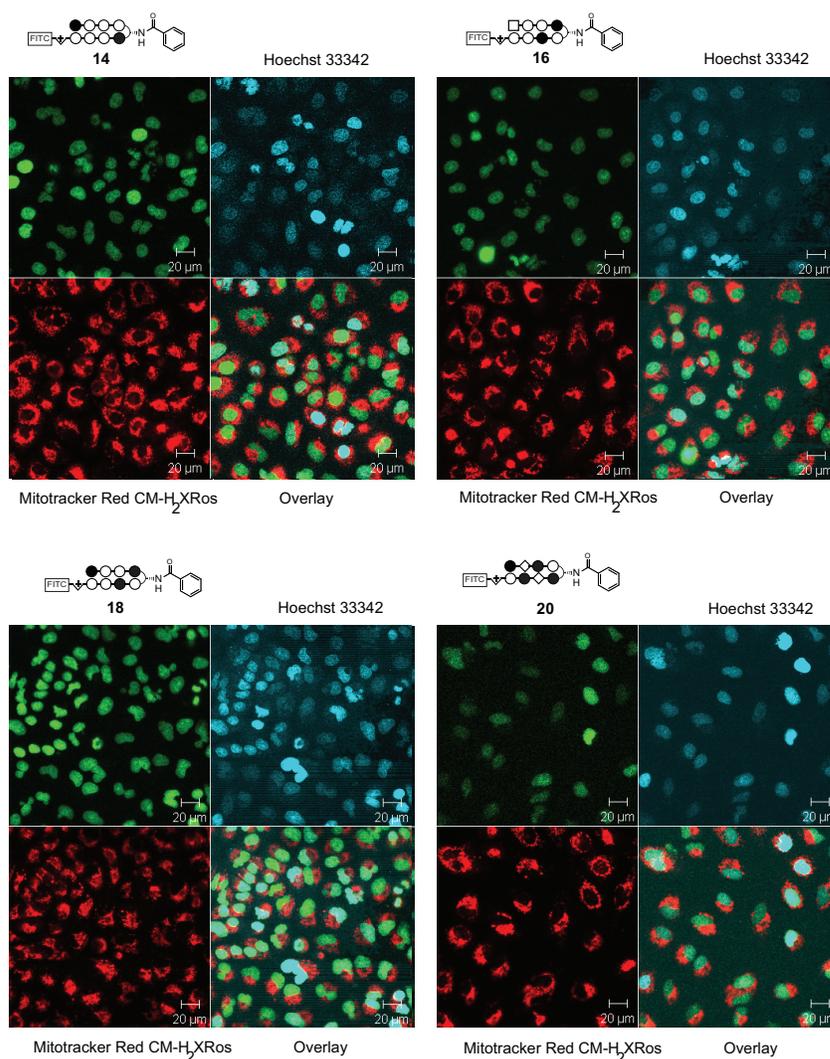


Figure 2.12: Confocal microscopy analysis of β -aryl polyamide nuclear localization. Adherent A549 cells were treated with 1 μ M β -aryl polyamide-FITC conjugates (**14/16/18/20**) for 16 hr, and counterstained with Hoechst 33342 (nuclear stain) and Mitotracker Red CM-H₂XRos (mitochondrial stain) just prior to imaging. Top left: polyamide-FITC ($\lambda_{\text{ex}} = 488$ nm $\lambda_{\text{em}} = 505$ -530 nm). Top right: Hoechst ($\lambda_{\text{ex}} = 750$ nm [2-photon] $\lambda_{\text{em}} = 390$ -465 nm). Bottom left: Mitotracker ($\lambda_{\text{ex}} = 543$ nm $\lambda_{\text{em}} = 565$ -615 nm). Bottom right: three-color overlay.

Each compound was further analyzed for time and concentration-dependent uptake by quantitative flow cytometry (Figures 2.13 and 2.14). When added to media at 100 nM, time-course experiments demonstrate that polyamide-FITC conjugates **15-18** rapidly accumulate in A549 cells (Figure 2.13). Analysis of Im-rich polyamide **19-20** at 100 nM was less informative, as these compounds required dosing at 1000 nM to label a significant percentage of treated cells (Figure 2.10).

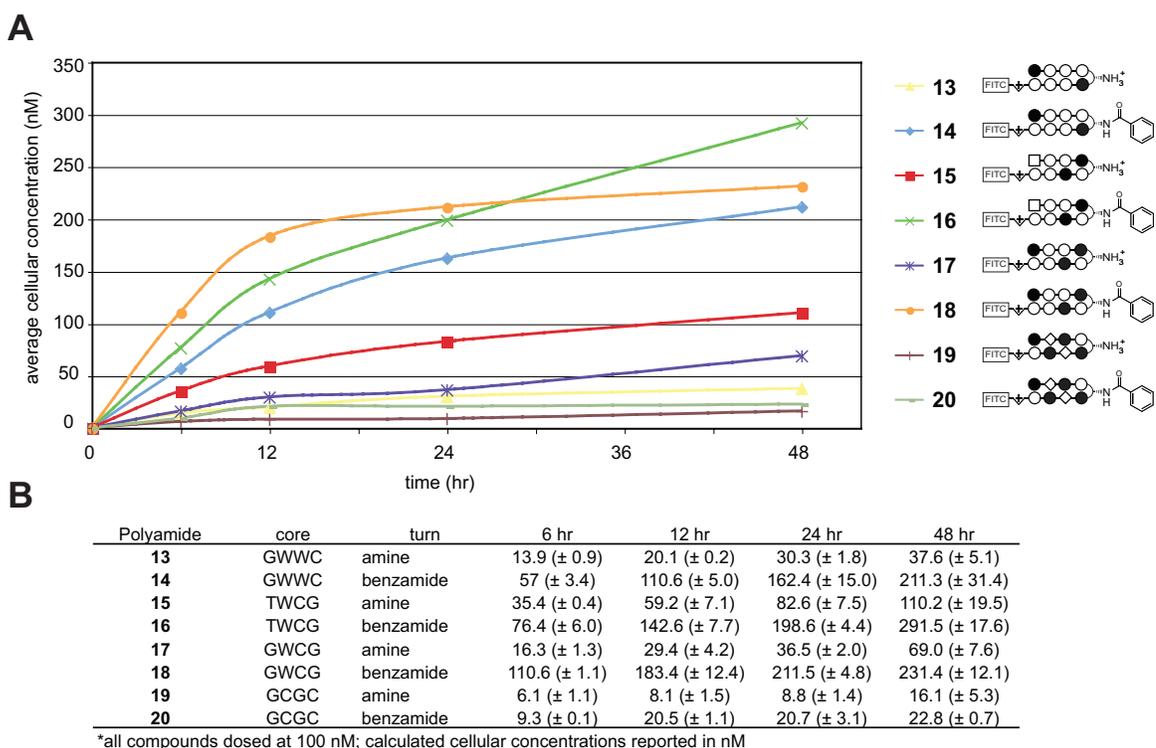


Figure 2.13: Time course analysis of uptake of polyamide-FITC conjugates incorporating diverse DNA sequence-recognition elements. (a) Cellular concentration of polyamide-FITC conjugates **13-20** as a function of time incubated with A549 cells. Polyamide structures are represented as ball and stick models according to the shorthand code: closed circle, Im monomer; open circle, Py monomer; diamond, β -alanine; square, 3-chlorothiophene 2-carboxylic acid. Complete structures can be found in Supplementary Data. Cellular concentration calculated from flow cytometry data as described in materials and methods. (b) Data displayed in tabular form with standard deviations. Core = DNA sequence targeted by the hairpin polyamide core heterocyclic ring pairs. Turn = identity of β -turn modification.

Under these treatment conditions (100 nM **15-18**; 1000 nM **19-20**) each β -aryl polyamide shows increased cellular uptake relative to its β -amino counterpart, demonstrating the general utility of this modification in promoting increased uptake (Figure 2.14).

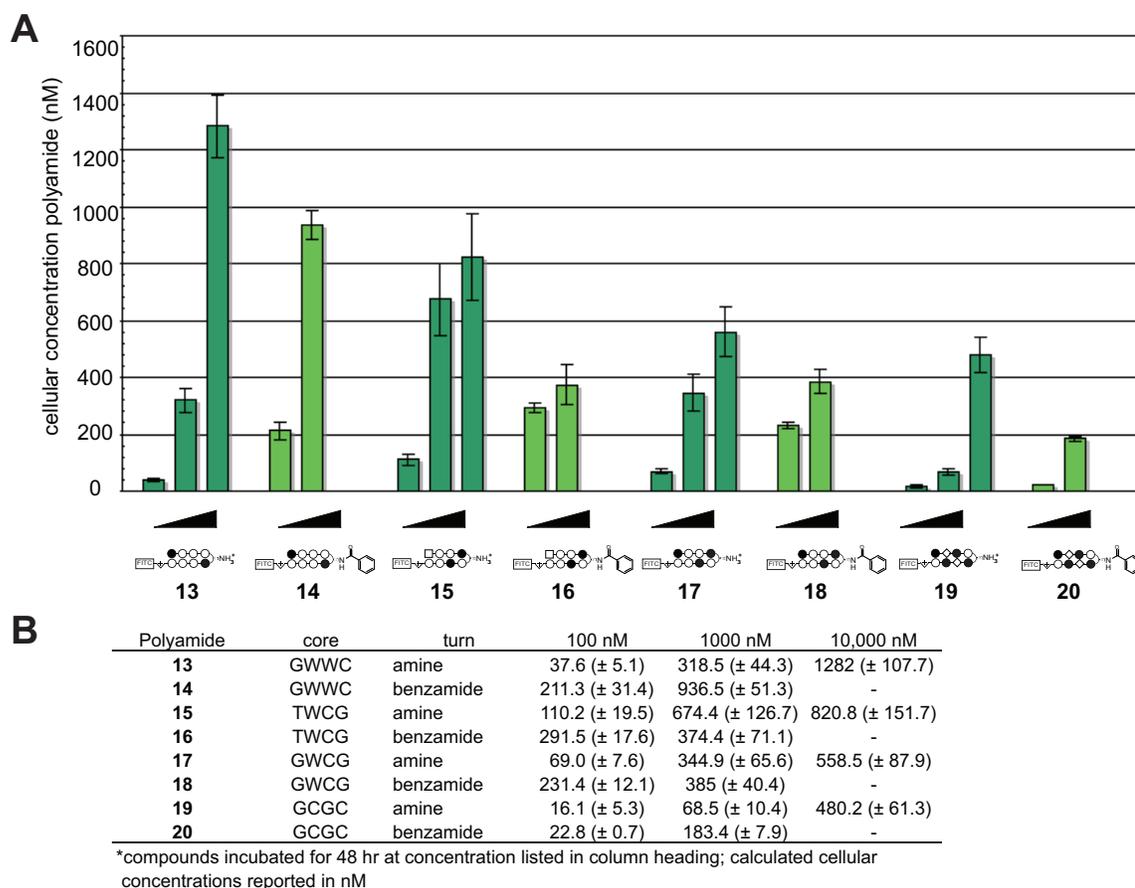


Figure 2.14: Influence of polyamide dosage on cellular concentration at 48 hr. (a) Graphical depiction of relative cellular concentrations of polyamides analyzed in this study. β -amino polyamides **13**, **15**, **17**, and **19** were dosed at 100, 1000, and 10,000 nM respectively (left to right). β -aryl polyamides **14**, **16**, **18**, and **20** were dosed at 100 and 1000 nM respectively (left to right). Attempts to dose β -aryl compounds at higher concentrations were hindered by insolubility. Cellular concentration calculated from flow cytometry data as described in materials and methods. (b) Data displayed in tabular form with standard deviations. Core = DNA sequence targeted by the hairpin polyamide core heterocyclic ring pairs. Turn = identity of β -turn modification.

This enhanced uptake is accompanied by a large increase in the cytotoxicity of non-fluorescent analogues of Im-rich polyamides **18** and **20**, but surprisingly not 5'-WTWCGW-3' targeting polyamide **16** (Figure 2.15).

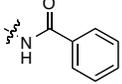
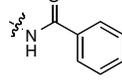
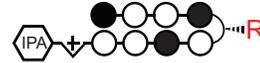
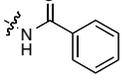
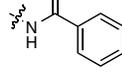
polyamide	R:	IC ₅₀ (nM)
	 2  4	3200 ± 900 35 ± 9
	 21  22	5200 ± 800 >1000
	 23  24	>30,000 580 ± 33
	 25  26	10000 ± 280 530 ± 26

Figure 2.15. Cytotoxicity of β -aryl polyamide cores conjugated to C-terminal isophthalic acid (IPA) tails in A549 cells. Complete structures are given in Scheme S2.4.

Once again flow cytometric analysis proved informative in explaining this unexpected observation, as A549 cells dosed at increasing concentrations (100 nM–10 μ M) of **15-16** show concentration-dependent accumulation of parent polyamide **15**, while β -aryl polyamide **16** shows relatively equivalent uptake at both 100 nM and 1000 nM (Figure 2.14). This is consistent with the small, but constant, inhibitory effect polyamide **15** has on HIF1- α -mediated transcription, as judged by qRT-PCR analysis of VEGF gene expression (Figure 2.16).

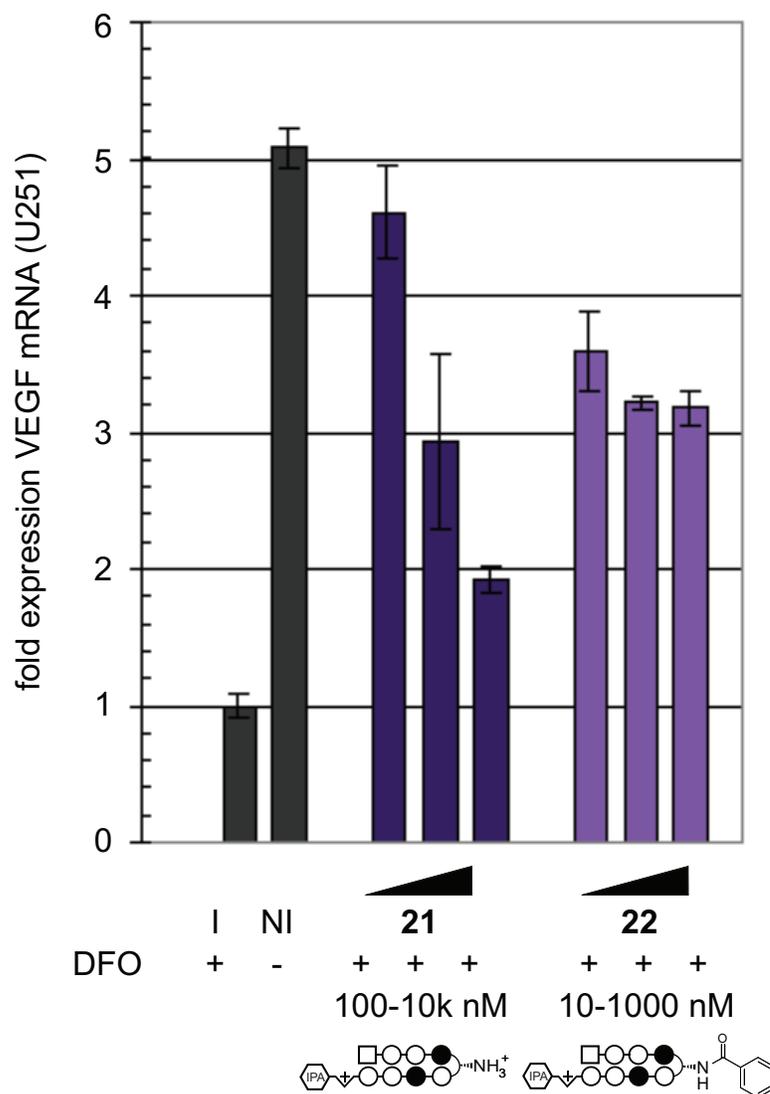


Figure 2.16. Inhibition of hypoxia-induced gene expression by β -turn polyamides targeted to the hypoxia response element (HRE) of the vascular endothelial growth factor (*VEGF*) locus. The effect of β -substituted polyamides **21** and **22** on desferoxamine (DFO)-induced *VEGF* gene expression was analyzed by qRT-PCR analysis. I = DFO induced. NI = non-induced. Polyamide **21** concentrations: 100, 1000, 10,000 nM. Polyamide **22** concentrations: 10, 100, 1000 nM.

Therefore, while our studies show β -aryl turns provide a generally applicable approach to increase the uptake of polyamides at reduced concentrations, this modification can have a variable consequence on the upper limits of polyamide nuclear accumulation that may be required for biological effects.

Conclusions

Chemical approaches for controlling gene expression at the protein-DNA interface require efficient nuclear delivery of gene regulatory agents. Thus far, comprehensive efforts to define structure-uptake relationships for hairpin polyamides in cell culture have largely focused on the optimization of the C-terminus. Insights from these studies are reflected in our use of polyamides modified with a C-terminal isophthalic acid modification for gene regulation and cytotoxicity studies (**1-12**, **21-26**), with the higher molecular weight FITC reserved for direct analysis of polyamide uptake (**14-20**).^{12,15} Here we explore integration of an additional uptake determinant at the hairpin polyamide β -aminobutyric acid turn. Introduction of functionality at the polyamide turn position has previously been most thoroughly explored in the design of covalent sequence-selective DNA alkylating agents.^{28,29} Our findings here represent an initial inquiry into the effect of turn modification on noncovalent sequence-selective DNA-binding agents, and have led to the discovery of a polyamide (**4**) that exhibits excellent affinity for DNA and nanomolar inhibition of dexamethasone and DHT-induced gene expression in human cancer cells. This molecule represents one of the most biologically potent members of this compound-class (eight-ring hairpin polyamides) identified to date.

The second phase of this study examined the mechanism and generality of β -aryl turn modification as a vehicle for increasing polyamide potency using fluorescent polyamide conjugates and quantitative flow cytometry analysis. The power of this approach lies in its relatively simple calibration and ability to sample a large number of cells for any given condition. Our findings indicate the β -aryl turn of **14** aids polyamide

uptake, allowing rapid permeation and nuclear accumulation as compared to β -amino-modified **13**. When extended to hairpin polyamides with alternative heterocycle composition and DNA-binding preferences (**15-20**), uptake of polyamide-FITC conjugates was well-correlated with cytotoxicity and gene regulatory effects, implicating membrane permeability as a primary determinant of the biological activity for this class of molecules. However, while informative, these methods are not without caveats. First, the calculated nuclear concentrations cannot be taken as absolute values, as they do not take into account the known fluorescence enhancements exhibited by polyamide-FITC conjugates upon binding to DNA,³⁰ subcellular localization, fluorescein photostability, or differences in the optical properties of cells as compared to fluorochrome-coated beads. Second, fluorescein modification can significantly alter the biological properties of hairpin polyamides (compare the IC_{50} of **4** with **14**; Figure 2.8). Integration of a low molecular weight reporter into the hairpin polyamide scaffold is therefore attractive from the standpoint of streamlining activity and uptake assays. This approach may benefit from recent incorporation of click chemistry methods into many flow cytometry workflows.^{31,32}

It is interesting to speculate as to the mechanism by which β -aryl turns expedite polyamide uptake. Studies of oligonucleotide-based therapies have shown that modification of these agents with highly lipophilic moieties, such as cholesterol, can facilitate association with the cell membrane and endocytosis.^{33,34} It is possible a similar effect mediates the delivery of β -aryl polyamides. Also interesting is the lack of nuclear accumulation of β -aryl polyamide **16** at higher concentrations. The finding that β -aryl modification promotes increased uptake at 100 nM but not 1000 nM suggests this may

result from a physical phenomena such as insolubility (due to aggregation) at the higher concentration. In general, β -aryl polyamides show decreased solubility relative to β -amino compounds owing to their reduced charge at physiological pH.

Other than delivery, perhaps the most important future challenge lies in developing new methods to define the concentration-dependent effects of Py-Im polyamides on gene expression in living cells. Analytical techniques such as MPE footprinting and affinity cleavage have proven essential to the design of sequence-selective DNA-binding agents,^{35,36} and reveal dose-dependent binding patterns that can be used to directly guide applications *in vitro*.^{37,38} Quantitative fluorescence analysis of polyamide uptake, as performed here, combined with recently developed high-throughput sequencing strategies for analysis of protein-DNA binding^{39,40} and gene expression³⁹ represent promising approaches to similarly footprint polyamide-induced perturbations and binding events *in vivo*, and thereby define the relationship between nuclear concentration and gene regulatory effects.

Overall, these findings highlight hairpin turn modification as a promising new strategy for intracellular delivery of Py-Im polyamides. In terms of applications, the increased potency of these analogues should prove immediately useful for testing in animal models, where the ability to work at lower concentrations will help overcome technical challenges of polyamide solubility and formulation. The amenability of the β -aryl turn to substitution also raises the possibility of using it as a selective handle to optimize polyamide pharmacokinetic properties, such as plasma protein binding, through attachment of pendant chemical functionalities. Finally, β -aryl turns may prove useful for the delivery of molecular probes using polyamides as tethered DNA-binding domains

into living cells. The relevance of such strategies is highlighted by the numerous studies which have used polyamides to target the activity of alkylating agents,²⁸ chromatin remodelling enzymes,⁴¹ and transcriptional activation domains⁴²⁻⁴³ to subsets of genomic loci. Future work will focus on characterizing the mechanism of β -aryl polyamide uptake and applying this technology to the manipulation of protein-DNA interactions in living systems.

Materials and Methods

Reagents and Equipment

Anhydrous N,N-dimethylformamide (DMF), diisopropylethylamine (DIEA), triethylsilane (Et₃SiH), trifluoroacetic acid (CF₃CO₂H [peptide synthesis grade]), isophthalic acid, and (±)-verapamil hydrochloride were purchased from Sigma-Aldrich. Fluorescein isothiocyanate (FITC, Isomer I) was purchased from Invitrogen. Kaiser oxime resin (LL, 200-400 mesh) and benzotriazole-1-yl-oxy-trispyrrolidinophosphonium hexafluorophosphate (PyBOP) were from Novabiochem. N-β-Cbz-N-γ-Boc-D-3,4-diaminobutyric acid [Z-D-β-Dab(Boc)-OH] was purchased from Sigma Aldrich (product code 28206). All Boc-protected *N*-methylpyrrole and *N*-methylimidazole monomers and dimers for polyamide synthesis were prepared according to the published protocols.⁴⁴⁻⁴⁵ Bulk grade solvents were purchased from Fisher Scientific. Centrifugation was performed in a Beckman Coulter bench-top centrifuge (Allegra 21R) equipped with a Beckman swing-out rotor (model S4180). Preparative HPLC purification was performed on an Agilent 1200 Series instrument equipped with a Phenomenex Gemini preparative column (250 x 21.2 mm, 5 μm) with the mobile phase consisting of a gradient of acetonitrile (MeCN) in 0.1% CF₃CO₂H (aqueous). Polyamide synthesis was monitored by analytical HPLC, with analysis 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 consisting of a gradient of MeCN in 0.1% CF₃CO₂H (aqueous). Polyamide concentrations were measured by UV analysis on a Hewlett-Packard model 8453 diode array spectrophotometer in distilled and deionized water (ddH₂O) with a molar extinction coefficient (ϵ) of 69,500 M⁻¹cm⁻¹ at λ_{max} of 310 nm.

Synthesis of Polyamides

All polyamide cores were synthesized by manual solid-phase synthesis on Kaiser oxime resin (Novabiochem) according to the previously published protocol.¹⁹ Polyamides were cleaved from resin by aminolysis with 3,3'-diamino-N-methyldipropylamine for 3 hr at 55°C. Repeated cycles of precipitation and washing with diethyl ether were used to remove excess 3,3'-diamino-N-methyldipropylamine from polyamides prior to purification by reverse phase HPLC.²⁰ The purified polyamide cores were modified at the C-terminal tail position by isophthalic acid (IPA) or fluorescein-5-isothiocyanate (FITC isomer I; Invitrogen) as previously described.¹⁵ Polyamides incorporating the β -Cbz- γ -aminobutyric acid turns were subjected to repeated cycles of precipitation and washing with diethyl ether to remove excess reagents, resuspended in 9:1 CF₃COOH/TFMSA (0.9 mL, 5 min) to remove the benzyl carbamate group,²¹ and purified by HPLC to afford polyamides bearing a β -amino group at the turn. These compounds were coupled to the designated acids by PyBOP and subjected to a final step of HPLC purification to yield β -aryl turn conjugates. Purity of all compounds was verified by analytical HPLC and matrix-assisted, LASER desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry.

General Synthetic Procedure for Conjugation of β -Aryl Turn

A solution of benzoic acid (2.4 mg, 0.02 mmol) and PyBOP (10.4 mg, 0.02 mmol) in DIEA (14 μ L, 0.2 mmol) and DMF (0.4 mL) was stirred at 23°C for 1 hr. Separately, polyamide **2** (800 nmol) was dissolved in DMF (0.1 mL) and slowly added to the pre-activated isophthalic acid solution. The reaction was allowed to stand at 23°C for 1 hr and

monitored by analytical HPLC. Upon completion, Et₂O (45 mL) was added and the reaction was vortexed thoroughly, resulting in formation of a brownish white precipitate that was isolated by centrifugation (~ 4500 rpm).⁴⁶ After removal of the organic layer, the residual solid was dissolved in DMF (0.5 mL) with ultrasonication, and diluted with 20% MeCN in 0.1% CF₃COOH (4.5 mL). Purification by reverse-phase HPLC followed by lyophilization provided β-aryl polyamide **4** (0.082 mmol, 91%). MS (MALDI-TOF) calc'd for C₇₂H₈₁N₂₂O₁₃ [M+H]⁺ 1461.6, found 1461.9.

Polyamide Characterization Data

- 1** MS (MALDI-TOF) calc'd for C₆₅H₇₆N₂₁O₁₂ [M+H]⁺ 1342.6, found 1589.8.
- 2** MS (MALDI-TOF) calc'd for C₆₅H₇₇N₂₂O₁₂ [M+H]⁺ 1357.6, found 1357.8.
- 3** MS (MALDI-TOF) calc'd for C₆₇H₇₉N₂₂O₁₃ [M+H]⁺ 1399.6, found 1399.9.
- 4** MS (MALDI-TOF) calc'd for C₇₂H₈₁N₂₂O₁₂ [M+H]⁺ 1461.6, found 1461.9.
- 5** MS (MALDI-TOF) calc'd for C₇₁H₈₀N₂₃O₁₃ [M+H]⁺ 1462.6, found 1462.7.
- 6** MS (MALDI-TOF) calc'd for C₇₂H₈₀N₂₃O₁₅ [M+H]⁺ 1506.6, found 1506.5.
- 7** MS (MALDI-TOF) calc'd for C₇₂H₈₂N₂₃O₁₃ [M+H]⁺ 1476.6, found 1476.2.
- 8** MS (MALDI-TOF) calc'd for C₇₈H₈₅N₂₂O₁₃ [M+H]⁺ 1537.7, found 1537.5.
- 9** MS (MALDI-TOF) calc'd for C₇₃H₈₁N₂₂O₁₅ [M+H]⁺ 1505.6, found 1505.8.
- 10** MS (MALDI-TOF) calc'd for C₇₃H₈₁N₂₂O₁₅ [M+H]⁺ 1505.6, found 1505.8.
- 11** MS (MALDI-TOF) calc'd for C₇₈H₉₃N₂₄O₁₄ [M+H]⁺ 1589.7, found 1589.8.
- 12** MS (MALDI-TOF) calc'd for C₇₈H₉₃N₂₄O₁₄ [M+H]⁺ 1589.7, found 1589.8.
- 13** MS (MALDI-TOF) calc'd for C₇₈H₈₃N₂₃O₁₄S [M+H]⁺ 1597.6, found 1598.0.
- 14** MS (MALDI-TOF) calc'd for C₈₅H₈₇N₂₃O₁₅S [M+H]⁺ 1701.6, found 1701.8.
- 15** MS (MALDI-TOF) calc'd for C₇₇H₈₀N₂₂O₁₄S₂Cl₁ [M+H]⁺ 1635.5, found 1635.6.
- 16** MS (MALDI-TOF) calc'd for C₈₄H₈₃N₂₂O₁₅S₂Cl₁ [M+H]⁺ 1738.6, found 1738.3.
- 17** MS (MALDI-TOF) calc'd for C₇₇H₈₀N₂₄O₁₄S₁ [M+H]⁺ 1596.6, found 1596.8.
- 18** MS (MALDI-TOF) calc'd for C₈₄H₈₆N₂₄O₁₅S₁ [M+H]⁺ 1702.6, found 1702.8.
- 19** MS (MALDI-TOF) calc'd for C₇₀H₈₀N₂₃O₁₄S₁ [M+H]⁺ 1498.6, found 1498.9.

20 MS (MALDI-TOF) calc'd for $C_{77}H_{83}N_{23}O_{15}S_1$ $[M+H]^+$ 1601.6, found 1601.7.

21 MS (MALDI-TOF) calc'd for $C_{64}H_{73}N_{21}O_{12}S_1Cl_1$ $[M+H]^+$ 1394.5, found 1394.6.

22 MS (MALDI-TOF) calc'd for $C_{71}H_{76}N_{21}O_{13}S_1Cl_1$ $[M+H]^+$ 1497.5, found 1497.7.

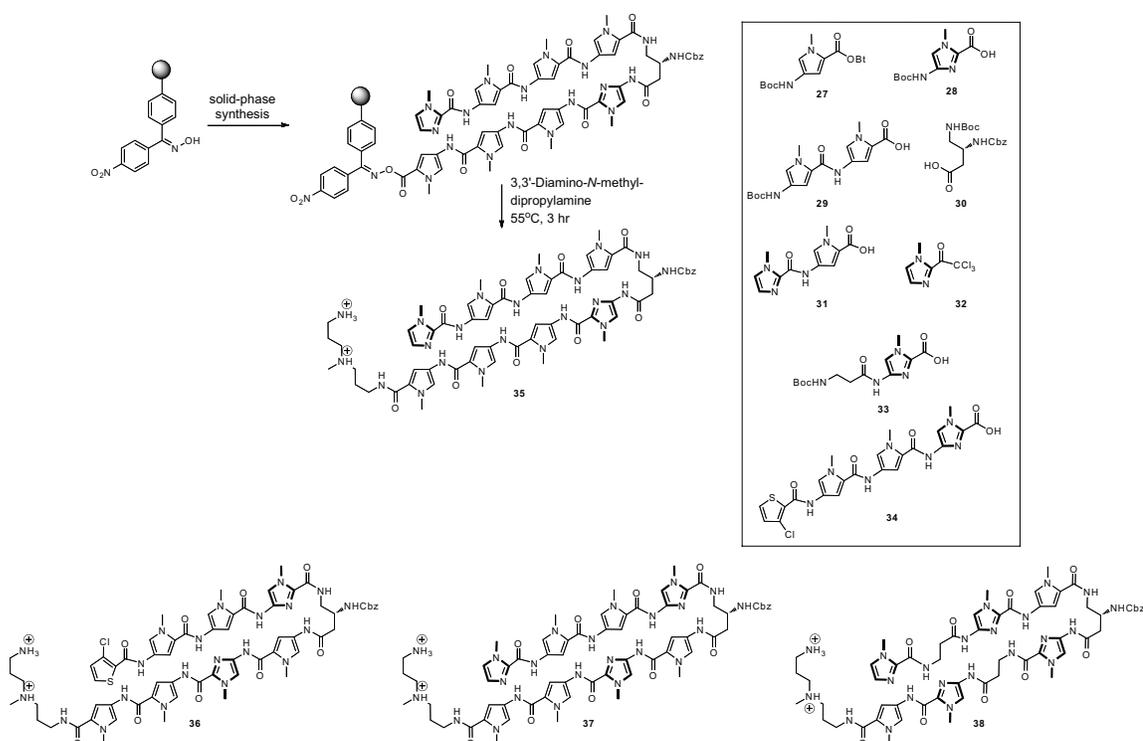
23 MS (MALDI-TOF) calc'd for $C_{64}H_{74}N_{23}O_{12}$ $[M+H]^+$ 1356.6, found 1356.8.

24 MS (MALDI-TOF) calc'd for $C_{71}H_{77}N_{23}O_{13}$ $[M+H]^+$ 1459.6, found 1459.6.

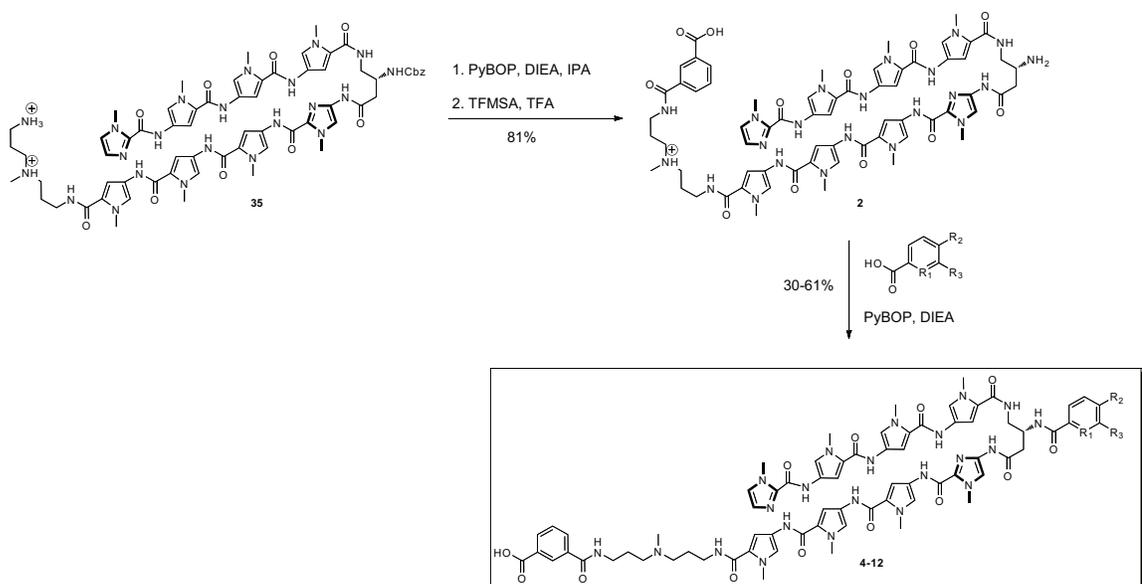
25 MS (MALDI-TOF) calc'd for $C_{57}H_{73}N_{22}O_{12}$ $[M+H]^+$ 1257.6, found 1257.5.

26 MS (MALDI-TOF) calc'd for $C_{64}H_{76}N_{22}O_{13}$ $[M+H]^+$ 1360.6, found 1360.8.

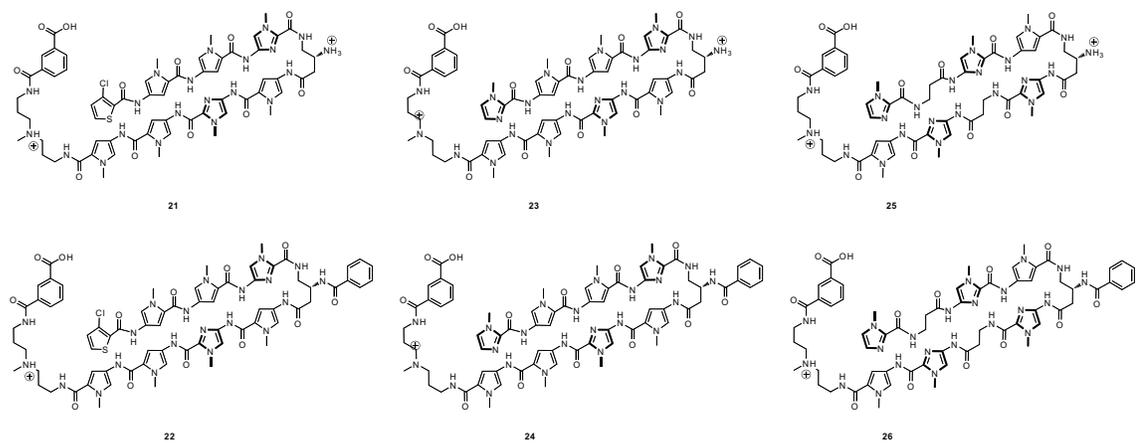
Synthetic Schemes



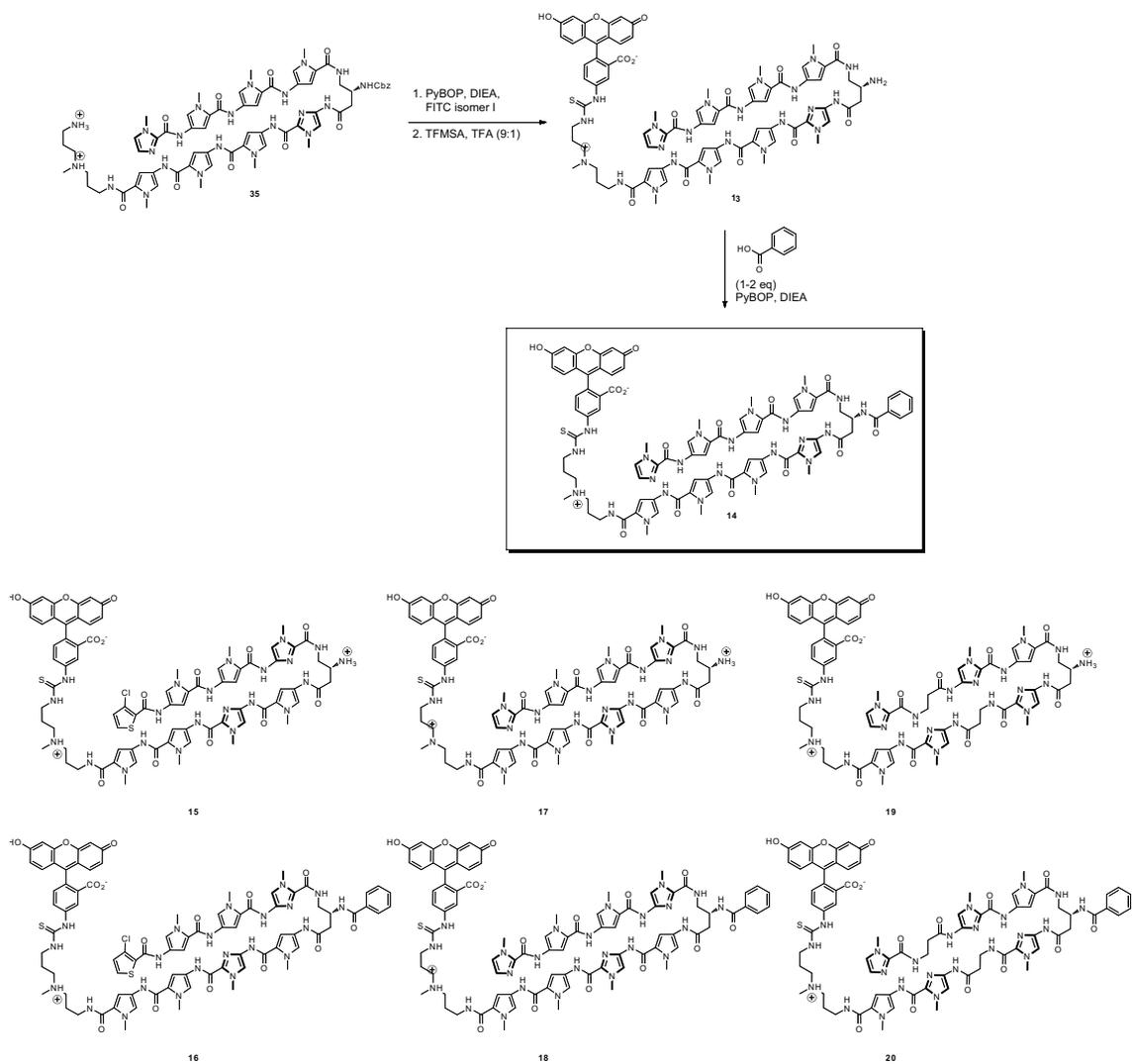
Scheme S2.1: Monomers, dimers, and general scheme for solid-phase synthesis of polyamides analyzed in this study.



Scheme S2.2: Scheme for solution-phase derivatization of polyamides and conjugation of β -aryl turns.



Scheme S2.3: Complete structures of polyamides 21-26.



Scheme S2.4: Scheme and complete chemical structures for fluorescent polyamides 14-20.

Cell Culture

All cell lines were purchased from ATCC (Manassas, VA) and maintained in the following media: A549 cells (F-12K); LNCaP (RPMI 1640); HCT-116 (McCoy's 5a Medium Modified); MCF-7 (Eagle's Minimum Essential Medium). All media were supplemented with 10% FBS and cultured at 37°C under 5% CO₂.

Sulforhodamine B Assay of Polyamide Cytotoxicity

IC₅₀ values for cytotoxicity were determined in 96-well microplates using the sulforhodamine B colorimetric assay for cellular protein content as previously described.²² All polyamide stock solutions were prepared in neat DMSO and dosed to give a final concentration of $\leq 0.3\%$ DMSO. Briefly, cell lines were plated in 100 μ L of the defined media at the following densities: A549 (1000 cells/well); LNCaP (5000 cells/well), HCT-116 (750 cells/well), MCF-7 (3000 cells/well). After 24 hr, polyamides were added to adhered cells in 100 μ L of media by serial dilution. Quadruplicate wells were used for each polyamide concentration. After 72 hr, the medium was replaced with 100 μ L fresh medium, and cells were allowed to recover for 24 hr. Following recovery, cells were fixed with 100 μ L 10% trichloroacetic acid solution, washed, stained, and dried as described. For 48 and 72 hr polyamide treatments, the procedure was followed as above with A549 cells plated at 3000 and 2000 cells per well, respectively. After solubilization of the bound dye in 10 mM Tris (pH 8), the absorbance was measured at 490 nm on a Victor microplate reader (PerkinElmer). The data are charted as a percentage of untreated controls, corrected for background absorbance. IC₅₀ is defined as the concentration that inhibits 50% of control cell growth. These values were determined

by non-linear least-squares regression fit to $Y = A + (B-A)/(1+10^{((\text{Log EC}_{50}-X)*H)}$, where $A=\text{max.}$, $B=\text{min.}$, and $H=\text{Hill Slope}$. All calculations were performed using Prism 4 (GraphPad) software. Three independent trials were averaged; stated IC_{50} values represent the mean and standard deviation.

Thermal Melting Temperature Analysis

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. An 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. Oligonucleotides (0.1 mM stock solutions dissolved in 10 mM Tris-Cl, 0.1 mM EDTA, pH 8.0) were purchased from Integrated DNA Technologies. DNA duplexes and hairpin polyamides were mixed to a final concentration of 2 μM and 2.4 μM , respectively, 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 $\lambda = 260$ nm from 25°C to 90°C with a heating rate of 0.5°C/min. The reported melting temperatures were defined as the maximum of the first derivative of the denaturation profile, and represent the average of four measurements.

Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR) Analysis of Nuclear Receptor-Mediated Gene Expression

To analyze the effects of polyamide-treatment on dexamethasone-induced gene expression, A549 cells were plated in 24-well plates at a density of $15\text{--}25 \times 10^3$ cells per well ($30\text{--}50 \times 10^3$ cells/ml).⁸ After 24 hr the medium was replaced by F-12K containing 10% charcoal-stripped FBS, and polyamides were added to the specified concentrations (1-10,000 nM). Cells were incubated with polyamides for 12, 24, or 48 hr, followed by induction with dexamethasone (100 nM). After 6 hr cells were harvested and mRNA was isolated (RNEasy 96 kit - Qiagen) and reverse transcribed (Transcriptor First Strand cDNA Synthesis kit – Roche). Quantitative real-time PCR was performed with FastStart Universal SYBR Green Master Mix (Roche) on an ABI 7300 qPCR instrument (Applied Biosystems) following the manufacturer's protocol. A similar protocol was utilized to measure 4,5 α -dihydrotestosterone (DHT)-induced gene expression in LNCaP cells, with the following modifications: (1) the initial plating density was $20\text{--}30 \times 10^3$ cells per well ($40\text{--}60 \times 10^3$ cells/ml), (2) cells were incubated with polyamides for 48 hr, (3) cells were induced with DHT (10 nM) for 16 hr.⁷ In both cases, cDNA corresponding to the genes of interest was measured relative to β -glucuronidase as an endogenous control. Primer sequences are provided in the Supplementary Data.

Confocal Microscopy Analysis

For microscopy experiments, A549 cells in F-12K medium were plated into culture dishes equipped with glass bottoms for direct imaging (MatTek) at a density of 20×10^3 cells per dish (100×10^3 cells/mL). Cells were grown in the glass-bottom dishes for 24 hr. Medium was then removed and replaced with 200 μ L of fresh medium supplemented with FITC-labelled polyamides (1 μ M) in DMSO (final concentration 0.1%). Cells were incubated at 37°C for 16 hr, followed by removal of media, gentle washing with 100 μ L PBS, and addition of fresh medium immediately prior to imaging. For colocalization experiments, 15 μ M Hoechst 33342 (0.5 mM stock in ddH₂O) and 1 μ M Mitotracker Red CM-H₂XRos (1 mM stock in DMSO) were added 2 hr prior to imaging. For time-course experiments, cells were imaged directly after polyamide addition using an environmentally-controlled microscopy chamber (37 °C, 5 % CO₂). Imaging was performed at the Caltech Beckman Imaging Center using a Zeiss LSM 510 Meta NLO 2-photon inverted laser scanning microscope equipped with a 40x oil-immersion objective lens. Polyamide–fluorescein conjugates and Mitotracker were imaged in multi-track mode using 488 nm and 543 nm laser excitation with a pinhole of 375 μ m and standard filter sets for fluorescein and rhodamine respectively. Hoechst was imaged using 800 nm two-photon excitation with an HFT KP680 dichroic and a 390-465 nm bandpass filter with a fully open pinhole. All images were analyzed using Zeiss LSM software.

Flow Cytometry Analysis.

For flow cytometry experiments A549 cells were plated in 6-well plates at a density of 500×10^3 cells per well (133×10^3 cells/ml), and allowed to adhere for 16-24 hr before treatment with polyamide-FITC conjugates (100-10,000 nM). Cells were grown for 6, 12, 24, or 48 hr, the medium removed, washed with cold PBS, and trypsinized for 5 min at 37 ° C. The trypsinized cells were combined with the cell culture supernatant and wash solution, and centrifuged for 5 min at 300 x g. This pellet was resuspended, washed with cold PBS, pelleted for 5 min at 300 x g and resuspended in 800 μ L Hank's Balanced Salt Solution (2.5 mg/mL BSA, 10 mM HEPES, pH 7.0, no Mg^{2+} , no Ca^{2+} , no phenol red). Cell viability was checked with trypan blue stain and found to be ≥ 90 -95% in all cases. Live cells were then diluted to a concentration of 5×10^5 cells/mL pipetted through a 40 μ m cell strainer (BD Falcon) into 5 mL polystyrene round-bottom tubes (BD Falcon). Just prior to analysis, cells were stained for viability using 7-amino-actinomycin D (7-AAD; eBioscience). Analyses were performed on a BD Bioscience FACSCalibur instrument at the Caltech Flow Cytometry Cell Sorting Facility using standard filter sets for fluorescein and 7-AAD. SPHERO Rainbow Calibration Particles (6 peaks, 3.0-3.4 μ m; Spherotech) were used as calibration standards. For each condition 10,000 cells were analyzed. Fluorescence values are representative of the relative median fluorescence (RMF) intensity of the main population, gated for viability based on 7-AAD dye exclusion. Comparison of RMFs of polyamide-labelled cells with SPHERO Rainbow Calibration Particles was used to calculate molecules of equivalent fluorescein per cell, which was converted to a nuclear concentration based on modelling the A549 nucleus as a cylinder with radius 10 μ m and height 5 μ m to give a calculated nuclear volume of 1 x

10^{-12} L (estimates based on confocal microscopy). All data were analyzed using FlowJo v8.8.2 (TreeStar) and indicate the average and standard deviation of two trials.

Funding

This work was supported by the National Institutes of Health (grant numbers GM51747 and GM27681). J.L.M was supported by a postdoctoral fellowship from the American Cancer Society (grant number PF-10-015-01-CDD). D.C.M. was supported by a National Institutes of Health Cellular, Biochemical, and Molecular Sciences Predoctoral Research training grant (grant number 5T32GM007616).

Acknowledgements

Mass spectrometry analyses were performed in the Mass Spectrometry Laboratory of the Division of Chemistry and Chemical Engineering at the California Institute of Technology. Flow cytometry analyses were performed at the Caltech Flow Cytometry Cell Sorting Facility. We thank Rochelle Diamond for support and technical expertise in performing flow cytometry experiments, and William Dempsey for helpful discussions regarding confocal microscopy.

References

1. Dervan, P. B., Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, 9, 2215-2235.
2. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, 13, 284-299.
3. Mrksich, M.; Parks, M. E.; Dervan, P. B., Hairpin peptide motif. A new class of oligopeptides for sequence-specific recognition in the minor groove of double-helical DNA. *J. Am. Chem. Soc.* **1994**, 116, 7983-7988.
4. Hsu, C. F., Phillips, J. W., Trauger, J. W., Farkas, M. E., Belitsky, J. M., Heckel, A., Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C.; Dervan, P. B., Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron* **2007**, 63, 6146-6151.
5. Vaquerizas, J. M.; Kummerfeld, S. K.; Teichmann, S. A.; Luscombe, N. M., A census of human transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **2009**, 10, 252-263.
6. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G.; Jr.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. USA* **2004**, 101, 16768-16773.
7. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci. USA* **2007**, 104, 10418-10423.
8. Muzikar, K. A.; Nickols, N. G.; Dervan, P. B., Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, 106, 16598-16603.
9. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.* **2007**, 2, 561-571.
10. Matsuda, H.; Fukuda, N.; Ueno, T.; Tahira, Y.; Ayame, H.; Zhang, W.; Bando, T.; Sugiyama, H.; Saito, S.; Matsumoto, K.; Mugishima H.; Serie, K., Development of gene silencing pyrrole-imidazole polyamide targeting the TGF-beta1 promoter for treatment of progressive renal diseases. *J. Am. Soc. Nephrol.* **2006**, 17, 422-432.
11. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA* **2003**, 100, 12063-12068.
12. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.* **2004**, 32, 2802-2818.

13. Jacobs, C. S.; Dervan, P. B., Modifications at the C-terminus to improve pyrrole-imidazole polyamide activity in cell culture. *J. Med. Chem.* **2009**, *52*, 7380-7388.
14. Nishijima, S.; Shinohara, K.; Bando, T.; Minoshima, M.; Kashiwazaki, G.; Sugiyama, H., Cell permeability of Py-Im-polyamide-fluorescein conjugates: Influence of molecular size and Py/Im content. *Bioorg. Med. Chem.* **2010**, *18*, 978-983.
15. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res.* **2007**, *35*, 363-370.
16. Herman, D. M.; Baird, E. E.; Dervan, P. B., Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.* **1998**, *120*, 1382-1391.
17. Chenoweth, D. M.; Dervan, P. B., Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 13175-13179.
18. Chenoweth, D. M.; Harki, D. A.; Phillips, J. W.; Dose, C.; Dervan, P. B., Cyclic pyrrole-imidazole polyamides targeted to the androgen response element. *J. Am. Chem. Soc.* **2009**, *131*, 7182-7188.
19. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.* **2002**, *10*, 2767-2774.
20. Krutzik, P. O.; Chamberlin, A. R., Synthesis of DNA-binding polyamides. Robust solid-phase methods for coupling heterocyclic aromatic amino acids. *Methods Mol. Biol.* **2002**, *201*, 77-92.
21. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J. Am. Chem. Soc.* **2008**, *130*, 6859-6866.
22. Vichai, V.; Kirtikara, K., Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat. Protoc.* **2006**, *1*, 1112-1116.
23. So, A. Y.; Chaivorapol, C.; Bolton, E. C.; Li, H.; Yamamoto, K. R., Determinants of cell- and gene-specific transcriptional regulation by the glucocorticoid receptor. *PLoS Genet.* **2007**, *3*, e94.
24. Reddy, T. E.; Pauli, F.; Sprouse, R. O.; Neff, N. F.; Newberry, K. M.; Garabedian, M. J.; Myers, R. M., Genomic determination of the glucocorticoid response reveals unexpected mechanisms of gene regulation. *Genome Res.* **2009**, *19*, 2163-2171.
25. Meijssing, S. H.; Pufall, M. A.; So, A. Y.; Bates, D. L.; Chen, L.; Yamamoto, K. R., DNA binding site sequence directs glucocorticoid receptor structure and activity. *Science* **2009**, *324*, 407-410.
26. Hsu, C. F.; Dervan, P. B., Quantitating the concentration of Py-Im polyamide-fluorescein conjugates in live cells. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5851-5855.

27. Crowley, K. S.; Phillion, D. P.; Woodard, S. S.; Schweitzer, B. A.; Singh, M.; Shabany, H.; Burnette, B.; Hippenmeyer, P.; Heitmeier, M.; Bashkin, J. K., Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorg. Med. Chem. Lett.* **2003**, 13, 1565-1570.
28. Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M., Arresting cancer proliferation by small-molecule gene regulation. *Chem. Biol.* **2004**, 11, 1583-1594.
29. Tsai, S. M.; Farkas, M. E.; Chou, C. J.; Gottesfeld, J. M.; Dervan, P. B., Unanticipated differences between alpha- and gamma-diaminobutyric acid-linked hairpin polyamide-alkylator conjugates. *Nucleic Acids Res.* **2007**, 35, 307-316.
30. Rucker, V. C.; Foister, S.; Melander, C.; Dervan, P. B., Sequence specific fluorescence detection of double strand DNA. *J. Am. Chem. Soc.* **2003**, 125, 1195-1202.
31. Bradford, J. A.; Clarke, S. T., Dual-pulse labeling using 5-ethynyl-2'-deoxyuridine (EdU) and 5-bromo-2'-deoxyuridine (BrdU) in flow cytometry. *Curr. Protoc. Cytom.* **2011**, 55, 7.38.1-7.38.15.
32. Poulin-Kerstien, A. T.; Dervan, P. B., DNA-templated dimerization of hairpin polyamides. *J. Am. Chem. Soc.* **2003**, 125, 15811-15821.
33. Gissot, A.; Camplo, M.; Grinstaff, M. W.; Barthelemy, P., Nucleoside, nucleotide and oligonucleotide based amphiphiles: a successful marriage of nucleic acids with lipids. *Org. Biomol. Chem.* **2008**, 6, 1324-1333.
34. Bijsterbosch, M. K.; Rump, E. T.; De Vruh, R. L.; Dorland, R.; van Veghel, R.; Tivel, K. L.; Biessen, E. A.; van Berkel, T. J.; Manoharan, M., Modulation of plasma protein binding and in vivo liver cell uptake of phosphorothioate oligodeoxynucleotides by cholesterol conjugation. *Nucleic Acids Res.* **2000**, 28, 2717-2725.
35. Van Dyke, M. W.; Hertzberg, R. P.; Dervan, P. B., Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA·Fe(II). *Proc. Natl. Acad. Sci. USA* **1982**, 79, 5470-5474.
36. Schultz, P. G.; Taylor, J. S.; Dervan, P. B., Design and Synthesis of a Sequence-Specific DNA Cleaving Molecule. (Distamycin-EDTA)iron(II). *J. Am. Chem. Soc.* **1982**, 104, 6861-6863.
37. Schmidt, T. L.; Heckel, A., Pyrrole/imidazole-polyamide anchors for DNA tertiary interactions. *Small* **2009**, 5, 1517-1520.
38. Cohen, J. D.; Sadowski, J. P.; Dervan, P. B., Programming multiple protein patterns on a single DNA nanostructure. *J. Am. Chem. Soc.* **2008**, 130, 402-403.
39. Johnson, D. S.; Mortazavi, A.; Myers, R. M.; Wold, B., Genome-wide mapping of in vivo protein-DNA interactions. *Science* **2007**, 316, 1497-1502.

40. Hesselberth, J. R.; Chen, X.; Zhang, Z.; Sabo, P. J.; Sandstrom, R.; Reynolds, A. P.; Thurman, R. E.; Neph, S.; Kuehn, M. S.; Noble, W. S.; Fields, S.; Stamatoyannopoulos, J. A., Global mapping of protein-DNA interactions in vivo by digital genomic footprinting. *Nat. Methods* **2009**, *6*, 283-289.
41. Ohtsukia, A.; Kimurab, M. T.; Minoshimaa, M.; Suzukib, T.; Ikedab, M.; Bandoa, T.; Nagase, H.; Shinohara, K.; Sugiyama, H., Synthesis and properties of PI polyamide-SAHA conjugate. *Tetrahedron Lett.* **2009**, *50*, 7288-7292.
42. Kwon, Y.; Arndt, H. D.; Mao, Q.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M., Small molecule transcription factor mimic. *J. Am. Chem. Soc.* **2004**, *126*, 15940-15941.
43. Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3930-3935.
44. Baird, E. E.; Dervan, P. B., Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146.
45. Chenoweth, D. M.; Harki, D. A.; Dervan, P. B., Solution-phase synthesis of pyrrole-imidazole polyamides. *J. Am. Chem. Soc.* **2009**, *131*, 7175-7181.
46. Krutzik, P. O.; Chamberlin, A. R., Synthesis of DNA-binding polyamides. Robust solid-phase methods for coupling heterocyclic aromatic amino acids. *Methods Mol. Biol.* **2002**, *201*, 77-92.

Chapter 3

Synthesis and Biological Activity of Cyclic Pyrrole-Imidazole Polyamide Libraries

The text of this chapter was taken in part from a manuscript co-authored with Benjamin C. Li, James W. Puckett, and Peter B. Dervan (California Institute of Technology)

(Li, B. C.; Montgomery, D. C.; Puckett, J. W.; Dervan, P. B. "Synthesis of Cyclic Py-Im Polyamide Libraries," *J. Org. Chem.* **2013**, 78, 124-133)

Abstract

Cyclic Py-Im polyamides containing two GABA turn units exhibit enhanced DNA binding affinity, but extensive studies of their biological properties have been hindered due to synthetic inaccessibility. A facile modular approach towards cyclic polyamides has been developed via microwave-assisted solid-phase synthesis of hairpin amino acid oligomer intermediates followed by macrocyclization. A focused library of cyclic polyamides **1-7** targeted to the androgen response element (ARE) and the estrogen response element (ERE) were synthesized in 12-17% overall yield. The Fmoc protection strategy also allows for selective modifications on the GABA turn units that have been shown to improve cellular uptake properties. The DNA binding affinities of a library of cyclic polyamides were measured by DNA thermal denaturation assays and compared to the corresponding hairpin polyamides. Fluorescein-labeled cyclic polyamides have been synthesized and imaged via confocal microscopy in A549 and T47D cell lines. The IC₅₀ values of compounds **1-7** and **9-11** were determined, revealing remarkably varying levels of cytotoxicity.

Introduction

The selective modulation of eukaryotic gene expression by small molecules may have important implications in the field of chemical biology and human medicine. Pyrrole-imidazole polyamides are a class of synthetic ligands that can be programmed to bind the minor groove of specific DNA sequences.¹ Antiparallel, side-by-side N-methylpyrrole (Py) and N-methylimidazole (Im) carboxamides (Im/Py) pairs distinguish G·C from C·G base pairs, N-methyl-3-hydroxypyrrole (Hp)/Py shows specificity for T·A over A·T, whereas Py/Py pairs are specific for both T·A and A·T.²⁻⁵ By linking two strands of these heterocyclic oligomers via a γ -amino butyric acid (GABA) turn unit, hairpin Py-Im polyamides can be programmed to bind a large library of DNA sequences with affinities comparable to natural DNA-binding proteins.⁶⁻⁸ Hairpin polyamides have been shown to localize to the nuclei of living cells, and regulate endogenous gene expression by disrupting protein/DNA interfaces.⁹⁻¹⁷ Cyclic polyamides containing a second GABA turn unit exhibit further enhanced DNA binding properties.¹⁸⁻²¹ We have recently demonstrated their gene regulatory effects on AR-activated gene expression in prostate cancer models.²²

This discovery has opened a new area of research towards transcriptional regulation with small molecules, but the relative synthetic inaccessibility of cyclic polyamides has remained a bottleneck for examining libraries of structural variants that would modulate affinity, cell uptake and biological activity. Initial solid-phase methods were low yielding, and required substantial pre-modifications of the PAM resin.¹⁸⁻²⁰ While the solution-phase synthesis of cyclic polyamides remains useful in large-scale target-oriented synthesis, it has limited practicality towards libraries for screening

biological activities.²² The recent report by Morinaga et al. offers a modular approach to achieve cyclic polyamides by intramolecular coupling of a cysteine and a chloroacetyl residue, but the modification of the optimal three-carbon GABA turn into a sulfur-containing four-atom linker compromises its DNA binding affinity and may alter its biological properties.²³

We report here a solid-phase polyamide synthesis of a key hairpin amino acid oligomer intermediate, which followed by intramolecular cyclization, affords cyclic polyamides **1-8** in good yields. The polyamides were synthesized step-wise on 2-chlorotrityl resin. The modular approach led to rapid access of a focused library of cyclic polyamides **1-7** with various core sequences and turn unit modifications. The utilization of Fmoc chemistry allowed for differentially protected turn units, which were modified selectively to complement existing cellular imaging and cell uptake enhancement technologies.²⁴ We examined the DNA binding properties and cytotoxicity profiles of compounds **1-11**, and the cellular localization of cyclic polyamides **12-14** by fluorescence microscopy.

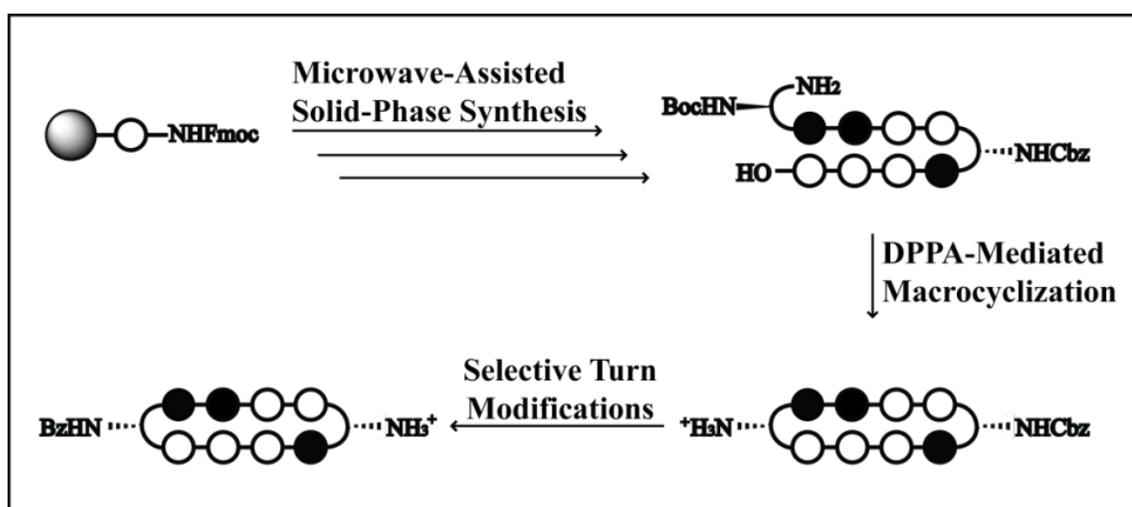


Figure 3.1: The basic strategy for synthesis of differentially protected cyclic Py-Im polyamides.

Results and Discussion

Microwave-Assisted Solid-Phase Synthesis

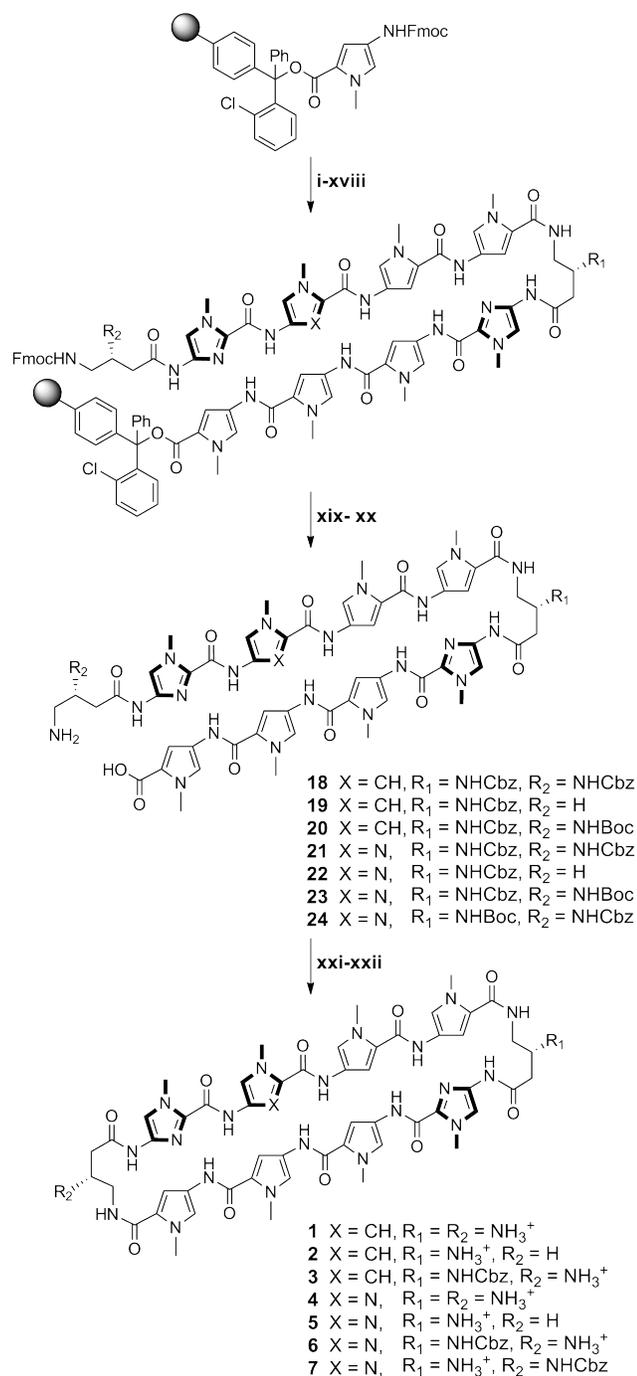
Due to previously observed decomposition of the conjugated C-terminal free carboxylic acid in polyamide intermediates, 2-chlorotrityl-chloride (2-Cl-Trt-Cl) resin was chosen for its mild synthesis and cleavage conditions. Polyamide synthesis on this resin has been previously reported by Aldrich-Wright and co-workers, but a resin-bound β -alanine linker was used in both instances, and a new loading procedure was therefore needed.^{25,26} 2-Cl-Trt-Cl resin was first loaded with the Fmoc-protected Py monomer in N,N-dimethylformamide (DMF) and capped with methanol. Resin substitution levels were determined by the Fmoc test and confirmed by weighing the dry mass of the loaded resin. Fmoc deprotection was achieved using a 50% piperidine in DMF solution. In light of the recent improvements in both efficiency and yield, the couplings were performed under microwave-assisted conditions using the desired PyBOP-activated monomers.²⁷

Resin-Bound Nucleophile	Deprotection Times ^a	Coupling Times ^b (min)		
		Py	Im	GABA/ β -Ala
Py	3 x 10min	20	20	20
Im	3 x 10min	IR ^c	30	30
GABA/ β -Ala	2 x 5min	20	20	20

Table 3.1 Standard Fmoc deprotection and microwave-assisted coupling times for solid-phase polyamide synthesis. ^aAll deprotections were performed in 50% piperidine in DMF. ^bAll coupling reactions were conducted under microwave-assisted conditions at 50 °C with a 0.3M solution of the activated monomers (3 eq. monomer acid, 3 eq. PyBOP, 8 eq. DIEA, DMF). ^cFmocPyOH coupling onto resin with N-terminal Im was incomplete even at 60°C for up to 1h. Synthesis of polyamide sequences that require this linkage should use the FmocPyImOH dimer instead, demonstrated later in the synthesis of **8**.

Initial syntheses performed at 60°C led to premature cleavage of intermediates off the 2-Cl-Trt resin, and 50°C couplings were therefore preferable. The challenging Im to Py coupling required an FmocPyImOH dimer, which was obtained via an optimized procedure

by Weltzer and Wemmer.²⁸ The deprotection and coupling conditions are detailed in Table 3.1. This 2-step deprotection-coupling procedure was repeated until the desired polyamide sequence was achieved. To build the small library of polyamides in a modular fashion, the resin was split into different batches at corresponding steps for further derivatization. Upon completion, the N-terminal Fmoc-protected polyamide oligomer was cleaved from the resin with 30% hexafluoroisopropanol (HFIP) in dichloromethane (DCM), concentrated *in vacuo*, and the resulting residue was subjected to a 20% piperidine solution to remove the Fmoc group. Direct cleavage of the free-amine polyamide oligomer was attempted, but found to be ineffective due to poor solubility of the zwitterion intermediate in the cleavage solution. After purification by high-performance liquid chromatography (HPLC), the desired polyamide intermediates **18-24** were obtained in 31-40% yields.



Scheme 3.1: Microwave-assisted synthesis of cyclic polyamides **1-7**. ¹All PyBOP-mediated coupling conditions were performed under microwave-assisted conditions (see Table 3.1). ²Reagents and conditions: (i) 50% piperidine, DMF; (ii) FmocPyOH, PyBOP, DIEA, DMF; (iii) 50% piperidine, DMF; (iv) FmocPyOH, PyBOP, DIEA, DMF; (v) 50% piperidine, DMF; (vi) FmocImOH, PyBOP, DIEA, DMF; (vii) 50% piperidine, DMF; (viii) Z-β-Dab(Fmoc)-OH (for **1-6**) or Boc-β-Dab(Fmoc)-OH (for **7**), PyBOP, DIEA, DMF; (ix) 50% piperidine, DMF; (x) FmocPyOH, PyBOP, DIEA, DMF; (xi) 50% piperidine, DMF; (xii) FmocPyOH, PyBOP, DIEA, DMF; (xiii) 50% piperidine, DMF; (xiv) FmocPyOH (for **1-3**) or FmocImOH (for **4-7**), PyBOP, DIEA, DMF; (xv) 50% piperidine, DMF; (xvi) FmocImOH, PyBOP, DIEA, DMF; (xvii) 50% piperidine, DMF; (xviii) Z-β-Dab(Fmoc)-OH (for **1 & 4 & 7**) or Fmoc-GABA-OH (for **2 & 5**) or Boc-β-Dab(Fmoc)-OH (for **3 & 6**), PyBOP, DIEA, DMF; (xix) 30% HFIP, DCM; (xx) 20% piperidine, DMF; (xxi) DPPA, DIEA, DMF; (xxii) 10% TFMSA, TFA (for **1, 2, 4 & 5**) or TFA (for **3, 6 & 7**).

DPPA-Mediated Macrocyclization

The polyamide macrocyclization step was achieved by a DPPA-mediated ring-closing reaction between the N-terminal amino group and the C-terminal carboxylic acid. This method was first employed by Cho et al. in the synthesis of cyclic polyamides.^{18,29,30} In order to obtain a general work-up procedure applicable to polyamides of various lipophilicities, diisopropylethylamine (DIEA) was used as the base in place of sodium bicarbonate (NaHCO₃) in the original conditions.³¹ Deprotection of the turn units with trifluoromethanesulfonic acid (TFMSA) or trifluoroacetic acid (TFA), followed by HPLC purification, afforded polyamides **1-7** in **37-43%** yields over 2 steps.

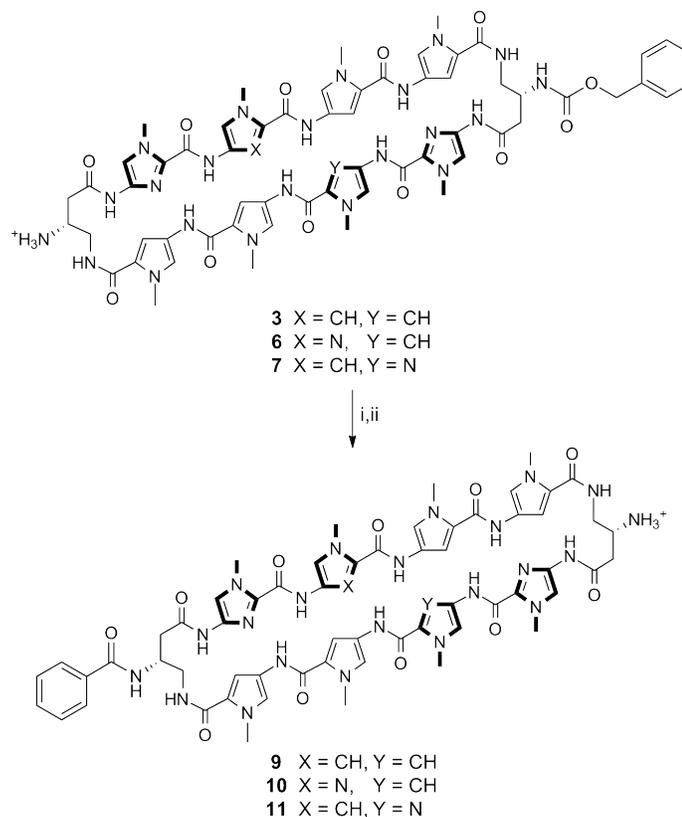
Polyamide Precursor	[M+H] ⁺ (expected)	m/z (observed)	Solid-Phase Yield (%)	Cyclic Polyamide	[M+H] ⁺ (expected)	m/z (observed)	Cyclization Yield (%)	Overall Yield (%)
	1465.4	1465.9	33		1179.5	1179.9	35	12
	1316.5	1316.9	40		1164.5	1164.9	42	17
	1431.6	1453.9 [M+Na] ⁺	34		1313.6	1314.0	47	16
	1466.5	1466.9	31		1180.5	1180.9	39	12
	1317.5	1317.2	34		1165.5	1165.5	48	16
	1432.6	1455.0 [M+Na] ⁺	32		1314.5	1314.7	37	12
	1432.6	1454.9 [M+Na] ⁺	33		1314.5	1314.8	38	13
	1414.6	1372.0 [M-CO2] ⁺	34		1129.5	1130.0	39	13

Table 3.2: Summary Table of MALDI-TOF Data and Synthetic Yields for Cyclic Polyamides **1-8** and Intermediates **18-25**.

Selective Derivatization of Cyclic Polyamide Turn Units

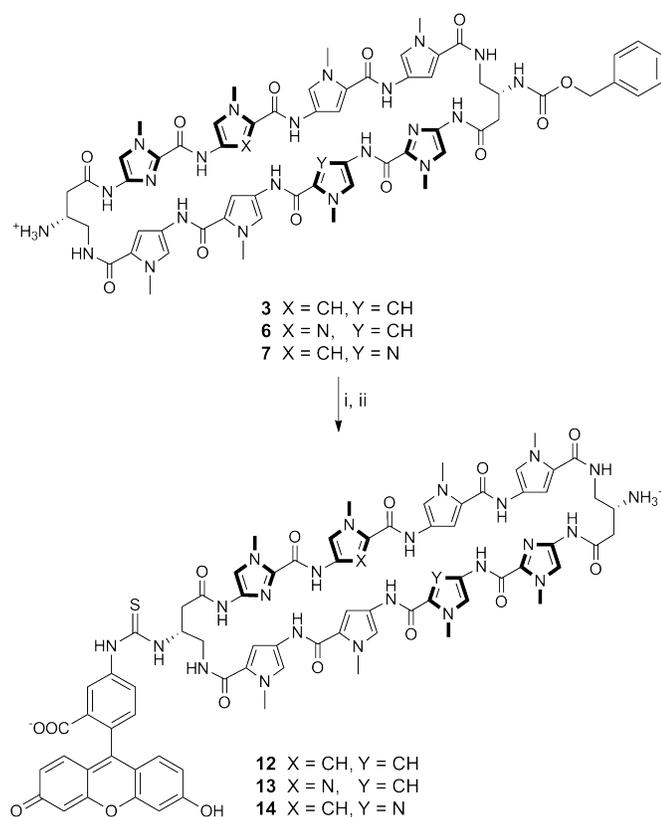
By taking advantage of the Fmoc protection scheme, the two GABA β-amino groups in **3** were differentially protected. This is further highlighted in **6** and **7**, which

share the same asymmetric polyamide core targeted to the 5'-WGGWCW-3' sequence, and allowed for the selective conjugation of a benzoic acid moiety on a single turn unit in polyamides **9-11** that has been recently developed to enhance the cellular localization properties of hairpin polyamides.²⁴



Scheme 3.2: Preparation of cyclic polyamides **9-11**. Reagents and conditions: (i) BzOH, PyBOP, DIEA, DMF; (ii) 10% TFMSA, TFA.

Mono-substituted benzyl carbamate (Cbz) polyamides **3**, **6** and **7** were chosen as targets based on unpublished results indicating Cbz-functionalized polyamides are biologically active (Supp. Fig. S16). Cyclic polyamides **12-14** with a fluorescein dye were synthesized in a similar fashion, and imaged in live cells via fluorescence microscopy. Furthermore, both the solubility and the pharmacokinetic profiles of cyclic polyamides can be highly dependent on subtle structural modifications, and this method allows for the modular synthesis of these structural variants in an efficient manner.³¹⁻³³

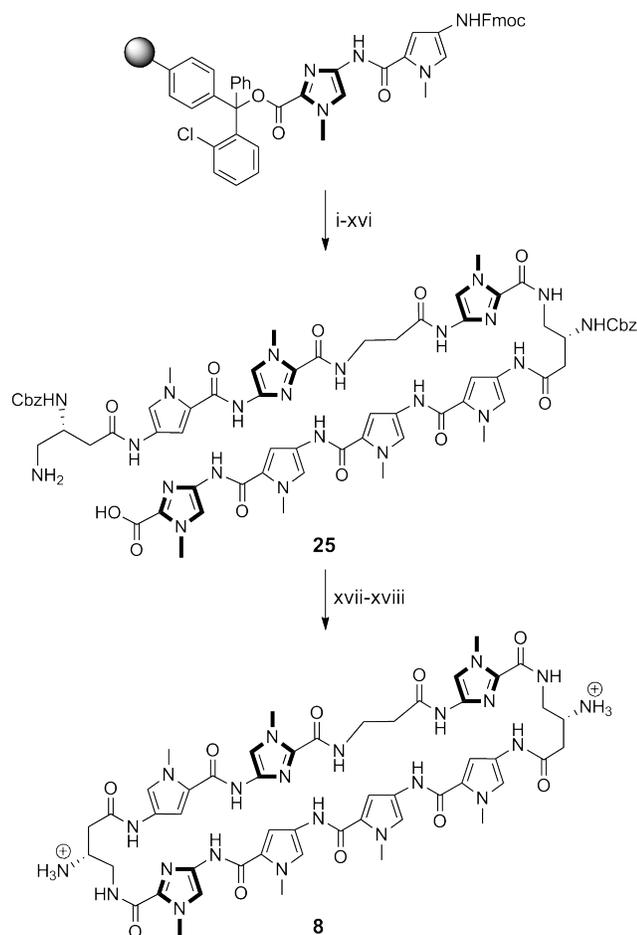


Scheme 3.3: Preparation of FITC-labeled cyclic polyamides **12-14**. Reagents and conditions: (i) FITC, DIEA, DMF; (ii) 10% TFMSA, TFA.

Synthesis of Cyclic and Hairpin Polyamides with C-terminal Imidazole Units

Previously established solid-phase polyamide synthesis methods have been generally limited to sequences beginning with a pyrrole monomer.^{34,35} Solid-phase synthesis of polyamides starting with imidazole units on the commonly used Kaiser oxime resin have been low yielding, mainly attributed to the sensitivity of the oxime-imidazole linkage that lead to premature cleavage of resin-bound intermediates. The addition of an aliphatic linker (e.g. Boc- β -Ala-PAM resin) circumvents this issue, but previous studies on hairpin polyamides with C-terminal β -alanine motifs have shown reduced cellular uptake properties and thus diminished gene regulatory effects of these products.⁹ Using the microwave-assisted conditions reported above, cyclic polyamide **8** targeted to the 5'-

WCGWGW-3' sequence found in E-Box binding sites have been synthesized in 13% yield overall.

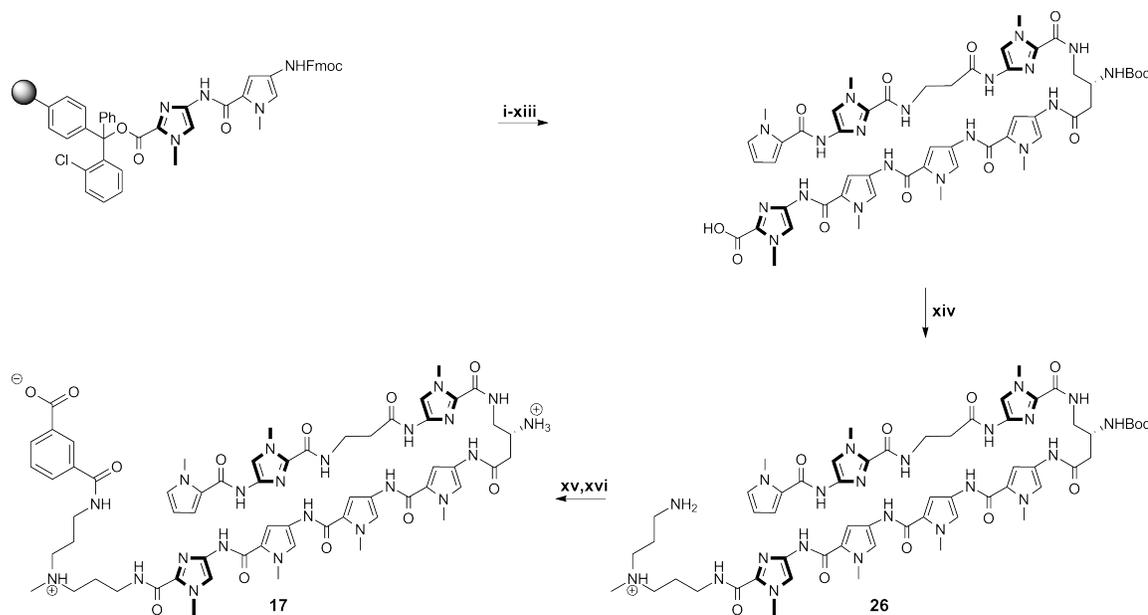


Scheme 3.4: Preparation of cyclic polyamide **8**. All PyBOP-mediated coupling conditions were performed under microwave-assisted conditions (see Table 3.1). Reagents and conditions: (i) 50% piperidine, DMF; (ii) FmocPyOH, PyBOP, DIEA, DMF; (iii) 50% piperidine, DMF; (iv) FmocPyOH, PyBOP, DIEA, DMF; (v) 50% piperidine, DMF; (vi) Z-β-Dab(Fmoc)-OH, PyBOP, DIEA, DMF; (vii) 50% piperidine, DMF; (viii) FmocImOH, PyBOP, DIEA, DMF; (ix) 50% piperidine, DMF; (x) Fmoc-β-Ala-OH, PyBOP, DIEA, DMF; (xi) 50% piperidine, DMF; (xii) FmocPyImOH, PyBOP, DIEA, DMF; (xiii) 50% piperidine, DMF; (xiv) Z-β-Dab(Fmoc)-OH, PyBOP, DIEA, DMF; (xv) 30% HFIP, DCM; (xvi) 20% piperidine, DMF; (xvii) DPPA, DIEA, DMF; (xviii) 10% TFMSA, TFA.

Initial attempts starting with FmocImOH-loaded resin resulted in undesired cleavage during subsequent steps, and so syntheses began with loading of the FmocPyImOH dimer onto 2-chlorotrityl resin. After deprotection-coupling of the corresponding monomer units, followed by resin cleavage and Fmoc removal, polyamide

intermediate **25** was isolated by HPLC purification in 34% yield. DPPA-mediated cyclization of **25**, followed by Cbz deprotection afforded cyclic polyamide **8** in 39% yield over 2 steps.

Hairpin polyamide **17** was synthesized in a similar stepwise manner to afford the Im-capped, C-terminal acid intermediate, which was then coupled to a 3,3'-diamino-N-methyldipropylamine linker, followed by isophthalic acid conjugation, Boc deprotection, and isolation in 24% yield over 16 steps (Scheme 3.5). This is a step forward which allows for the synthesis of both cyclic and non- β -alanine-linked hairpin polyamides with sequences beginning with an imidazole unit, further expanding the scope of targetable DNA sequences and inhibition of transcription factor-mediated gene expression by Py-Im polyamides.



Scheme 3.5: Preparation of hairpin polyamide **17**. All PyBOP-mediated coupling conditions were performed under microwave-assisted conditions (see Table 3.1). Reagents and conditions: (i) 50% piperidine, DMF; (ii) FmocPyOH, PyBOP, DIEA, DMF; (iii) 50% piperidine, DMF; (iv) FmocPyOH, PyBOP, DIEA, DMF; (v) 50% piperidine, DMF; (vi) Boc-b-Dab(Fmoc)-OH, PyBOP, DIEA, DMF; (vii) 50% piperidine, DMF; (viii) FmocImOH, PyBOP, DIEA, DMF; (ix) 50% piperidine, DMF; (x) Fmoc- β -Ala-OH, PyBOP, DIEA, DMF; (xi) 50% piperidine, DMF; (xii) PyImOH, PyBOP, DIEA, DMF; (xiii) 30% HFIP, DCM; (xiv) 3,3'-diamino-N-methyldipropylamine, PyBOP, DMF; (xv) Isophthalic acid, PyBOP, DIEA, DMF; (xvi) TFA.

Thermal Stabilization of DNA Duplexes by Polyamides

Py-Im polyamide-DNA binding affinities and specificities have historically been measured by quantitative DNase I footprinting assays.³⁶ As previously reported, however, cyclic polyamides have exceptionally high DNA binding affinities that exceed the detection limit of this experiment (i.e. $K_a \geq 2 \times 10^{10} \text{ M}^{-1}$).^{22,37,38} The DNA binding affinities of cyclic polyamides **1-14** have been rank ordered by magnitude of DNA thermal stabilization (ΔT_m), and compared to the corresponding hairpin polyamides **14-16**. Spectroscopic measurements were performed on 12-mer DNA duplexes with sequences 5'-CGATGTTCAAGC-3', 5'-CGATGGTCAAGC-3', and 5'-CGATCGTGAAGC-3', each containing a match binding site for the corresponding polyamides.

Consistent with the findings of Chenoweth et al, the ΔT_m value for bis-amino cyclic polyamide **1** ($\Delta T_m = 26.1^\circ\text{C}$) was calculated to be significantly higher than that of hairpin polyamide **15** ($\Delta T_m = 22.0^\circ\text{C}$).²² So while **15** has an established binding affinity to the match androgen response element (ARE) half-site (5'-WGWWCW-3') in the subnanomolar range, cyclic polyamide **1** provides even greater stabilization to such DNA duplexes. Mono-unsubstituted cyclic polyamide **2** provides less DNA stabilization compared to **1** ($\Delta T_m = 20.4^\circ\text{C}$), which is likely due to the loss of a positive charge and thus the loss of favorable electrostatic interactions with the negatively-charged DNA backbone. Perhaps more surprising is the high ΔT_m values retained by mono-protected cyclic polyamides **3** ($\Delta T_m = 27.3^\circ\text{C}$) and **9** ($\Delta T_m = 28.0^\circ\text{C}$), each containing a lone free amino group and net +1 charge. As shown in Figure 3.2, the benzoyl (Bz) group in **9** projects straight down the minor groove, avoiding unfavorable steric interactions with the groove

wall, and may offer insight into the high degree of DNA stabilization by **3** and **9** comparable to the bis-amino cycle **1**.

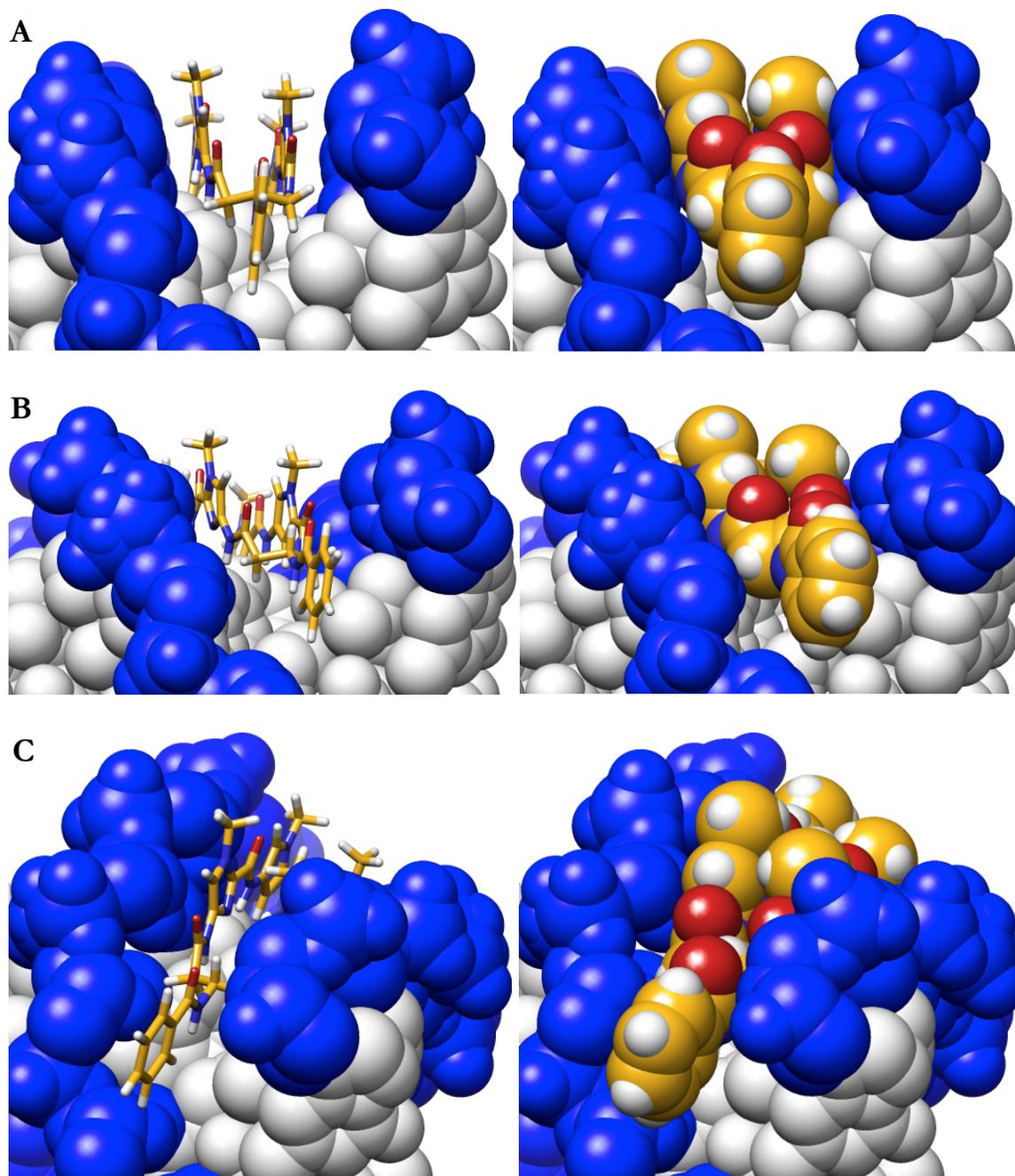


Figure 3.2 (Left) Stick model and (right) space-filling model of benzoyl-substituted turn along the DNA minor groove from three different angles (A-C). Based on published crystal structure by Chenoweth et al (PDB ID: 3OMJ).

Fluorescein-conjugate **12** ($\Delta T_m = 21.8^\circ\text{C}$) has a DNA binding affinity lower than **3** and **9**, perhaps due to increased steric clashes from the larger substitution group and unfavorable electrostatic interactions from the negatively-charged fluorescein group, but still binds DNA at a similar level to **15**.

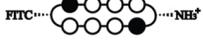
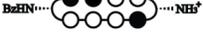
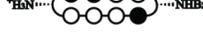
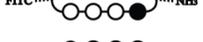
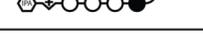
Polyamides	5'- CGA TGGTCA AGC -3'	
	$T_m / ^\circ\text{C}$	$\Delta T_m / ^\circ\text{C}$
—	53.1 (± 0.2)	—
 (1)	79.2 (± 0.2)	26.1
 (2)	73.5 (± 0.2)	20.4
 (3)	80.4 (± 0.2)	27.3
 (9)	81.1 (± 0.2)	28.0
 (12)	74.9 (± 0.1)	21.8
 (15)	75.1 (± 0.1)	22.0
5'- CGA TGGTCA AGC -3'		
—	56.2 (± 0.3)	—
 (4)	79.4 (± 0.3)	23.2
 (5)	74.8 (± 0.2)	18.6
 (6)	79.8 (± 0.2)	23.6
 (7)	80.5 (± 0.3)	24.3
 (10)	82.2 (± 0.1)	26.0
 (11)	81.5 (± 0.1)	25.1
 (13)	78.4 (± 0.6)	22.2
 (14)	77.2 (± 0.1)	21.0
 (16)	72.9 (± 0.2)	16.7
5'- CGA TCGTGA AGC -3'		
—	56.7 (± 0.2)	—
 (8)	70.1 (± 0.1)	13.4
 (17)	63.3 (± 0.3)	6.6

Table 3.3: T_m Values for Polyamide Library.

The magnitude of stabilization provided by cyclic polyamides **4-7**, **10-11**, and **13-14** targeted to estrogen response element (ERE) half-sites (5'-WGGWCW-3') follows a similar pattern to the aforementioned ARE-targeting series. Mono-unsubstituted cycle **5** ($\Delta T_m = 18.6^\circ\text{C}$) stabilizes the duplex at a comparable level to hairpin polyamide **16** ($\Delta T_m = 16.7^\circ\text{C}$), whereas bis-amino compound **4** ($\Delta T_m = 23.2^\circ\text{C}$) has a higher ΔT_m value. The mono-Cbz cycles **6** ($\Delta T_m = 23.6^\circ\text{C}$) and **7** ($\Delta T_m = 24.3^\circ\text{C}$), mono-Bz substituted **10** ($\Delta T_m = 26.0^\circ\text{C}$) and **11** ($\Delta T_m = 25.1^\circ\text{C}$), and mono-fluorescein conjugates **13** ($\Delta T_m = 22.2^\circ\text{C}$) and **14** ($\Delta T_m = 21.0^\circ\text{C}$) each bind DNA similar to **4**.

Cyclic polyamide **8** ($\Delta T_m = 13.4^\circ\text{C}$) binds the match 5'-WCGWGW-3' sequence at an elevated level compared to hairpin **17** ($\Delta T_m = 6.6^\circ\text{C}$), which may prove important towards targeting oncogenic transcription factors such as c-Myc that act through binding canonical E-Box (5'-CACGTG-3') sequences.

Sulforhodamine B Cytotoxicity Assay for Compounds 1-11

The cytotoxicities of compounds **1-11** were assessed in A549 human lung carcinoma and T47D human breast cancer cell lines (Tables 3.4 & 3.5). The cyclic polyamides targeted to the 5'-WGWWCW-3' sequence generally exhibit a higher level of cytotoxicity than the 5'-WGGWCW-3' series, which is consistent with the trends observed in the DNA thermal denaturation analysis and confocal microscopy studies. Detailed inspection of the IC_{50} values within the series of compounds and across the cell lines, however, offers some unanticipated insights into the different biological properties of these minor structural variants.

Bis-amino cycle **1**, which has previously been shown to be biologically active, did not display any significant level of cytotoxicity in either cell line ($IC_{50} > 30 \mu M$). Mono-unsubstituted compound **2** ($IC_{50} = 4.9 \mu M$) and mono-Bz **9** ($IC_{50} = 1.0 \mu M$) were comparably cytotoxic to hairpin polyamide **15** ($IC_{50} = 3.1 \mu M$) in A549 cells, but an order of magnitude more cytotoxic ($IC_{50} = 74 \text{ nM}$ & 79 nM , respectively) than **15** ($IC_{50} = 710 \text{ nM}$) in T47D cells. The mono-Cbz cycle **3** was consistently the most cytotoxic compound in both A549 ($IC_{50} = 160 \text{ nM}$) and T47D ($IC_{50} = 25 \text{ nM}$) cell lines.

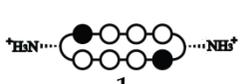
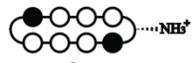
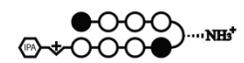
Cyclic Polyamide	A549	T47D
	$IC_{50} (\mu M)$	$IC_{50} (\mu M)$
 1	>30	>30
 2	4.9 ± 1.8	0.074 ± 0.011
 3	0.16 ± 0.05	0.025 ± 0.005
 9	1.0 ± 0.2	0.079 ± 0.003
 15	3.1 ± 0.6	0.71 ± 0.10

Table 3.4: SRB cytotoxicity data on compounds **1-3**, **9**, and **15**, in A549 and T47D cells, 72 hr incubation.

For the 5'-WGGWCW-3' targeting polyamides, the only compound that exhibited an IC_{50} value lower than $30 \mu M$ in A549 cells was the mono-Cbz compound **6** ($IC_{50} = 1.9 \mu M$). In T47D cells, consistent with the 5'-WGWWCW-3' series, bis-amino cycle **4** ($IC_{50} > 30 \mu M$) was found to be not significantly cytotoxic and **6** ($IC_{50} = 460 \text{ nM}$) was the most cytotoxic compound. Mono-unsubstituted cycle **5** ($IC_{50} = 0.82 \mu M$) again shares a comparable level of cytotoxicity with the reference hairpin **16** ($IC_{50} = 1.1 \mu M$), whereas mono-Bz compounds **10** ($IC_{50} = 13.3 \mu M$) and **11** ($IC_{50} = 7.5 \mu M$) are both an order of magnitude less cytotoxic.

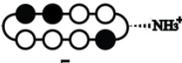
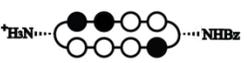
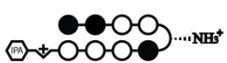
Cyclic Polyamide	A549	T47D
	IC ₅₀ (μM)	IC ₅₀ (μM)
 4	>30	>30
 5	>30	0.82 ± 0.21
 6	1.9 ± 0.2	0.46 ± 0.27
 7	>30	>30
 10	>30	13.3 ± 0.6
 11	>30	7.5 ± 0.7
 16	>30	1.1 ± 0.1

Table 3.5: SRB cytotoxicity data on compounds **4-7**, **10-11**, and **16**, in A549 and T47D cells, 72 hr incubation.

Perhaps most interestingly, mono-Cbz compound **7** did not exhibit observable levels of cytotoxicity ($IC_{50} > 30 \mu M$) in either A549 or T47D cells. Considering that **7** is a regioisomer of **6**, where the Cbz group is simply swapped onto the other turn, and that **7** only differs from **3** by a single $-CH$ to $-N$ substitution, it is rather surprising that **7** is more than 15- to 65- fold less cytotoxic than **6** and at least 180- to 1200- fold less cytotoxic than **3** in the two examined cell lines. Given the comparable DNA stabilization properties between **6** and **7**, and their common core sequence, we would not have predicted this vast discrepancy in cytotoxicity.

This study has demonstrated the large and somewhat unpredictable effects in biological activity induced by small structural variations of cyclic polyamides. Based on our preliminary work, the aggregation and pharmacokinetic properties of polyamides also vary greatly depending on structural modifications.³¹⁻³³ All this combines to highlight the

importance of a fast and reliable method to generate focused libraries of cyclic polyamides for future research.

Confocal Microscopy of Cyclic Polyamide-Fluorescein Conjugates 12-14

To directly examine the cellular localization of cyclic polyamides, fluorescein conjugates **12-14** were synthesized and visualized in living cells via confocal microscopy (Figs. 2 & 3). The selective conjugation of a single fluorescein molecule not only helped retain a free amino group for solubility purposes, but also allowed for the qualitative comparison of the two fluorescein conjugates **13** and **14** that share the same asymmetric polyamide core.

In each of the cases examined, cyclic polyamides **12-14** appear to permeate through the cellular membrane and localize in the cell nucleus, which was confirmed by the co-localization with Hoechst 33258 DNA stain. For ease of qualitatively assessing compound uptake, all A549 images were taken at a 660 fluorescence gain level, and all T47D images were taken at 600 fluorescence gain. Comparing Figures 3.3 and 3.4, the fluorescence levels of compounds **12-14** in T47D cells are all significantly higher than in A549, and this is only further amplified by this difference in gain levels.

Compound **12** matched to the 5'-WGWWCW-3' sequence exhibits the highest level of nuclear localization in both cell lines. Among the two 5'-WGGWCW-3' targeting cycles, polyamide **14** qualitatively appears to have a relatively higher fluorescence signal in the cell nuclei in both A549 and T47D cells. This may help explain the cytotoxicity data reported above, where compounds **6** and **11** with Cbz and Bz substitutions on the same side

as the fluorescein in **14** consistently display larger biological effects than **7** and **10**, which are more structurally similar to cyclic polyamide **13**.

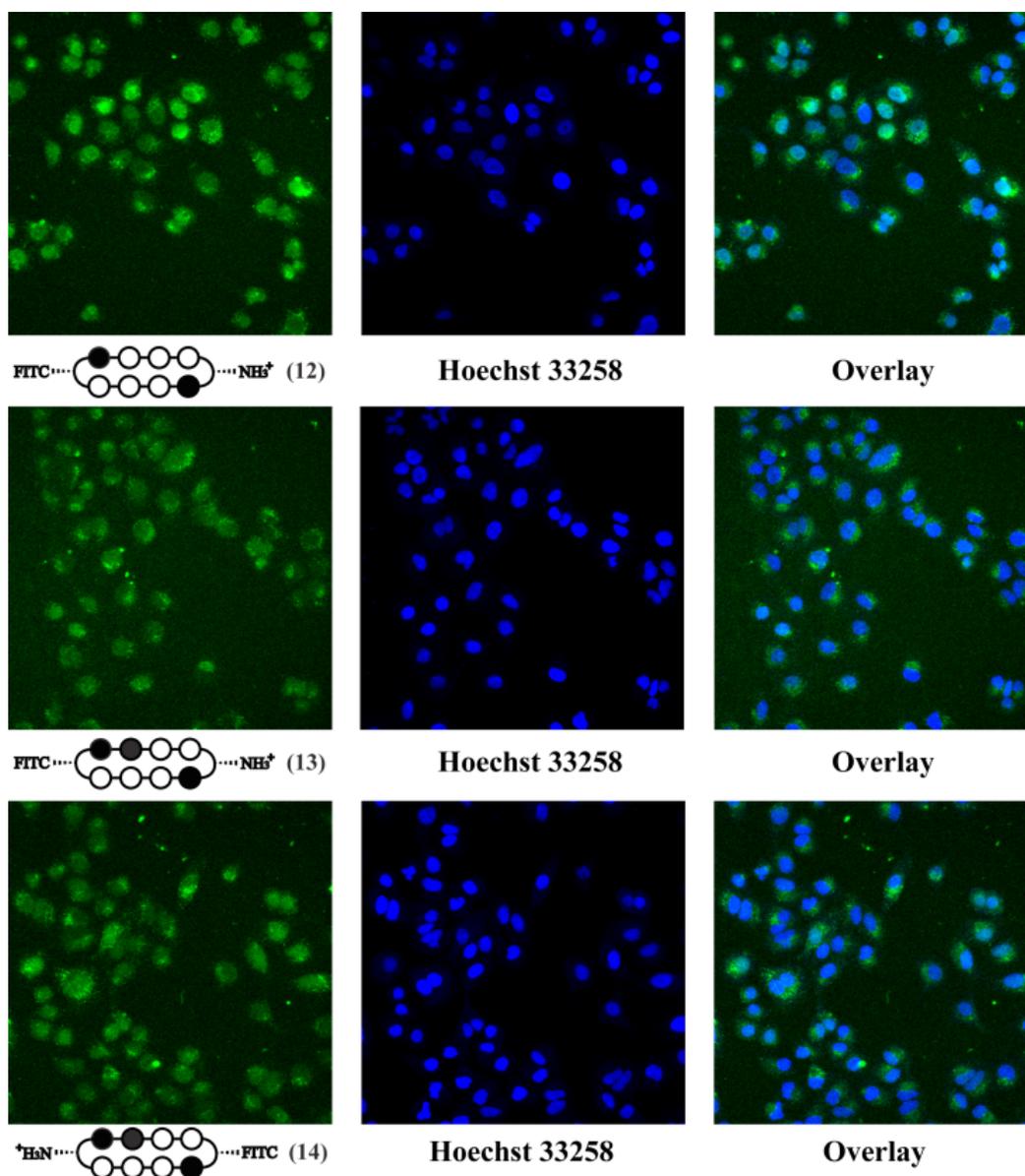


Figure 3.3: Confocal microscopy of cyclic polyamide-fluorescein conjugates **12** (top), **13** (middle), and **14** (bottom) in A549 cells. In order to confirm nuclear localization, the fluorescence panel (left) was compared with Hoechst 33258 staining (middle) and overlay (right).

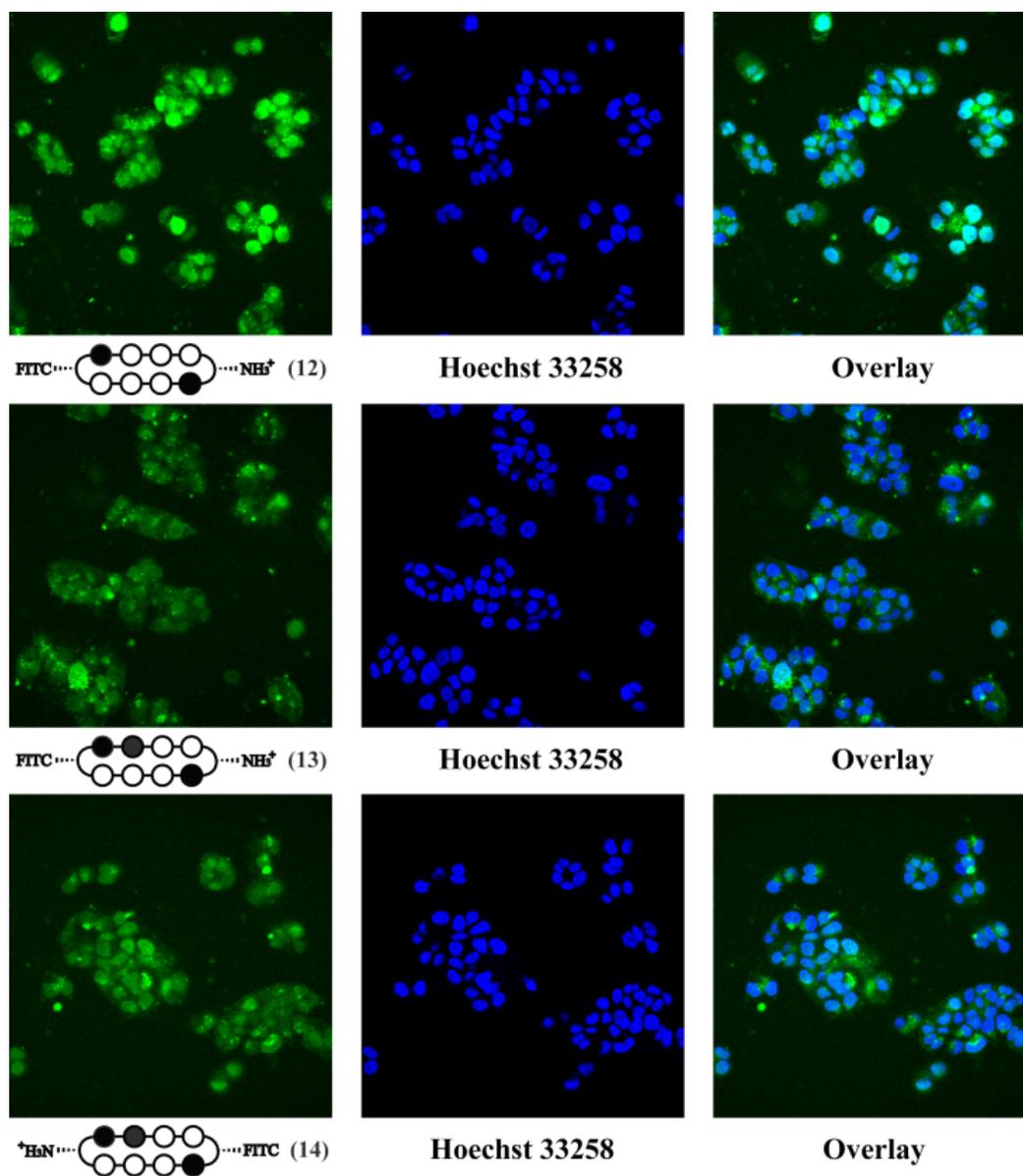


Figure 3.4: Confocal microscopy of cyclic polyamide-fluorescein conjugates **12** (top), **13** (middle), and **14** (bottom) in T47D cells. In order to confirm nuclear localization, the fluorescence panel (left) was compared with Hoechst 33258 staining (middle) and overlay (right).

Conclusions

We have described a modular solid-phase synthesis method, which, when combined with an established DPPA-mediated macrocyclization step, afforded cyclic polyamides in a high-yielding and time-efficient manner. Using this method, we have overcome previous limitations and synthesized both cyclic and hairpin polyamides that start with an imidazole unit. The binding affinities of all synthesized cycles have been assessed by DNA thermal denaturation assays and compare favorably to hairpin polyamides that bind their match DNA sequences at subnanomolar concentrations. Furthermore, the protection strategy of our method allows for selective modification of the GABA turn units, which we have used to rapidly generate a focused library of compounds. The cytotoxicity and uptake analysis of the cyclic polyamides revealed unexpected properties that further highlight the need for an efficient method to synthesize structural variants of cyclic polyamides for future studies.

Materials and Methods

Reagents and Equipment

2-Chlorotrityl chloride (2-Cl-Trt-Cl) resin was purchased from Bachem. FmocPyOH and FmocImOH monomers were purchased from Wako. PyBOP was purchased from NovaBioChem. Boc- β -Dab(Fmoc)-OH was purchased from Peptides International. All DNA oligomers were purchased HPLC purified from Integrated DNA Technologies. Cell culture media was purchased from Gibco. Fetal bovine serum was purchased from Omega Scientific. Microwave-assisted coupling reactions were conducted on a Biotage Initiator Eight synthesizer.

Polyamide Quantification

Polyamide concentrations were measured in 20% MeCN in 0.1% (v/v) aqueous TFA using an approximated extinction coefficient of $69\,200\text{ M}^{-1}\text{ cm}^{-1}$ at λ_{max} near 310 nm, unless otherwise specified.^{32,36}

Monomer Loading onto 2-Cl-Trt Resin

Prior to manual microwave-assisted synthesis, 2-Cl-Trt-Cl resin (1.0 g, 1.59 mmol/g) was first loaded by mixing with 576 mg (1.59 mmol, 1 eq.) of FmocPyOH monomer, followed by addition of 6 mL dimethylformamide (DMF) and 1.38 mL diisopropylethylamine (DIEA) (7.59 mmol, 5 eq.). The suspended mixture was stirred for 18 hr, then capped by addition of 1 mL methanol (MeOH) and stirred for 1 hr. The orange-colored, loaded resin was then collected on a fritted peptide synthesis vessel, washed with DMF (2x), MeOH (2x), DMF (2x), MeOH (2x), and diethyl ether (Et₂O). (Owing to the

sensitivity of the 2-Cl-Trt resin towards hydrolysis/methanolysis, this final Et₂O wash was found to be essential and all loaded resin was sealed and stored at -20°C.) The loading efficiency was quantitated via the Fmoc test, and confirmed by measuring the mass of the dried resin. Typical monomer loading was calculated to be 0.4-0.8 mmol/g.

Microwave-Assisted Solid-Phase Synthesis (18-24)

All solid-phase polyamide coupling reactions were performed manually on a Biotage Initiator Eight microwave synthesizer on a 200-500 mg scale of loaded resin. Prior to each monomer coupling reaction, the N-terminal Fmoc group was first removed in a piperidine solution. The Fmoc deprotections were performed in a fritted peptide synthesis vessel at room temperature, and the specific conditions for each N-terminal monomer are as follows:

N-Fmoc-Pyrrole / Imidazole: (a) swell resin in DCM; (b) wash with DMF; (c) add 50% piperidine in DMF; (d) shake suspension for 10 min; (e) wash with DMF; (f) repeat steps a-e twice.

N-Fmoc-GABA / β -Alanine: (a) swell resin in DCM; (b) wash with DMF; (c) add 50% piperidine in DMF; (d) shake suspension for 5 min; (e) wash with DMF; (f) repeat steps a-e once.

Following Fmoc removal, the resin was de-swelled in MeOH, washed with Et₂O, dried *in vacuo*, and transferred to a microwave synthesis vessel as a dry powder. The corresponding monomer acid (3 eq.) was activated with PyBOP (3 eq.) and DIEA (6 eq.) in DMF (0.3M concentration of monomer), and added to the resin. The coupling reactions were then set up in the microwave reactor at 50°C for the time durations described in **Table**

3.1. After the listed microwave-assisted coupling times, the reaction mixture was filtered into a peptide synthesis vessel, and the collected resin was washed with DMF (3x), MeOH (3x), Et₂O, and dried *in vacuo*. To ensure completion of each deprotection and coupling step, analytical HPLC spectra were taken by cleaving a small resin sample in 30% hexafluoroisopropanol (HFIP) in DCM.

The polyamide core was synthesized on 2-Cl-Trt resin in an iterative manner by repeating the deprotection-coupling procedures described above using the corresponding monomeric units. Upon completion of the sequence, 100-200 mg of the resin was suspended in 1 mL 30% HFIP in DCM and stirred for 1 hr to yield the crude N-terminal Fmoc-protected polyamide intermediate. The reaction mixture was then run through a cotton filter to remove the resin, and the filtrate was concentrated *in vacuo*. The residual oil was resuspended in 5 mL of a 1:1 MeOH:DCM mixture, and reconcentrated *in vacuo* to give an off-white/beige solid. To remove the N-terminal Fmoc group, the solid was redissolved in 800 μL DMF, followed by addition of 200 μL piperidine, and the solution was stirred for 30 min. Upon confirmation of complete deprotection by analytical HPLC, the solution was added to 4 mL 30% MeCN in 0.1% aqueous TFA. The precipitated 9-methylene fluorene side product was then removed by centrifugation, and washed twice with 2 mL 30% MeCN in 0.1% aqueous TFA. The combined aqueous solution was purified by reverse-phase HPLC and lyophilized to dryness to yield pre-cyclic polyamide intermediates **18-24**. All dried samples of **18-24** were stored at -20°C prior to DPPA-mediated macrocyclization.

Synthetic yields and MALDI-TOF Characterization Data for **18-24**:

(18): 12.6 μmol recovered (38.0 μmol theoretical, 33% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{70}\text{H}_{77}\text{N}_{22}\text{O}_{15}^+ = 1465.4$, observed = 1465.9.

(19): 15.2 μmol recovered (38.0 μmol theoretical, 40% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{62}\text{H}_{70}\text{N}_{21}\text{O}_{13}^+ = 1316.5$, observed = 1316.9.

(20): 12.8 μmol recovered (38 μmol theoretical, 34% yield). MALDI-TOF $[\text{M} + \text{Na}]^+$ calc'd for $\text{C}_{67}\text{H}_{79}\text{N}_{22}\text{NaO}_{15}^+ = 1453.6$, observed = 1453.9.

(21): 15.1 μmol recovered (49 μmol theoretical, 31% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{69}\text{H}_{76}\text{N}_{23}\text{O}_{15}^+ = 1466.5$, observed = 1466.9.

(22): 17.2 μmol recovered (51 μmol theoretical, 34% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{61}\text{H}_{69}\text{N}_{22}\text{O}_{13}^+ = 1317.5$, observed = 1317.2.

(23): 16.2 μmol recovered (51 μmol theoretical, 32% yield). MALDI-TOF $[\text{M} + \text{Na}]^+$ calc'd for $\text{C}_{66}\text{H}_{77}\text{N}_{23}\text{NaO}_{15}^+ = 1454.5$, observed = 1455.0.

(24): 8.2 μmol recovered (25 μmol theoretical, 33% yield). MALDI-TOF $[\text{M} + \text{Na}]^+$ calc'd for $\text{C}_{66}\text{H}_{78}\text{N}_{23}\text{NaO}_{15}^+ = 1454.5$, observed = 1454.9.

DPPA-mediated Macrocyclization (1-7)

The macrocyclization reactions were run on a 2 μmol to 16 μmol scale. Intermediates **18-24** were first dissolved in DMF (0.25 mM) in a round-bottom flask equipped with a magnetic stir bar, followed by addition of DIEA (200 eq.), and purged with argon for 15 min. Diphenylphosphoryl azide (DPPA) (50 eq.) was then added to the reaction mixture in a drop wise manner, while rapidly stirring. Upon full addition of the DPPA, the solution was allowed to react and stirred at room temperature for 16-20 hr. After

confirmation of reaction completion by analytical HPLC, the reaction mixture was concentrated *in vacuo*, and the resulting oil residue was dissolved in 3 mL MeCN and transferred to a 15 mL Falcon Tube. The MeCN was then removed with air flow and 3 mL 0.1% aqueous TFA was added to the remaining oil layer to yield an off-white suspension, which was isolated via centrifugation and lyophilized to dryness.

For reactions starting with **18**, **19**, **21**, and **22**, the lyophilized residue was submitted to 1 mL 10% trifluoromethanesulfonic acid (TFMSA) in TFA, stirred for 5 min, frozen in LN₂, and thawed by layering 1 mL DMF. For reactions starting with **20**, **23**, and **24**, the lyophilized residue was submitted to 1 mL neat TFA, stirred for 15 min, frozen in LN₂, and thawed by layering 1 mL DMF. All the thawed solutions were then diluted with 6 mL 0.1% aqueous TFA, purified by reverse-phase HPLC and lyophilized to dryness to yield cyclic polyamides **1-7**.

Synthetic yields and MALDI-TOF characterization data for **1-7**:

(1): 4.5 μmol recovered (12.9 μmol theoretical, 35% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{54}\text{H}_{63}\text{N}_{22}\text{O}_{10}^+$ = 1179.5, observed = 1179.9.

(2): 0.84 μmol recovered (2.0 μmol theoretical, 42% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{54}\text{H}_{62}\text{N}_{21}\text{O}_{10}^+$ = 1164.5, observed = 1164.6.

(3): 3.2 μmol recovered (6.7 μmol theoretical, 47% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{62}\text{H}_{69}\text{N}_{22}\text{O}_{12}^+$ = 1313.6, observed = 1314.0.

(4): 3.1 μmol recovered (8.0 μmol theoretical, 39% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{53}\text{H}_{62}\text{N}_{23}\text{O}_{10}^+$ = 1180.5, observed = 1180.9.

(5): 0.96 μmol recovered (2.0 μmol theoretical, 48% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{53}\text{H}_{61}\text{N}_{22}\text{O}_{10}^+$ = 1165.5, observed = 1165.5.

(6): 3.7 μmol recovered (10.0 μmol theoretical, 37% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{61}\text{H}_{68}\text{N}_{23}\text{O}_{12}^+$ = 1314.5, observed = 1314.5.

(7): 2.2 μmol recovered (5.7 μmol theoretical, 38% yield). MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{61}\text{H}_{68}\text{N}_{23}\text{O}_{12}^+$ = 1314.5, observed = 1314.8.

Selective Conjugation of Benzoic Acid Derivatives (9-11)

A solution of benzoic acid (3.0 mg, 0.025 mmol, 25 eq.) and PyBOP (13 mg, 0.025 mmol, 25 eq.) in DMF (0.5 mL) and DIEA (44 μL , 0.25 mmol, 250 eq.) was stirred at room temperature for 10 min. The activated solution was then added to **3** (1.0 μmol) and stirred for 3 hr. After confirmation of complete reaction by analytical HPLC, 12 mL cold Et_2O was added to the reaction mixture and cooled at -20°C for 16 hr. The precipitate was then isolated by centrifugation and allowed to air dry. The resulting residue was submitted to 1 mL 10% trifluoromethanesulfonic acid (TFMSA) in TFA, stirred for 5 min, frozen in LN₂, and thawed by layering 1 mL DMF. The thawed solution was then diluted with 6 mL 0.1% aqueous TFA, purified by reverse-phase HPLC and lyophilized to dryness to yield cyclic polyamides **9** (684 nmol, 68% yield). Using the same procedure described above, starting with **6** (1.60 μmol) and **7** (750 nmol), yielded mono-benzoyl substituted cyclic polyamides **10** (1.05 μmol , 67% yield) and **11** (367 nmol, 49% yield), respectively.

(9): MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{61}\text{H}_{67}\text{N}_{22}\text{O}_{11}^+$ = 1283.5, observed = 1284.1.

(10): MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{60}\text{H}_{66}\text{N}_{23}\text{O}_{11}^+$ = 1284.5, observed = 1284.5.

(11): MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{60}\text{H}_{66}\text{N}_{23}\text{O}_{11}^+$ = 1284.5, observed = 1284.9.

Synthesis of Cyclic Polyamide-Fluorescein Conjugates (12-14)

A solution of fluorescein isothiocyanate (FITC) (2.7 mg, 7.0 μmol , 25 eq.) in DMF (0.2 mL) and DIEA (12 μL , 0.07 mmol, 250 eq.) was added to **3** (0.28 μmol) and stirred for 2 hr. After confirmation of complete reaction by analytical HPLC, 12 mL cold Et_2O was added to the reaction mixture and cooled at -20°C for 16 hr. The precipitate was then isolated by centrifugation and allowed to air dry. The resulting residue was submitted to 1 mL 10% trifluoromethanesulfonic acid (TFMSA) in TFA, stirred for 5 min, frozen in LN₂, and thawed by layering 1 mL DMF. The thawed solution was then diluted with 6 mL 0.1% aqueous TFA, purified by reverse-phase HPLC and lyophilized to dryness to yield cyclic polyamide **12** (65 nmol, 23% yield). Using the same procedure described above, starting with **6** (0.40 μmol) and **7** (0.40 μmol), yielded cyclic polyamide-fluorescein conjugated **13** (345 nmol, 86% yield) and **14** (118 nmol, 29% yield), respectively.

(12): ESI-MS $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{75}\text{H}_{74}\text{N}_{23}\text{O}_{15}\text{S}^+$ = 1568.6, observed = 1568.3.

(13): ESI-MS $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{74}\text{H}_{73}\text{N}_{24}\text{O}_{15}\text{S}^+$ = 1569.5, observed = 1569.2.

(14): ESI-MS $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{74}\text{H}_{73}\text{N}_{24}\text{O}_{15}\text{S}^+$ = 1569.5, observed = 1569.3.

Synthesis of Cyclic Polyamide Targeted to 5'-WCGWGW-3' Sequence (8)

2-Cl-Trt-Cl resin (200 mg, 1.59 mmol/g) was first loaded with FmocPyImOH dimer (96 mg, 0.20 mmol), which was obtained from published procedures.¹³ Experimental details were analogous to the monomer loading protocol reported above. The obtained Fmoc-Py-Im-(2-Cl-Trt) resin (265 mg, 0.59 mmol/g) was subjected to the previously described microwave-assisted solid-phase synthesis conditions to build the corresponding

polyamide sequence. *A quarter* of the resin (0.15 μmol theoretical) was then cleaved and purified by reverse-phase HPLC to yield pre-cyclic polyamide intermediate **25** as an off-white powder (13.3 μmol , 34% yield). The isolated **25** (2.0 μmol) was subjected to DPPA-mediated macrocyclization conditions analogous to that for compounds **1-7**, with the Cbz groups removed with 10% TFMSA in TFA, and purified by reverse-phase HPLC to afford cyclic polyamide **8** (773 nmol, 39% yield).

(25): MALDI-TOF $[\text{M} - \text{CO}_2 + \text{H}]^+$ calc'd for $\text{C}_{65}\text{H}_{675}\text{N}_{22}\text{O}_{13}^+$ = 1371.6, observed = 1371.7.

(8): MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{61}\text{H}_{67}\text{N}_{22}\text{O}_{11}^+$ = 1129.5, observed = 1130.0

Synthesis of Hairpin Polyamide Targeted to 5'-WCGWGW-3' Sequence (17)

Fmoc-Py-Im-(2-Cl-Trt) resin, obtained via the same procedure as for **8**, was subjected to the previously described microwave-assisted solid-phase synthesis conditions to build the corresponding polyamide sequence with an N-terminal PyImOH cap. The resin (70 mg, 0.039 mmol) was then suspended in 2 mL 30% HFIP in DCM, stirred for 1 hr, filtered, washed, and concentrated to afford the crude C-terminal free acid. The residue was then dissolved in DMF (1.8 mL) and added drop wise to a pre-activated solution of 3,3'-diamino-N-methyldipropylamine (252 μL , 1.56 mmol, 40 eq.) and PyBOP (40 mg, 0.078 mmol, 2 eq.) in DMF (6.0 mL). After stirring for 1 hr and confirmation of complete conjugation by HPLC, the crude product was isolated by Et_2O precipitation, air dried, redissolved in 8 mL 15% MeCN: 0.1% aqueous TFA, and purified by reverse-phase HPLC to yield polyamide intermediate **26** as an off-white powder (11.8 μmol , 30% yield).

The isolated **26** (3.1 μmol) was dissolved in DMF (200 μL) and added drop wise to a pre-activated solution of isophthalic acid (12 mg, 0.072 mmol) and PyBOP (5 mg, 9 μmol) in DMF (800 μL) and DIEA (13 μL). After stirring for 1 hr and confirmation of complete conjugation by HPLC, the reaction mixture was then precipitated in Et_2O , isolated by centrifugation, and air-dried. The remaining residue was then subjected to 500 μL neat TFA, stirred for 15 min, frozen in LN_2 , thawed by addition of 1 mL DMF, diluted with 6 mL 0.1% aqueous TFA, and purified by reverse-phase HPLC to afford cyclic polyamide **8** (3.1 nmol, 86% yield). (Note: Extinction coefficient for compounds **26** and **17** were established as $40,000 \text{ M}^{-1} \text{ cm}^{-1}$ at $\lambda_{\text{max}} = 302\text{nm}$.)

(26): MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{58}\text{H}_{79}\text{N}_{22}\text{O}_{11}^+$ = 1259.6, observed = 1260.1.

(17): MALDI-TOF $[\text{M} + \text{H}]^+$ calc'd for $\text{C}_{61}\text{H}_{75}\text{N}_{22}\text{O}_{12}^+$ = 1307.6, observed = 1307.8.

Thermal Denaturation Analysis

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 MgCl_2 , and 5 mM CaCl_2 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 95°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 $\lambda = 260 \text{ nm}$ from 25 to 95°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.

Cell Culture

Cell lines were cultured at 37°C under 5% CO₂ using standard cell culture and sterile techniques. Cell medium was supplemented with 10% fetal bovine serum. Ham's F-12K (Kaighn's) medium was used for A549 cells, and RPMI 1640 was used for T47D cells.

Confocal Microscopy

For each experiment, cells were plated in 200 µL of the proper medium onto glass-bottom cell culture plates at a density of 1×10^5 (A549) cells/mL, or 1.5×10^5 cells/mL (T47D). Cells were grown for 24 hr, and media was replaced with fresh media containing polyamide to give a final DMSO concentration of 0.1%. Next, cells were incubated for 16 hr, followed by removal of media, washing, and addition of fresh media. Hoechst 33258 was added 2 hr prior to imaging. Imaging was performed at the Caltech Beckman Imaging Center using a Zeiss LSM 510 Meta NLO 2-photon inverted laser scanning microscope equipped with a 40x oil-immersion objective lens. Polyamide–fluorescein conjugates **12-14** were imaged in multi-track mode using 488 nm laser excitation at 15% output with a pinhole of 375 µm and a standard fluorescein filter set. Hoechst was imaged using 800 nm two-photon excitation with an HFT KP680 dichroic and a 390- to 465-nm bandpass filter with a fully open pinhole. All images were analyzed using Zeiss LSM Zen software.

Sulforhodamine B Cytotoxicity Assay

For cytotoxicity assays, cell lines were plated in 96-well cell culture plates in 100 µL media at a density of 1×10^4 cells/mL (A549), or 5×10^4 cells/mL (T47D). IC₅₀ values were determined using the sulforhodamine B (SRB) colorimetric assay as previously

described.³⁹ Cells were grown for 24 hr, before polyamides in 100 μ L media were added in serial dilution, in quadruplicate for each concentration. After incubation for 72 hr, cell media was replaced with 100 μ L fresh media, and cells were allowed to recover for an additional 24 hr. Cells were then fixed by adding 100 μ L of 10% trichloroacetic acid directly to each well and stored at 4°C for 1 hr, before being washed, dried, stained with 100 μ L 0.057% SRB solution per well for 30 min, and washed and dried again as described. After solubilizing the bound dye with 200 μ L of 10 mM Tris (pH 10.5) per well, absorbance at 490nm was measured on a PerkinElmer Victor microplate reader. The data are charted as a percentage of untreated controls, corrected for background absorbance. IC₅₀ is defined as the concentration that inhibits 50% of control cell growth. These values were determined by non-linear least-squares regression fit to $Y = A + (B - A) / (1 + 10^{((\text{Log EC}_{50} - X) \times H)})$, where $A = \text{max.}$, $B = \text{min.}$ and $H = \text{Hill Slope}$. All calculations were performed using Prism 4 (GraphPad) software. Stated IC₅₀ values represent the mean and standard deviation of three independent biological replicates.

Acknowledgements

This work was supported by the National Institutes of Health (GM27681). BCL is grateful to the California Tobacco-Related Disease Research Program (18DT-0015) for a dissertation award. DCM is thankful to the National Institutes of Health for a predoctoral research training grant (5T32GM007616). The authors would like to thank Dr. Amanda E. Hargrove for helpful discussions, and Dr. Jordan L. Meier for assistance with molecular modeling.

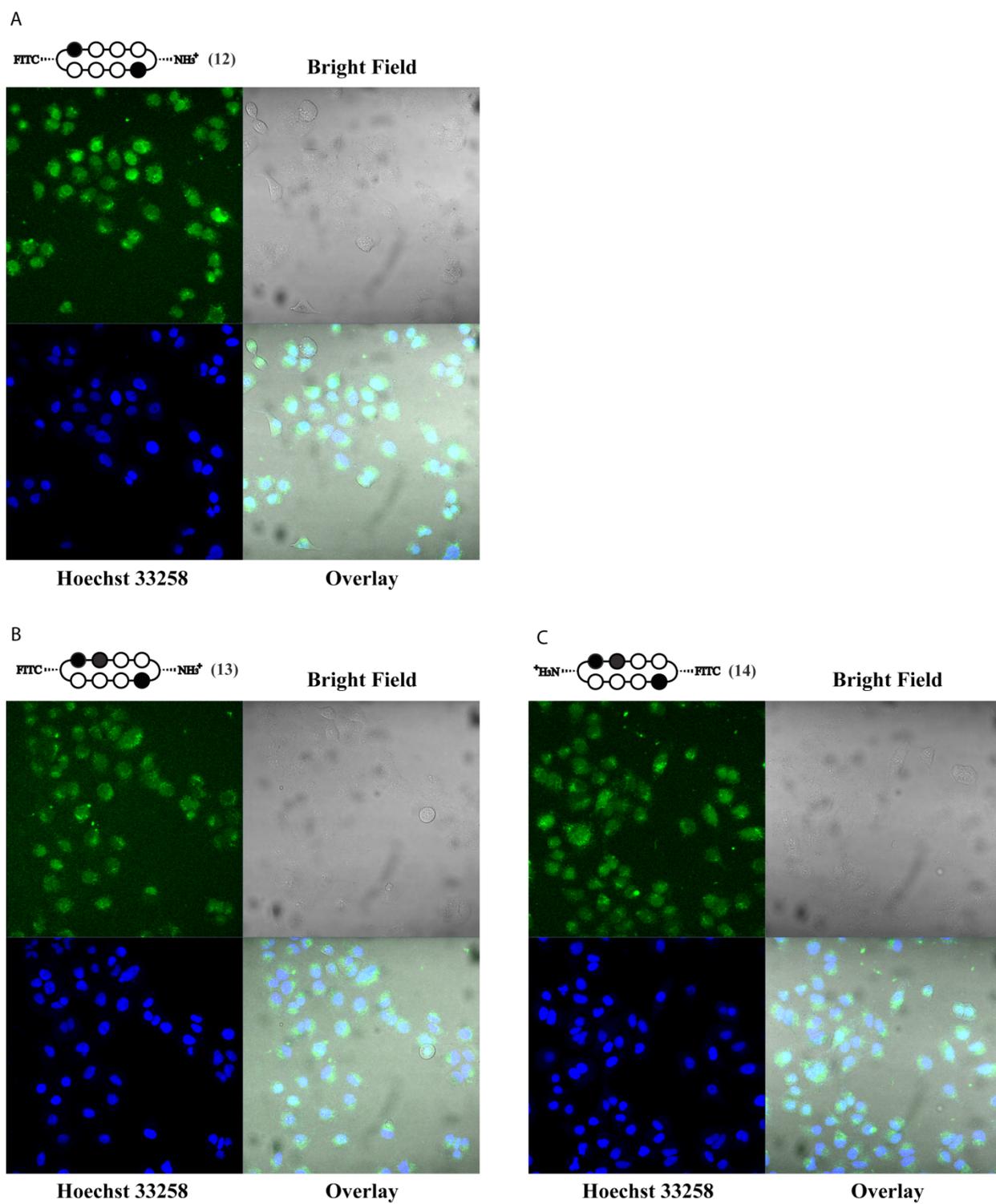


Figure S3.1: Confocal microscopy analysis of 12-14 in A549 cells with included bright field images.

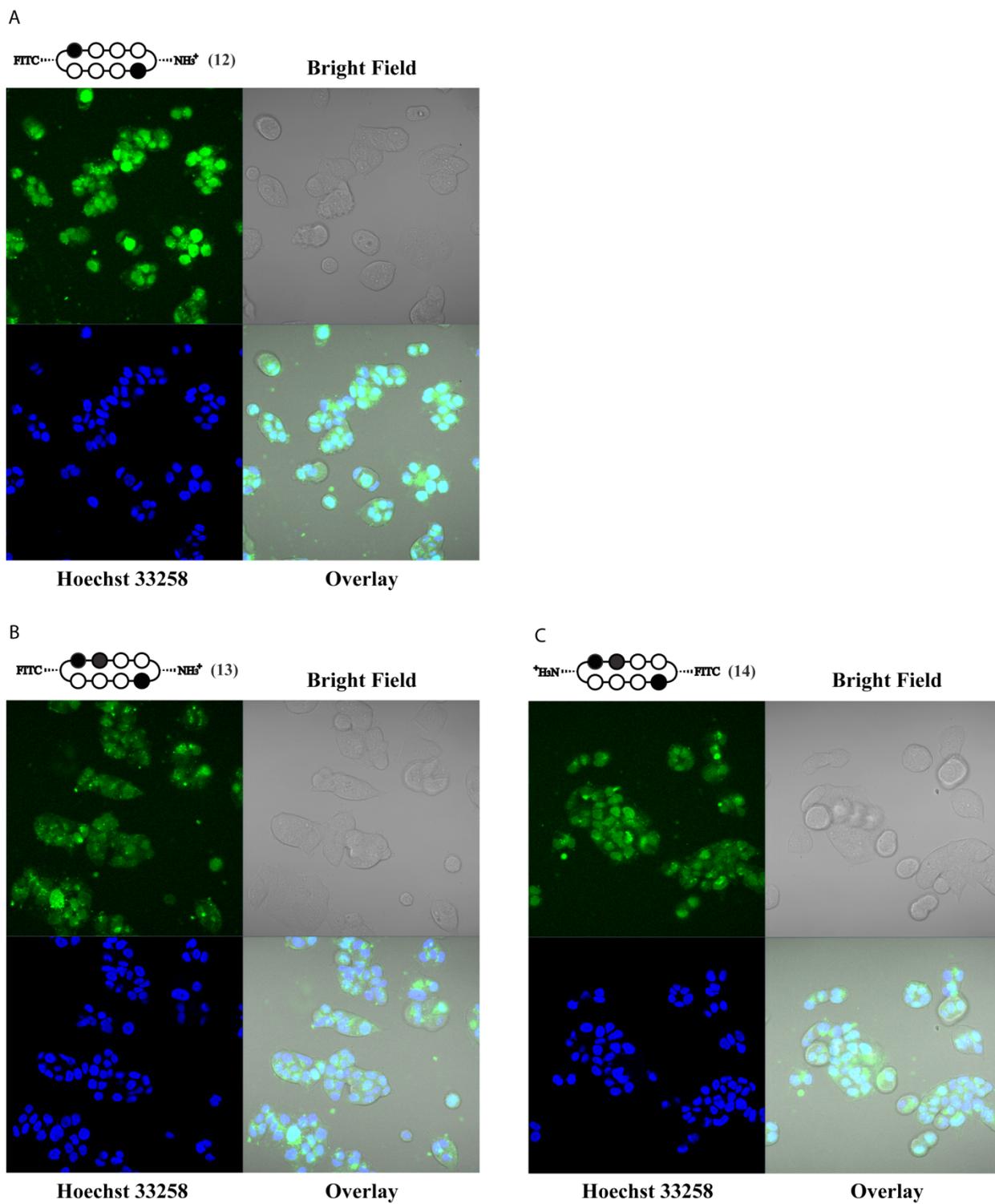


Figure S3.2: Confocal microscopy analysis of 12-14 in T47D cells with included bright field images.

References

1. Dervan, P. B., Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, 9, 2215–2235.
2. Trauger, J. W.; Baird, E. E.; Dervan, P. B., Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **1996**, 382, 559–561.
3. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **1998**, 391, 468–470.
4. Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C., Structural basis for G·C recognition in the DNA minor groove. *Nat. Struct. Biol.* **1998**, 5, 104–109.
5. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., A structural basis for recognition of A·T and T·A base pairs in the minor groove of B-DNA. *Science* **1998**, 282, 111–115.
6. Mrksich, M.; Parks, M. E.; Dervan, P. B., Hairpin peptide motif. A new class of oligopeptides for sequence-specific recognition in the minor groove of double-helical DNA. *J. Am. Chem. Soc.* **1994**, 116, 7983–7988.
7. Herman, D. M.; Baird, E. E.; Dervan, P. B., Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.* **1998**, 120, 1382–1391.
8. Hsu, C. F.; Phillips, J. W.; Trauger, J. W.; Farkas, M. E.; Belitsky, J. M.; Heckel, A.; Olenyuk, B. Z.; Puckett, J. W.; Wang, C. C. C.; Dervan, P. B., Completion of a programmable DNA-binding small molecule library. *Tetrahedron* **2007**, 63, 6146–6151.
9. Best, T. B.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA* **2003**, 100, 12063–12068.
10. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, 13, 284–299.
11. Hsu, C. F.; Dervan, P. B., Quantitating the concentration of Py-Im polyamide-fluorescein conjugates in live cells. *Bioorg. Med. Chem. Lett.* **2008**, 18, 5851–5855.
12. Olenyuk, B. Z.; Zhang, G. J.; Klco, J. M.; Nickols, N. G.; Kaelin Jr., W. G.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, 101, 16768–16773.
13. Kageyama, Y.; Sugiyama, H.; Ayame, H.; Iwai, A.; Fujii, Y.; Huang, L. E.; Kizaka-Kondoh, S.; Hiraoka, M.; Kihara, K., Suppression of VEGF transcription in renal cell carcinoma cells by pyrrole-imidazole hairpin polyamides targeting the hypoxia responsive element. *Acta Oncol.* **2006**, 45, 317–324.

14. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.* **2007**, *2*, 561–571.
15. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10418–10423.
16. Matsuda, H.; Fukuda, N.; Ueno, T.; Tahira, Y.; Ayame, H.; Zhang, W.; Bando, T.; Sugiyama, H.; Saito, S.; Matsumoto, K.; Mugishima H.; Serie, K., Development of gene silencing pyrrole-imidazole polyamide targeting the TGF-beta1 promoter for treatment of progressive renal diseases. *J. Am. Soc. Nephrol.* **2006**, *17*, 422-432.
17. Raskatov, J. A.; Meier, J. L.; Puckett, J. W.; Yang, F.; Ramakrishnan, P.; Dervan, P. B., Modulation of NF-kappaB-dependent gene transcription using programmable DNA minor groove binders. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1023–1028.
18. Cho, J.; Parks, M. E.; Dervan, P. B., Cyclic polyamides for recognition in the minor groove of DNA. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 10389-10392.
19. Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Cycle polyamide motif for recognition of the minor groove of DNA. *J. Am. Chem. Soc.* **1999**, *121*, 1121-1129.
20. Melander, C.; Herman, D. M.; Dervan, P. B., Discrimination of A/T sequences in the minor groove of DNA within a cyclic polyamide motif. *Chem. Eur. J.* **2000**, *6*, 4487-4497.
21. Zhang, Q.; Dwyer, T. J.; Tsui, V.; Case, D. A.; Cho, J.; Dervan, P. B.; Wemmer, D. E., NMR structure of a cyclic polyamide-DNA complex. *J. Am. Chem. Soc.* **2004**, *126*, 7958-7966.
22. Chenoweth, D. M.; Harki, D. A.; Phillips, J. W.; Dose, C.; Dervan, P. B., Cyclic pyrrole-imidazole polyamides targeted to the androgen response element. *J. Am. Chem. Soc.* **2009**, *131*, 7182-7188.
23. Morinaga, H.; Bando, T.; Takagaki, T.; Yamamoto, M.; Hashiya, K.; Sugiyama, H., Cysteine cyclic pyrrole-imidazole polyamide for sequence-specific recognition in the DNA minor groove. *J. Am. Chem. Soc.* **2011**, *133*, 18924-18930.
24. Meier, J. L.; Montgomery, D. C.; Dervan, P. B., Enhancing the cellular uptake of Py-Im polyamides through next-generation aryl turns. *Nucleic Acids Res.* **2012**, *40*, 2345-2356.
25. Taleb, R. I.; Jaramillo, D.; Wheate, N. J.; Aldrich-Wright, J. R., Synthesis of DNA-sequence-selective hairpin polyamide platinum complexes. *Chem. Eur. J.* **2007**, *13*, 3177-3186.
26. Van Holst, M.; Le Pevelen, D.; Aldrich-Wright, J. R., Terpyridineplatinum(II) incorporation in N-methylpyrrole-based polyamides by solid phase techniques. *Eur. J. Inorg. Chem.* **2008**, *29*, 4608-4615.

27. Puckett, J. W.; Green, J. T.; Dervan, P. B., Microwave assisted synthesis of Py-Im polyamides. *Org. Lett.* **2012**, 14, 2774-2777.
28. Weltzer, M.; Wemmer, D. E., Facile dimer synthesis for DNA-binding polyamide ligands. *Org. Lett.* **2010**, 12, 3488-3490.
29. Brady, S. F.; Varga, S. L.; Freidinger, R. M.; Schwenk, D. A.; Mendlowski, M.; Holly, F. W.; Veber, D. F., Practical synthesis of cyclic peptides, with an example of dependence of cyclization yield upon linear sequence. *J. Org. Chem.* **1979**, 44, 3101-3105.
30. Boger, D. L.; Johannes, D., Studies on the total synthesis of bouvardin and deoxybouvardin: cyclic hexapeptide cyclization studies and preparation of key partial structures. *J. Org. Chem.* **1988**, 53, 487-499.
31. Raskatov, J. A.; Hargrove, A. E.; So, A. Y. Dervan, P. B., Pharmacokinetics of Py-Im polyamides depend on architecture: cyclic versus linear. *J. Am. Chem. Soc.* **2012**, 134, 7995-7999.
32. Hargrove, A. E.; Raskatov, J. A.; Meier, J. L.; Montgomery, D. C.; Dervan, P. B., Characterization and solubilization of pyrrole-imidazole polyamide aggregates. *J. Med. Chem.* **2012**, 55, 5425-5432.
33. Synold, T. W.; Xi, B.; Wu, J.; Yen, Y.; Li, B. C.; Yang, F.; Phillips, J. W.; Nickols, N. G.; Dervan, P. B., Single-dose pharmacokinetic and toxicity analysis of pyrrole-imidazole polyamides in mice. *Cancer Chemoth. Pharm.* **2012**, 70, 617-625.
34. Baird, E. E.; Dervan, P. B., Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **1996**, 118, 6141-6146.
35. Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan P. B., Fmoc solid phase synthesis of polyamides containing pyrrole and imidazole amino acids. *Org. Lett.* **2001**, 3, 1201-1203.
36. Trauger, J. W.; Dervan, P. B., Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. *Methods Enzymol.* **2001**, 340, 450-466.
37. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J. Am. Chem. Soc.* **2009**, 130, 6859-6866.
38. Farkas, M. E.; Li, B. C.; Dose, C.; Dervan, P. B., DNA sequence selectivity of hairpin polyamide turn units. *Bioorg. Med. Chem. Lett.* **2009**, 19, 3919-3923.
39. Vichai, V.; Kirtikara, K., Sulforhodamine B colorimetric assay for cytotoxicity screening *Nat. Protoc.* **2006**, 1, 1112-1116.

Chapter 4

Characterization and Solubilization of Pyrrole–Imidazole Polyamide Aggregates

The text of this chapter was taken in part from a manuscript co-authored with Amanda E. Hargrove, Jevgenij A. Raskatov, Jordan L. Meier, and Peter B. Dervan (California Institute of Technology)

(Hargrove, A. E.; Raskatov, J. A.; Meier, J. L.; Montgomery, D. C.; Dervan, P. B. "Characterization and Solubilization of Pyrrole–Imidazole Polyamide Aggregates," *J. Med. Chem.* **2012**, 55, 5425-5432)

Abstract

To optimize the biological activity of pyrrole–imidazole polyamide DNA-binding molecules, we characterized the aggregation propensity of these compounds through dynamic light scattering and fractional solubility analysis. Nearly all studied polyamides were found to form measurable particles 50–500 nm in size under biologically relevant conditions, while HPLC-based analyses revealed solubility trends in both core sequences and peripheral substituents that did not correlate with overall ionic charge. The solubility of both hairpin and cyclic polyamides was increased upon addition of carbohydrate solubilizing agents, in particular, 2-hydroxypropyl- β -cyclodextrin (Hp β CD). In mice, the use of Hp β CD allowed for improved injection conditions and subsequent investigations of the availability of polyamides in mouse plasma to human cells. The results of these studies will influence the further design of Py-Im polyamides and facilitate their study in animal models.

Introduction

N-methylpyrrole (Py) and N-methylimidazole (Im) polyamides are heterocycle-based oligomers that bind the minor groove of DNA in a sequence-specific manner.¹⁻³ Investigations of Py-Im polyamide biological properties have demonstrated that these compounds are cell permeable,³⁻⁵ localize to the nucleus,³⁻⁵ and display transcriptional inhibition, likely through an allosteric mechanism by disrupting the transcription factor-DNA interface.^{6,7} This compression may be responsible for the observed reduction in transcription factor occupancy upon polyamide-DNA complexation.⁸⁻¹⁰ Gene regulation properties have been illustrated in cell culture models targeting transcription factors androgen receptor (AR),⁸ glucocorticoid receptor (GR),¹⁰ hypoxia inducible factor (HIF),^{9,11} nuclear factor kappaB (NF- κ B),¹² AURKA/AURKB,¹³ and TGF- β .^{14,15} We have investigated the utility of Py-Im polyamides in organismal models through *in vitro* ADMET studies,¹⁶ real-time biodistribution monitoring methods,¹⁷ and, most recently, the development of mouse pharmacokinetic and toxicity profiles.¹⁸ Recent efforts to develop more potent polyamides, however, have been hindered by poor solubility.¹⁹ These observations raise concerns about the likely aggregation of Py-Im polyamides. If aggregation is an issue, how does particle size correlate with structural features such as size, charge, shape, turn substitution, and Py/Im composition of the oligomer?

Recent studies of the aggregation of small molecule drug candidates through dynamic light scattering (DLS) and detergent-based assays have highlighted the importance of such considerations in drug design.²⁰⁻²³ Indeed, a screen of over 70,000 potential drug candidates by Shoichet and co-workers found that 95% of the initial hits acted as aggregate-based inhibitors.²⁴ At the same time, several currently approved drugs

can be classified as aggregate-based inhibitors²⁵ and in some cases aggregate particle size may be linked to pharmaceutical efficacy.^{23,26,27} We thus decided to investigate this important pharmacokinetic parameter and its relationship to the biological activity of Py-Im polyamides. As our laboratory explores the efficacy of polyamides in animal disease models, there becomes a pressing need to characterize the aggregation and solubility properties of these compounds as well as to investigate the use of formulating reagents to solubilize polyamides at the high concentrations required for animal injections.

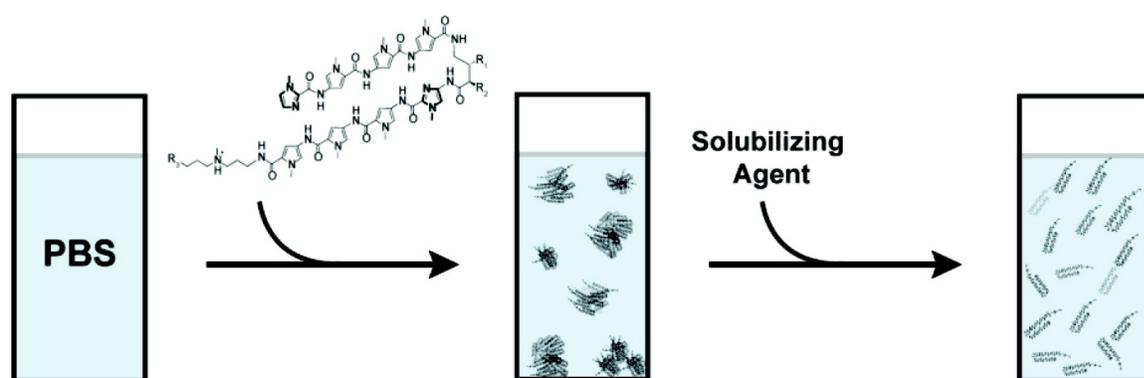


Figure 4.1: The use of solubilizing agents can disrupt the aggregation of polyamides and allow increased solubility.

Results and Discussion

Selection of a Panel of Polyamides

We selected two libraries of Py-Im hairpin polyamides, **1-6** and **7-11**, targeting the AR/GR consensus sequence 5'-WGWCW-3' (W = A/T)^{8,28} or the NF-κB consensus sequence 5'-WGGWW-3', respectively (Figure 4.2).¹²

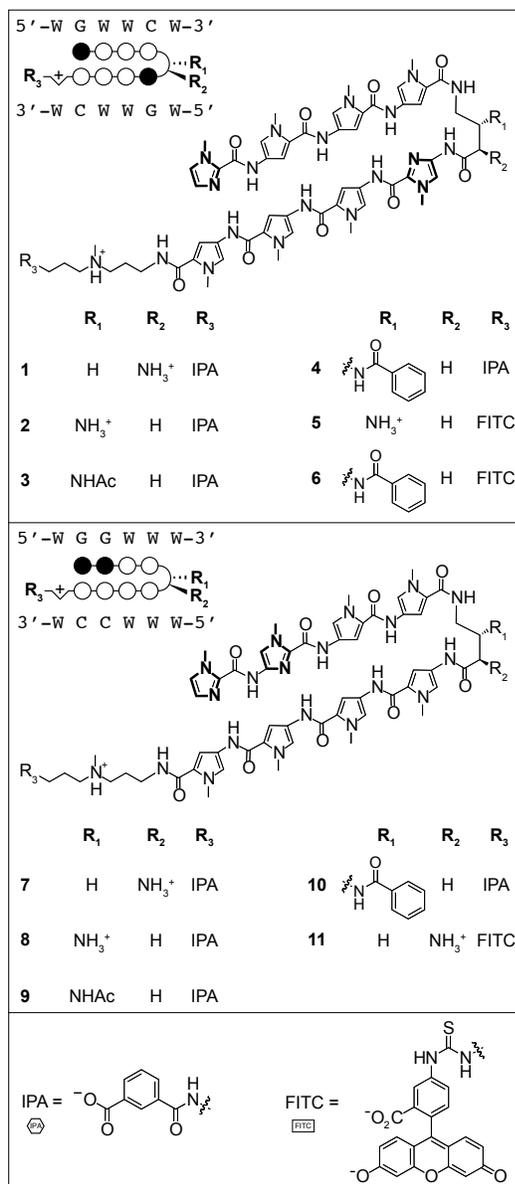


Figure 4.2 Chemical structures of hairpin polyamide library along with the corresponding circle-stick models and target DNA sequences. Legend: Black circle = Im; white circle = Py; semi-circle = γ -aminobutyric acid unit with dashed (R, β) or wedge (R, α) substituents; hexagon = isophthalic acid / IPA; rectangle = fluorescein / FITC; W = A/T bases.

These two different heterocyclic cores were diversified with a variety of substituents at the 4-aminobutyric acid (GABA) turn positions (R_1 , R_2) and the C-terminal tail position (R_3). Substituents at these positions are known to affect DNA binding and biological activities.^{19,29,30} In addition, the employment of cyclic polyamide architectures has resulted in increased DNA binding affinities and selectivities³¹ as well as improved efficacy against AR-regulated genes.²⁸ Taken in context with our recent finding of increased murine toxicity of cyclic polyamides,¹⁸ we decided to also investigate the properties of cycles **12-14** (Figure 4.3). All Py-Im polyamides were synthesized according to previously published solid-phase procedures.^{12,18,19,32}

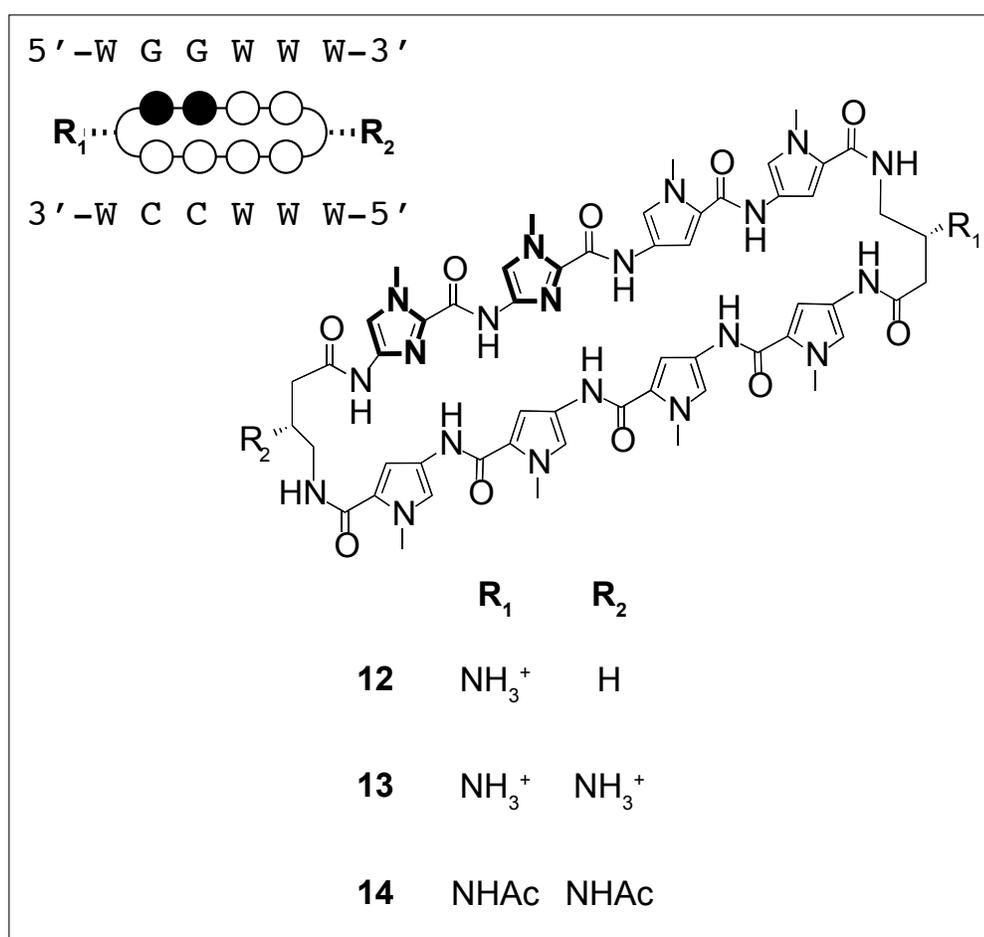


Figure 4.3: Chemical structures of cyclic polyamide library along with the corresponding circle-stick model and target DNA sequence. Legend: black circle = Im; white circle = Py; semi-circle = γ -aminobutyric acid unit with dashed (R , β) substituents; W = A/T bases.

Dynamic Light Scattering

The aggregation propensity of Py-Im polyamides was investigated through dynamic light scattering (DLS). Compounds **1-14** were studied at 1, 4, and 10 μM concentrations in a 0.1% DMSO/PBS solution in order to approximate the DMSO concentration and salt content present in cell culture experiments. Stock solutions (1000x) of each polyamide in DMSO were rapidly mixed with PBS, and the scattered light intensity was measured over the course of 10 minutes. The minimum concentration at which each compound was found to give a significant signal intensity (3x the buffer signal as per manufacturer guidelines), along with the respective particle sizes derived from a cumulant fit of the autocorrelation functions, are listed in Table 4.1.³³ Hairpin polyamides generally formed particles with radii of 70-200 nm at 4 μM concentration. One notable trend is that the benzamide substituted compounds (**4** and **10**) formed measurable particles at lower concentrations (1 μM vs. 4 μM) than their free amine counterparts (**2** and **8**). Polyamides containing fluorescein substituents (**5** and **6**) formed significantly larger particles when compared to the isophthalic acid conjugates (**2** and **4**), and compound **11** precipitated from the solution before particle size could be determined. Interestingly, cyclic polyamides **12-14** formed larger particles than the hairpin polyamides, with the bis- β -amino substituted cycle **13** forming the largest particles in this data set. Similar results were observed for a number of additional polyamides (Table S4.1), except for compounds containing three or four consecutive imidazole rings (**17**, **18**), which formed particles too large to be accurately measured (radii > 1 μM).

		Conc. (μM)	Radius (nm) ^a		Conc. (μM)	Radius (nm) ^a
1		4	182 ± 38	8	4	108 ± 17
2		4	104 ± 11	9	4	100 ± 1
3		4	74 ± 7	10	1	123 ± 28
4		1	87 ± 27	11	4	176 ± 41
		4	127 ± 32	11	-	n.d. ^b
5		1	294 ± 67	12	4	197 ± 28
6		1	147 ± 40	13	4	347 ± 113
7		4	124 ± 19	14	4	182 ± 14

Table 4.1: Estimated radii of polyamide aggregate particles at the concentration of minimum signal in 0.1% DMSO/PBS at 25 °C. ^aRadii derived from a cumulant fit of the average autocorrelation functions collected over 10 min. Errors represent standard deviation of at least three independent measurements. ^bRadius could not be determined due to rapid precipitation of the compound at 1 μM concentration.

Fractional Solubility Analysis

Next, the macroscopic solubility properties of these compounds were investigated by measuring the concentration of selected Py-Im polyamides in the soluble fraction of solutions with similar maximum concentrations (4 μM , Figure 4.4). Each compound was added as a 1000x (4 mM) stock in DMSO to PBS in accord with the light scattering experiments. Solutions were sonicated and then allowed to equilibrate for 2 hours at room temperature before aggregates were removed through centrifugation. In order to measure the concentration of polyamide in the supernatant, a plot of HPLC peak area vs. concentration was generated using polyamide 7 with detection at 310 nm, the wavelength at which each compound was quantified (Figure S4.1).

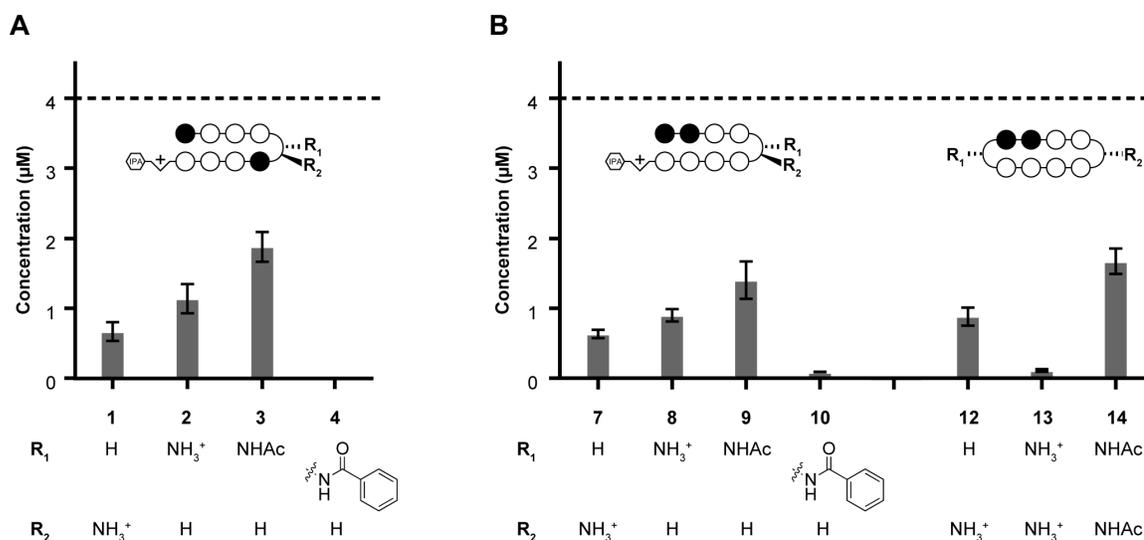


Figure 4.4: Calculated soluble concentration of select polyamides in 0.1% DMSO/PBS at 25°C. Maximum concentration estimated at 4 μM (dotted line) based on quantitation of starting material in 0.1% DMSO/water. Resultant concentrations determined by HPLC peak area at 310 nm detection after comparison with a standard curve (see SI). Error bars represent standard deviation of at least three independent measurements.

In general, the 5'-WGWWCW-3'-targeted hairpin polyamides (**1-4**) were found to be more soluble than their 5'-WGGWWW-3'-targeted counterparts (**7-10**). Within each set of polyamide hairpin cores, a relationship between turn substituents and solubility was observed. Polyamides with β -amine turns (**2, 8**) were found to be more soluble than those compounds with α -amine-substituted turns (**1, 7**), and the former compounds were further solubilized upon acetylation (**3, 9**). Increased solubility upon incorporation of acetylated turn units was also observed in the cyclic architecture (**13** vs. **14**). The benzamide-substituted compounds (**4, 10**) and the bis- β -amine substituted cycle **14** were found to be the least soluble, in good agreement with the light scattering measurements. Interestingly, none of the studied polyamides were fully soluble under these conditions.

Addition of Formulating Reagents

We thus decided to investigate the ability of known formulating reagents, in particular cyclodextrins (CDs),³⁴ to decrease aggregation and/or precipitation among the less soluble hairpin polyamides **7-10** (Figure 4.5). 2-Hydroxypropyl- β -cyclodextrin (Hp β CD) was chosen initially due to its high water solubility and low animal toxicity.³⁵ Using procedures identical to the solubility analyses, peak areas of Py-Im polyamides in the presence of 0, 5, or 50 mM Hp β CD were measured. A cyclodextrin-dependent increase of soluble polyamide concentration was observed for all compounds studied, with compounds **7** and **8** near the maximum expected concentration in solutions of 50 mM Hp β CD (Figure 4.5A). Impressively, 50 mM Hp β CD increased the concentration of the least soluble derivative (benzamide-substituted polyamide **10**) over 50-fold.

Surprisingly, the soluble concentration observed for polyamide **9** was significantly higher than expected based on the quantitation of the corresponding DMSO stock solution. This result likely derives from aggregation and/or precipitation of compound **9** upon dilution of the polyamide in water before the absorbance is measured, resulting in an underestimation of the stock concentration. We further probed the specificity of these effects by studying the solubilization of polyamide **7** by other carbohydrate formulating reagents, namely α -CD, γ -CD, hydroxypropyl methylcellulose (hypromellose), and dextrose (Figure S4.2). The three cyclodextrin derivatives were studied at 5 mM concentrations while hypromellose, a linear substituted glucose polymer, and dextrose, the glucose monomer, were normalized for total sugar content against 5 mM Hp β CD. In addition to Hp β CD, polyamide **7** was solubilized by γ -CD and hypromellose (Figure 4.5B).

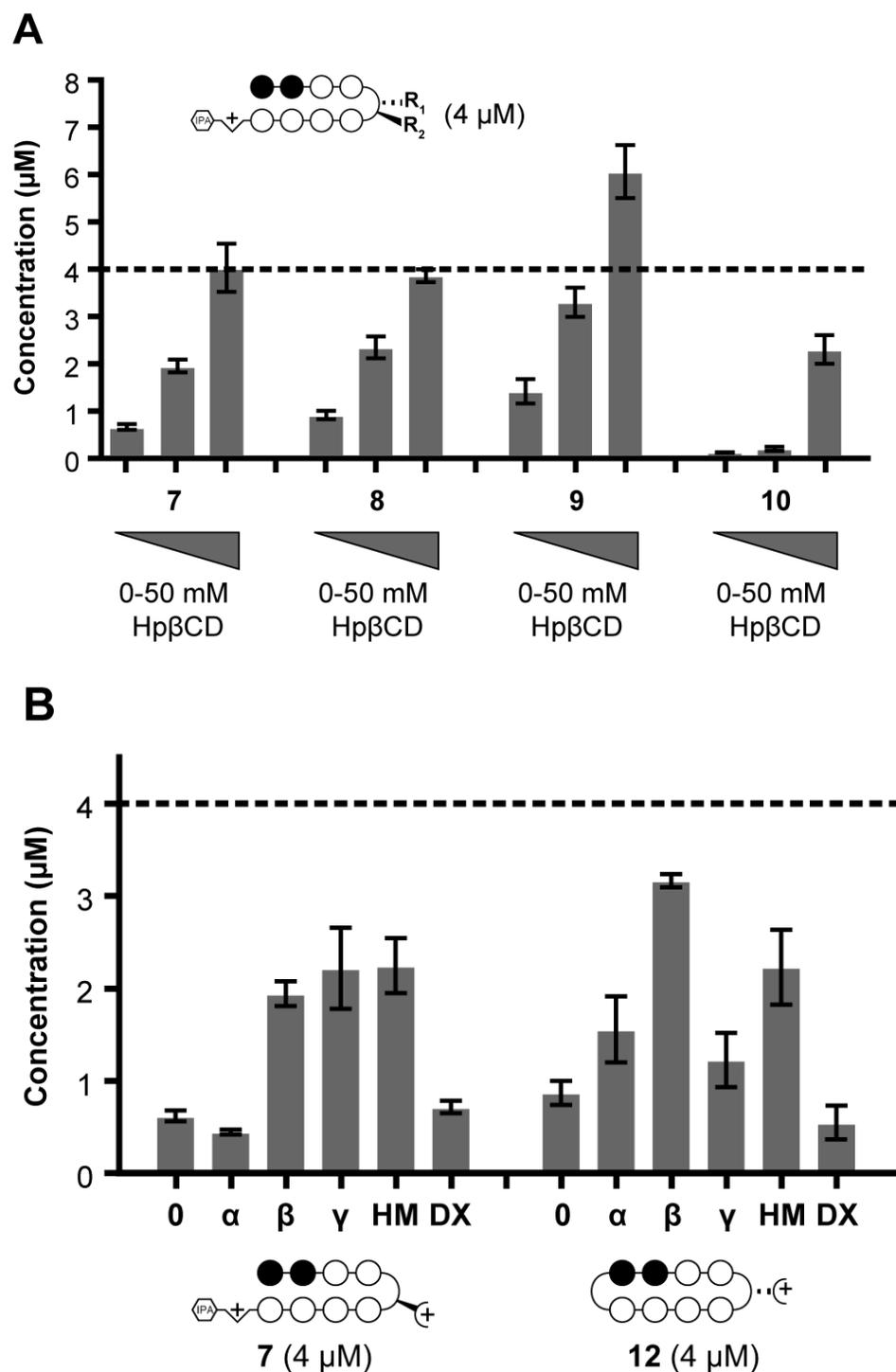


Figure 4.5: A) Calculated soluble concentration of polyamides 7 - 10 in 0.1% DMSO/PBS containing 0, 5, or 50 mM HP β CD at 25°C. B) Calculated soluble concentration of polyamides 7 and 12 in 0.1% DMSO/PBS containing: 5 mM α -, Hp β -, γ -cyclodextrin (α , β , γ , respectively); 6 mg/mL hypromellose (HM); 35 mM dextrose (DX). Maximum concentration estimated at 4 μM (dotted line) based on quantitation of starting material in 0.1% DMSO/water. Resultant concentrations determined by HPLC peak area ($\lambda = 310$ nm) after comparison with a standard curve (see SI). Error bars represent standard deviation of at least three independent measurements.

Polyamide **12**, which would seem less likely to form an inclusion complex with cyclodextrin due to its cyclic form, was screened against the same formulating agents. Hp β CD and hypromellose also solubilized cyclic compound **12**. Notably, neither polyamide displayed an increased solubility in the monomer (dextrose) solution.

Mouse Model Experiments

The utility of these results was further probed in an animal model system. Our laboratory recently found that high blood levels of polyamide **7** can be achieved in mice following an intraperitoneal (IP) injection of 120 nmol compound in a vehicle of 20% DMSO/PBS (600 μ M concentration, Figure 4.6A).¹⁸ Using Hp β CD, the DMSO content could be reduced to 1% with no loss in solubility. IP injections of 120 nmol polyamide **7** in a 1% DMSO/80 mM Hp β CD / PBS vehicle and the subsequent blood collection were performed under identical conditions to those previously reported. After blood collection, the plasma was isolated through centrifugation and the bulk proteins removed through methanol precipitation. The supernatant was then mixed with dilute aqueous trifluoroacetic acid (TFA) and a reference compound in acetonitrile was added. The injection vehicle containing Hp β CD yielded circulating polyamide concentrations comparable to those previously reported (Figure 4.6B). In both cases, polyamide concentrations of 13-14 μ M were detected in mouse plasma 1.5 hours after injection, with no polyamide detected after 24 hours. Furthermore, FITC-labeled compound **11**, which formed a precipitate in 20% DMSO/PBS solutions, was fully solubilized upon addition of Hp β CD (80 mM), allowing the compound to be injected into mice.

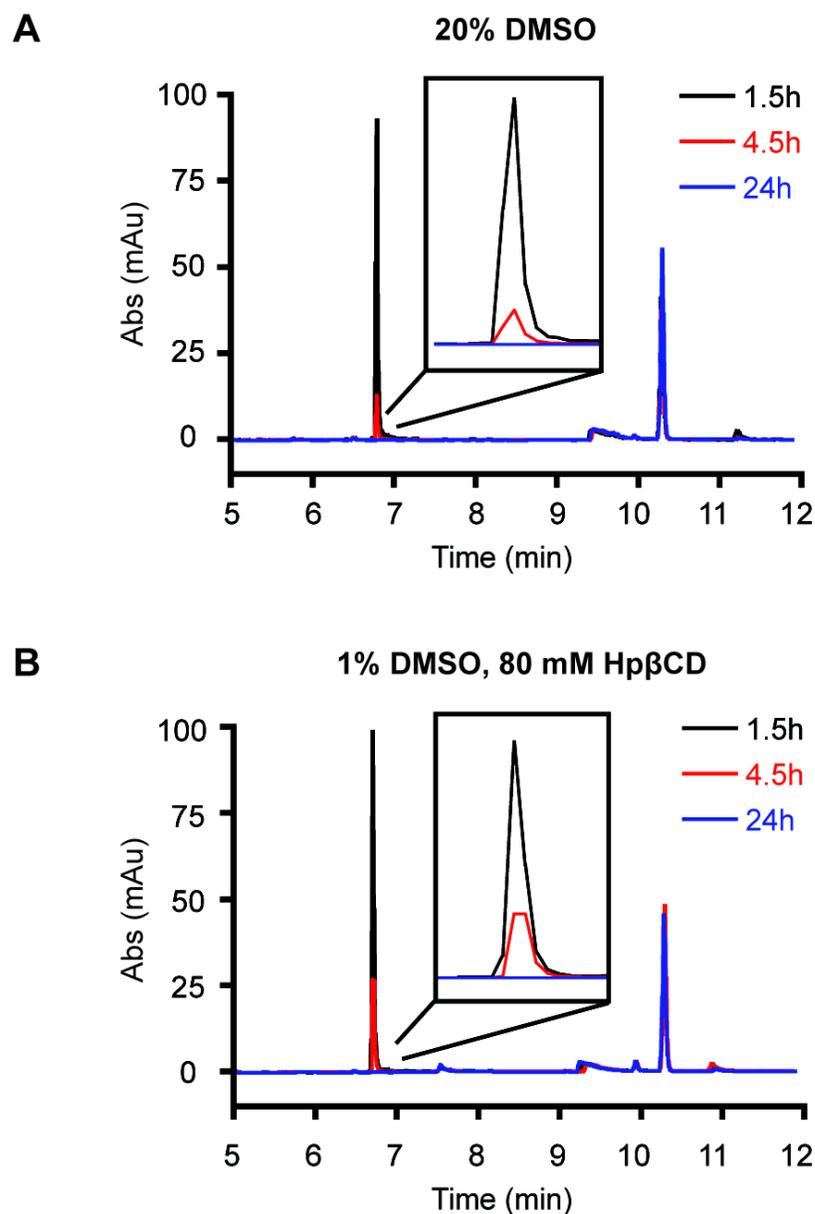


Figure 4.6: HPLC traces of mouse plasma isolated from four mice at three time points after injection with 120 nmol polyamide **7** in two different vehicles: 20% PBS / DMSO (A) and 1% DMSO/80 mM HP β CD / PBS (B).

Slightly reduced plasma concentrations of compound **11** were observed as compared to compound **7**, which may indicate reduced bioavailability of the FITC-modified polyamide (Figure 4.7A). We then sought to investigate the availability of the circulating polyamide to human cells by taking advantage of the nuclear staining generally observed with FITC-polyamide conjugates. Plasma samples isolated from mice

injected with compound **11** were added to A549 (human lung cancer) cells 16 hours prior to imaging live cells with confocal microscopy. In cells treated with plasma collected at 1.5 hours post-injection, strong nuclear fluorescent signals were observed (Figure 4.7B). Greatly reduced levels were observed with plasma isolated 4.5 hours post-injection, and no significant signal was observed with the addition of the 24 hours plasma sample (Figure 4.8).

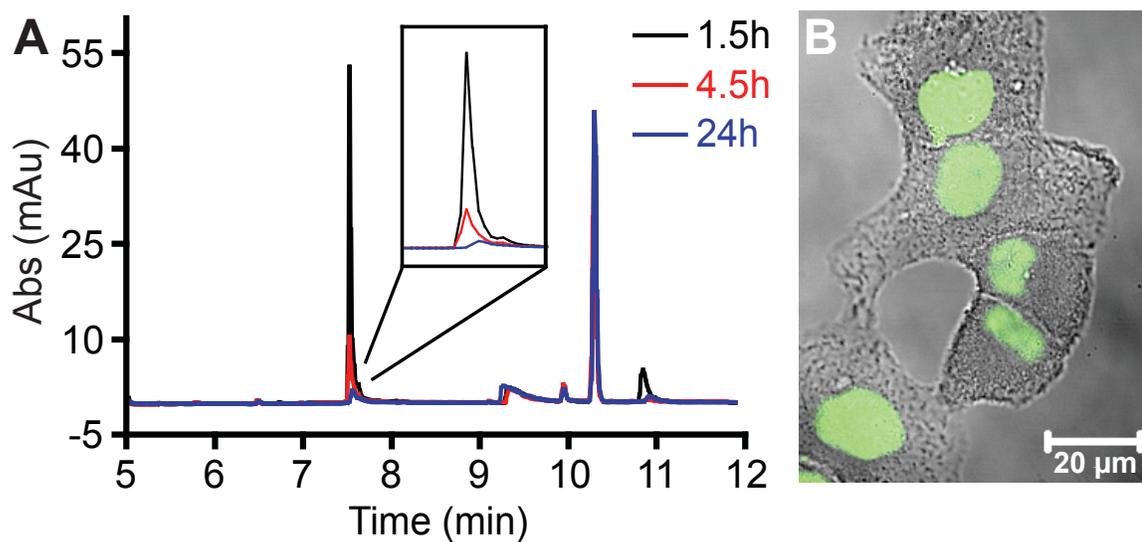


Figure 4.7: A) HPLC traces of mouse plasma isolated from four mice at three time points after injection with 120 nmol polyamide **11** in 20% DMSO/80 mM HP β CD / PBS. B) Confocal image of A549 cells after 16 hr incubation with mouse plasma isolated 1.5 hr after injection with polyamide **11**.

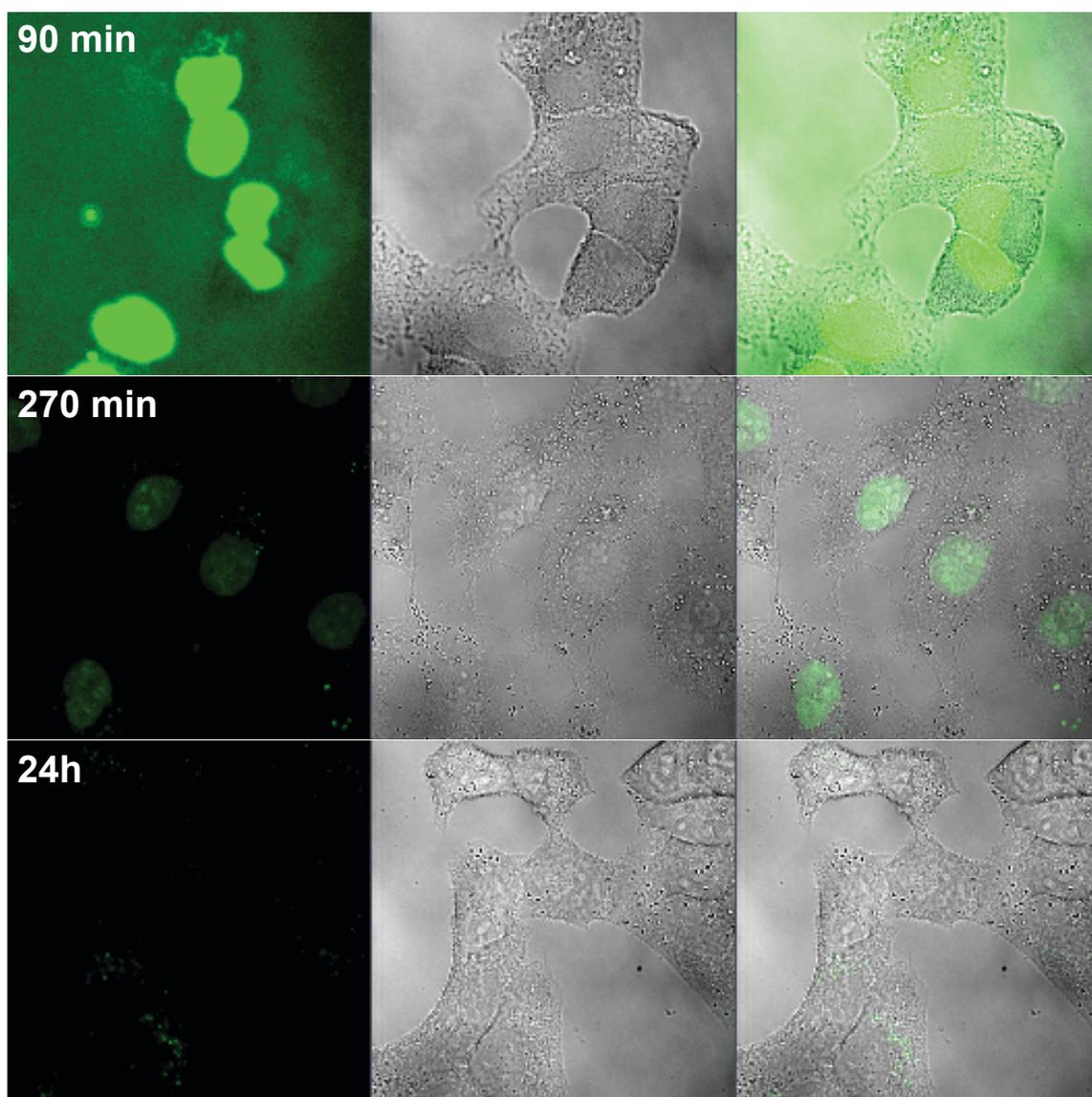


Figure 4.8: Confocal images of A549 cells after 16 hr incubation with mouse plasma isolated 90 min, 270 min, and 24 hr after injection with polyamide **11**. Images have been analyzed with identical intensity, brightness, and contrast settings to allow a direct comparison. Left: Fluorescence channel; Middle: Bright field image; Right: Overlay.

Conclusions

Dynamic light scattering measurements revealed that both hairpin and cyclic polyamides form measurable particles between 5 and 500 nm in size at biologically relevant concentrations (Table 4.1). Particle size was observed to be dependent on both cyclic vs. hairpin architecture and the terminal (tail) substituent. Interestingly, particles of similar size were observed for all polyamides containing an isophthalic acid (IPA) at the C-terminus despite the differing activities of these compounds in cell culture. These results support a mechanism of polyamide activity in which aggregation may not be a key factor.

Another interesting observation is that neither aggregation nor solubility is affected by the overall ionic charge of the polyamide. While organic compounds with ionizable groups are generally expected to be more soluble in aqueous salt solutions, neither light scattering nor solubility analyses revealed such a dependence. Indeed, hairpins and cycles in which the GABA amino turn units were modified with acetyl groups were found to be the most soluble.

At first glance, the lack of solubility observed for some polyamides is surprising as similar concentrations are commonly used in cell culture experiments, often without evidence of aggregation or precipitation. However, the experimental conditions required for the solubility experiments are a limited comparison to those in cell culture. For example, cell media generally contain a variety of small molecule and protein nutrients, and cell cultures are kept at higher temperatures (37 °C vs. 25 °C). In addition, the soluble fraction isolated by centrifugation is not necessarily representative of the available polyamide concentration during a typical cell incubation period (48-72 h),

particularly if aggregation is a dynamic process. We note that similar solubility problems have been reported by Sugiyama and co-workers, who enhanced the biological activity of seco-CBI polyamide conjugates through PEGylation³⁶ or liposomal formulations.³⁷

We were able to mitigate the problem of polyamide solubility through the addition of carbohydrate formulating reagents, in particular Hp β CD. As cyclodextrins are generally thought to form discrete inclusion complexes with small organic molecules,^{35,38} we postulated that the linear conformation of hairpin polyamides may be well solubilized by such additives. Indeed, the solubilization of polyamide **7** by the larger cyclodextrins (β and γ) is consistent with the formation of possible cyclodextrin inclusion complexes,³⁹ which in this case may result from interactions with the isophthalic acid unit at the C-terminus or the N-methylimidazole group at the N-terminus. Such interactions would not be expected, however, between cyclodextrins and cyclic polyamide **12**. Interestingly, solubilization of compound **12** was observed with Hp β CD but not the other cyclodextrin derivatives. The lack of solubilization with γ CD would be consistent with a model in which the interactions between Hp β CD and cycle **12** rely more on the hydroxypropyl substituents unique to Hp β CD, perhaps through additional hydrogen bonding interactions, rather than encapsulation. While both the linear **7** and cyclic **12** polyamide compounds were solubilized by hypromellose, presumably through encapsulation within the polymer matrix, it is notable that no significant solubilization was observed with the dextrose monomer. This latter observation may indicate the importance of an ordered carbohydrate structure, such as that available with cyclodextrins and hypromellose, for efficient polyamide solubilization. Further studies are necessary, however, before

conclusions can be drawn regarding the interactions between Py-Im polyamides and carbohydrate derivatives, and such investigations fall outside the scope of this work.

Further evidence of the utility of Hp β CD as a formulating reagent was gathered in mouse experiments. First, we demonstrated that the Hp β CD vehicle did not significantly affect circulating levels of polyamide **7**. On the other hand, the fluorescently labeled derivative **11** was only sufficiently soluble in Hp β CD solutions. As a result, hairpin **11** could be injected into mice using this vehicle. This tagged compound was of particular interest due to the high plasma protein binding levels (>99%) that had been previously reported for Py-Im polyamides during ADMET studies.²⁸ Evidence of nuclear uptake was observed in A549 cells following incubation with plasma from hairpin **11**-treated mice, thus demonstrating the availability of circulating polyamides to human cancer cells.

In summary, these studies have provided evidence that the aggregation propensity of Py-Im polyamides likely does not contribute to biological activity and may not be a critical concern in pharmacokinetic analyses. Solubility experiments revealed important trends, such as the increased solubility achieved by acetylation of the GABA amino turn unit, which will impact the design of next generation polyamides. Furthermore, the identification of an effective delivery vehicle will allow for the *in vivo* study of otherwise inaccessible Py-Im polyamides. These studies represent a valuable contribution to the field of small molecule transcriptional inhibitors and their ultimate utility as tools for perturbing gene expression networks *in vivo*.

Materials and Methods

Synthesis of Hairpin Py-Im Polyamides (1-11, 15-32)

The synthesis of Py-Im polyamides has been extensively described in previous work^{5,12,19,29,32} and is summarized as follows: Reagents were purchased from Sigma-Aldrich or Novabiochem. Py-Im cores were synthesized on Kaiser oxime resin using Boc-based chemistry, cleaved using 3,3'-diamino-N-methyldipropylamine, and purified by reverse phase preparative HPLC. The C-terminal amine was then derivatized with either isophthalic acid (IPA) or fluorescein isothiocyanate (FITC) and the crude intermediate isolated through ether precipitation. The GABA turn protecting groups (α -NHBoc or β -NHCBz) were removed under acidic conditions. If applicable, the crude intermediate was again isolated through ether precipitation and further derivatized at the GABA turn amine with either acetic anhydride or PyBOP-activated benzoic acid. Final products were purified through reverse phase HPLC and the identity confirmed through matrix-assisted laser desorption ionization – time-of-flight (MALDI-TOF) mass spectrometry. The synthesis and characterization of compounds **1**,¹⁰ **2-6**,¹⁹ **7**, **11**,¹² **15**, **25-32**¹⁹ were in line with literature reports. Results from MALDI-TOF characterization for compounds **8-10**, **12-14**, and **16-24** are available (Table S4.2).

Synthesis of Cyclic Py-Im Polyamides (12-14)

The synthesis and characterization of polyamides **12** and **13** have been previously described.¹⁸ In brief, the heterocyclic cores of these polyamides were synthesized on Kaiser oxime resin as above, except that a terminal GABA turn unit (Boc-GABA-OH or (R)-4-(Boc-amino)-3-(Z-amino)butyric acid) was added. Following deprotection of the

terminal Boc unit, the core was cleaved from the resin with DBU / H₂O and the resulting acid purified by reverse phase HPLC. The precursor acid was then cyclized using diphenylphosphorylazide under basic conditions. The crude intermediate was isolated through ether precipitation and the Cbz group(s) removed as above. Polyamides **12** and **13** were then isolated through reverse phase preparative HPLC. Polyamide **14** was synthesized by reaction of **13** with acetic anhydride under basic conditions and then purified by reverse phase HPLC. Results from MALDI-TOF characterization for compound **14** are available (Table S4.2).

Polyamide Quantification

Polyamide concentrations were measured by UV-absorption analysis on an Agilent 8453 diode array spectrophotometer in distilled and deionized water containing up to 0.1% DMSO using a molar extinction coefficient (ϵ) of 69,500 M⁻¹cm⁻¹ at 310 nm.

Dynamic Light Scattering

DMSO and PBS were passed through a 0.02 μ M syringe filter (Whatman) immediately prior to use. Stock solutions of each polyamide in DMSO were quantified as above and the purity determined by HPLC to be greater than 95%. Solutions of 1, 4, and 10 mM in DMSO were prepared and then centrifuged for 15 min at 16 x g to remove particulates. Immediately before measurement, 0.5 μ L of the DMSO stock was added to 500 μ L of PBS in a microcentrifuge tube. The solution was mixed briefly with a pipette tip and transferred to a disposable plastic cuvette (Fisher). Measurements were performed on a Wyatt Dynapro Nanostar instrument using a 659 nm / 100 mW laser at

100% power and a 90 degree detection angle at 25°C. Acquisition times of 10-15 sec were collected over 10 minutes and analyzed using the cumulant fit tool in the Dynamics (6.11.1.3) software with PBS as the referenced solvent. Acquisitions in which the baseline value of the fit was greater than ± 0.1 were omitted and the remaining traces averaged. Measurements in which the intensity (cts/s) was less than 3x the buffer signal intensity were considered below the detection limit.

Solubility Analysis

Stock solutions of each polyamide in DMSO were quantified as above and the purity checked by HPLC. Solutions of 4 mM stock were prepared in DMSO. Polyamide (0.5 μ L) was added to 500 μ L PBS in a microcentrifuge tube, and the solution was immediately vortexed and placed in a sonicating water bath at 25°C for 20 min. The tubes were then removed from the bath and allowed to equilibrate for 2 hr at room temperature. Samples were centrifuged for 20 min at 16 xg and 100 μ L of the supernatant removed for HPLC analysis. Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250 x 4.6 mm, 5 μ m) and a diode array detector (Mobile phase: 10-80% CH₃CN in 0.1% CF₃CO₂H (aqueous) over 17.5 min; Flow rate: 1.50 mL / min; Injection volume: 40 μ L). Peaks were detected and integrated at 310 nm absorbance using the Karat32 software. Sample concentrations were determined through comparison to a standard curve of concentration vs. peak area that was generated using compound 7 (Figure S1). Solubilization by formulating agents proceeded similarly except that the DMSO stock solutions were added to PBS containing 5 or 50 mM Hp β CD, 5 mM α CD, 5 mM γ CD,

35 mM dextrose, or 6.00 mg / mL hypromellose (approximately 35 mM relative glucose units based on reported substitution for Aldrich lot #128k0214v).

Animal Experiments

Murine experiments were performed as described previously.¹⁸ In brief, C57bl/6 mice (8 to 12 weeks of age, Jackson Laboratory) were injected intraperitoneally with 200 μ L of a PBS solution containing: a) 120 nmol compound **7**, 20% DMSO; b) 120 nmol compound **7**, 1% DMSO, 80 mM Hp β CD; or c) 120 nmol compound **11**, 20% DMSO, 80 mM Hp β CD. Blood was collected from anesthetized animals (2-5% isoflurane) by retro-orbital withdrawal. Immediately after the third blood draw, animals were euthanized by asphyxiation in a CO₂ chamber (2 atm).

Plasma was isolated by centrifugation of the collected blood. The samples from the four replicate mice were combined at 5 μ L / sample, yielding 20 μ L combined plasma that was then treated with 40 μ L CH₃OH, vortexed, and centrifuged. Fifty μ L of the supernatant were combined with one equivalent of the HPLC loading solution (4:1 water/CH₃CN, 0.08 % CF₃CO₂H) containing Boc-Py-OMe (methyl 4-((tert-butoxycarbonyl)amino)-1-methyl-pyrrole-2-carboxylate) as an internal spike-in control. Analytical HPLC analyses were conducted with a Phenomenex Kinetex C18 analytical column (100 \times 4.6 mm, 2.6 μ m, 100 Å) and a diode array detector (Mobile phase: 5-60% CH₃CN in 0.1% (v/v) aqueous CF₃CO₂H over 12.5 min; Flow rate: 2.0 mL / min; Injection volume: 40 μ L). Peaks were detected and integrated at 310 nm absorbance, and sample concentrations were determined through comparison to the previously published standard curve for this column.¹⁸

Confocal Microscopy

For confocal microscopy experiments, A549 cells in F-12K medium supplemented with 10% FBS (1 mL, 100k cells / mL) were applied to culture dishes equipped with glass bottoms for direct imaging (MatTek). Cells were allowed to adhere for 18 hr in a 5% CO₂ atmosphere at 37 °C. The medium was then removed and replaced with 200 µl of fresh medium supplemented with 20 µL of plasma collected 1.5, 4.5, or 24 hr after injection of compound **11**. After an additional 16 hr incubation period, 100 µL of untreated medium was added to each slide prior to imaging. Imaging was performed at the Caltech Beckman Imaging Center using a Zeiss LSM 5 Pascal inverted laser scanning microscope equipped with a 63x oil-immersion objective lens. Fluorescence and visible-light images were obtained using standard filter sets for fluorescein and analyzed using Zeiss LSM software.

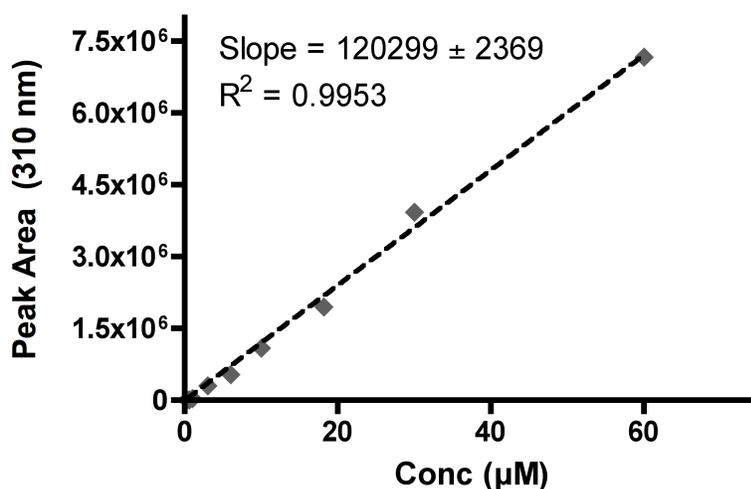


Figure S4.1. Plot of concentration vs. peak area plot derived from HPLC traces of compound **7** at 310 nm. Points were fit to a first order polynomial using the Prism software program (dotted line).

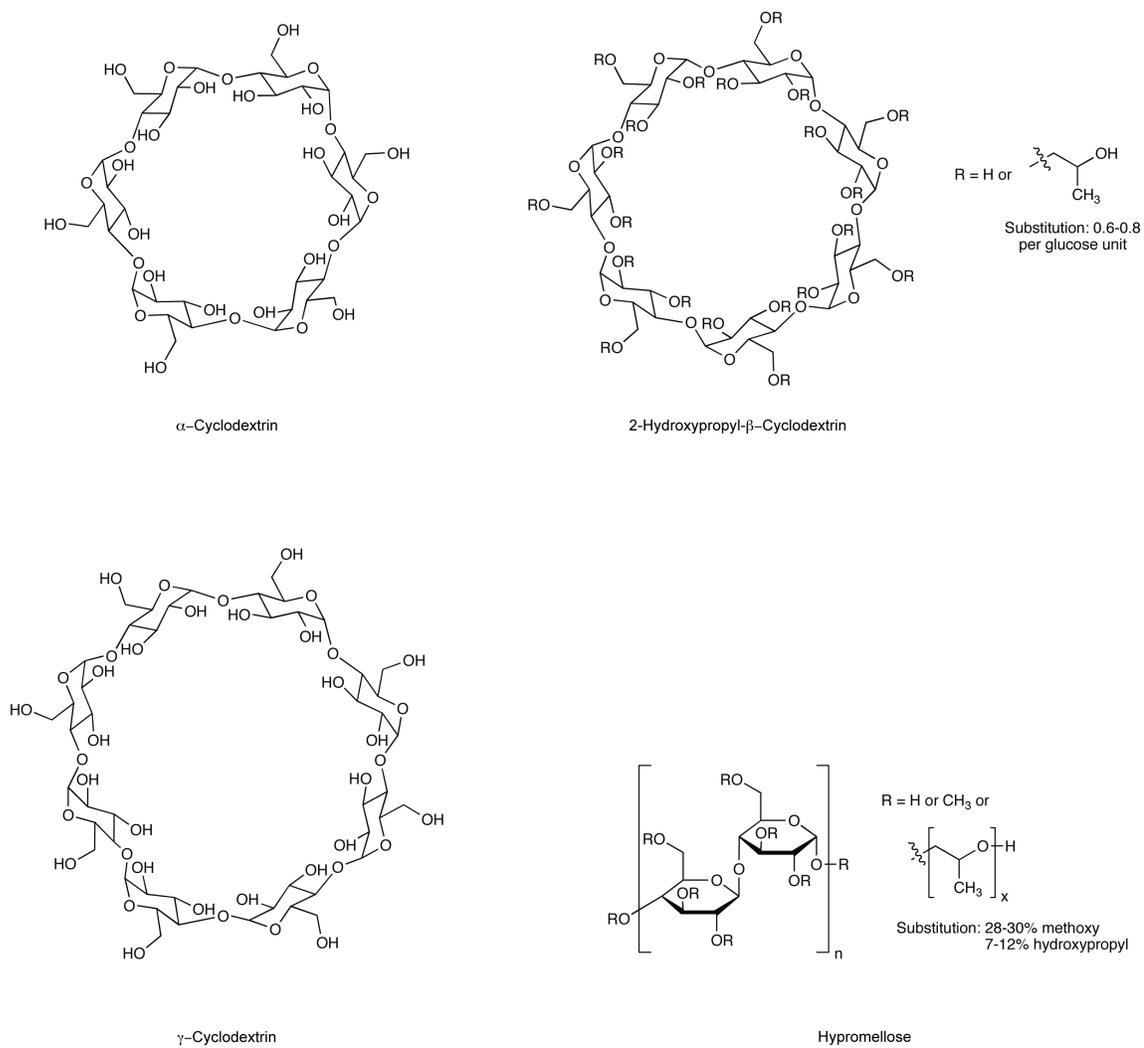


Figure S4.2. Structures of carbohydrate solubilizing agents.

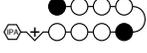
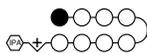
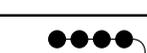
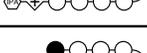
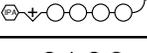
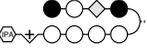
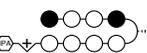
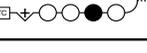
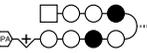
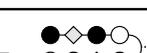
			Radius (nm) ^a		Intensity (cts/s)	
		Conc (μM)	Average	Std. Dev.	Average	Std. Dev.
0.1% DMSO / PBS	-	-	-	-	2.2E+05	7.9E+04
	15	4	157	33	4.6E+06	2.0E+06
	16	4	128	11	5.1E+06	2.2E+06
	17	4	> 1000 ^b	-	3.5E+06	1.2E+06
	18	4	> 1000 ^b	-	4.8E+06	1.2E+06
	19	4	90	20	1.5E+06	9.3E+05
	20	10	111	13	1.0E+06	7.4E+05
	21	4	108	2	6.1E+06	3.1E+06
	22	4	123	8	6.0E+06	4.5E+04
	23	4	189	2	7.9E+06	2.4E+05
	24	10	100	12	8.7E+05	1.3E+05
	25	1	344	41	4.4E+06	1.4E+05
	26	1	112	8	8.7E+06	2.3E+05
	27	1	119	10	7.7E+06	5.3E+05
	28	1	132	5	2.4E+06	6.8E+05
	29	1	100	1	1.9E+06	1.9E+05
	30	4	101	4	5.8E+06	1.6E+06
	31	10	120	7	9.9E+05	2.9E+05
	32	10	99	2	7.7E+05	3.6E+04

Table S4.1: Estimated radii of polyamide aggregate particles at the concentration of minimum signal in 0.1% DMSO/PBS at 25 °C. *a.* Radii derived from a cumulant fit of the average autocorrelation functions Collected over 10 min. *b.* Compounds formed particles with radii too large to be accurately determined with this method (radius > 1 μM).

	Technique	Formula	Calculated Mass	Measured Mass (m/z)
8	MALDI-TOF	C ₆₅ H ₇₇ N ₂₂ O ₁₂	[M+H] ⁺ 1357.6	1357.8
9	MALDI-TOF	C ₆₇ H ₇₉ N ₂₂ O ₁₃	[M+H] ⁺ 1399.6	1399.8
10	MALDI-TOF	C ₇₂ H ₈₁ N ₂₂ O ₁₃	[M+H] ⁺ 1461.6	1461.8
14	MALDI-TOF	C ₅₈ H ₆₇ N ₂₂ O ₁₂	[M+H] ⁺ 1263.5	1263.9
16	MALDI-TOF	C ₆₆ H ₇₈ N ₂₁ O ₁₂	[M+H] ⁺ 1356.6	1356.6
17	MALDI-TOF	C ₆₄ H ₇₆ N ₂₃ O ₁₂	[M+H] ⁺ 1358.6	1358.8
18	MALDI-TOF	C ₆₃ H ₇₅ N ₂₄ O ₁₂	[M+H] ⁺ 1359.6	1360.0
19	MALDI-TOF	C ₆₆ H ₇₈ N ₂₁ O ₁₂	[M+H] ⁺ 1356.6	1356.8
20	MALDI-TOF	C ₆₂ H ₇₆ N ₂₁ O ₁₂	[M+H] ⁺ 1306.6	1306.9
21	MALDI-TOF	C ₆₂ H ₇₆ N ₂₁ O ₁₂	[M+H] ⁺ 1306.6	1307.3
22	MALDI-TOF	C ₆₅ H ₇₇ N ₂₂ O ₁₂	[M+H] ⁺ 1357.6	1357.8
23	MALDI-TOF	C ₆₅ H ₇₇ N ₂₂ O ₁₂	[M+H] ⁺ 1357.6	1358.0
24	MALDI-TOF	C ₆₂ H ₇₆ N ₂₁ O ₁₂	[M+H] ⁺ 1306.6	1306.9

Table S4.2. Mass spectroscopy results for unpublished compounds **8-10, 12-14, 16-24**.

Acknowledgements

We thank Dr. Jennifer Keefe (California Institute of Technology) and Sigrid Kuebler (Wyatt Technology Corporation) for assistance with DLS experimentation and data analysis and Dr. Adam Urbach (Trinity University) for helpful discussions. A.E.H. thanks the California Tobacco-Related Disease Research Program (19FT-0105) and the NIH (NRSA number 1F32CA156833) for postdoctoral support. J.A.R. is grateful to the Alexander von Humboldt foundation for the award of a Feodor Lynen postdoctoral fellowship. J.L.M. acknowledges the American Cancer Society for a postdoctoral fellowship (PF-10-015-01-CDD). D.C.M. thanks the National Institutes of Health for a Cellular, Biochemical, and Molecular Sciences Predoctoral Research training grant (5T32GM007616). This work was supported by the Ellison Medical Foundation (AG-SS-2256-09) and NIH grants (GM27681, GM51747).

References

1. Dervan, P. B., Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, 9, 2215-2235.
2. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, 13, 284-299.
3. Hsu, C. F.; Dervan, P. B., Quantitating the concentration of Py-Im polyamide-fluorescein conjugates in live cells. *Bioorg. Med. Chem. Lett.* **2008**, 18, 5851-5855.
4. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.* **2004**, 32, 2802-2818.
5. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res.* **2007**, 35, 363-370.
6. Chenoweth, D. M.; Dervan, P. B., Allosteric modulation of DNA by small molecules. *Proc. Natl. Acad. Sci. USA* **2009**, 106, 13175-13179.
7. Chenoweth, D. M.; Dervan, P. B., Structural Basis for Cyclic Py-Im Polyamide Allosteric Inhibition of Nuclear Receptor Binding. *J. Am. Chem. Soc.* **2010**, 132, 14521-14529.
8. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci. USA* **2007**, 104, 10418-10423.
9. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chemical Biology.* **2007**, 2, 561-571.
10. Muzikar, K. A.; Nickols, N. G.; Dervan, P. B., Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, 106, 16598-16603.
11. Olenyuk, B. Z.; Zhang, G.-J.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G., Jr.; Dervan, P. B., Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. USA* **2004**, 101, 16768-16773.
12. Raskatov, J. A.; Meier, J. L.; Puckett, J. W.; Yang, F.; Ramakrishnan, P.; Dervan, P. B., Modulation of NF-kappaB-dependent gene transcription using programmable DNA minor groove binders. *Proc. Natl. Acad. Sci. USA* **2012**, 109, 1023-1028.
13. Takahashi, T.; Asami, Y.; Kitamura, E.; Suzuki, T.; Wang, X.; Igarashi, J.; Morohashi, A.; Shinojima, Y.; Kanou, H.; Saito, K.; Takasu, T.; Nagase, H.; Harada, Y.; Kuroda, K.; Watanabe, T.; Kumamoto, S.; Aoyama, T.; Matsumoto, Y.; Bando, T.; Sugiyama, H.; Yoshida-Noro, C.; Fukuda, N.; Hayashi, N., Development of pyrrole-

imidazole polyamide for specific regulation of human aurora kinase-A and -B gene expression. *Chem. Biol.* **2008**, 15, 829-841.

14. Yao, E.-H.; Fukuda, N.; Ueno, T.; Matsuda, H.; Nagase, H.; Matsumoto, Y.; Sugiyama, H.; Matsumoto, K., A pyrrole-imidazole polyamide targeting transforming growth factor-beta1 inhibits restenosis and preserves endothelialization in the injured artery. *Cardiovasc. Res.* **2009**, 81, 797-804.

15. Matsuda, H.; Fukuda, N.; Ueno, T.; Katakawa, M.; Wang, X.; Watanabe, T.; Matsui, S.; Aoyama, T.; Saito, K.; Bando, T.; Matsumoto, Y.; Nagase, H.; Matsumoto, K.; Sugiyama, H., Transcriptional inhibition of progressive renal disease by gene silencing pyrrole-imidazole polyamide targeting of the transforming growth factor-beta1 promoter. *Kidney Int.* **2011**, 79, 46-56.

16. Chenoweth, D. M.; Harki, D. A.; Phillips, J. W.; Dose, C.; Dervan, P. B., Cyclic pyrrole-imidazole polyamides targeted to the androgen response element. *J. Am. Chem. Soc.* **2009**, 131, 7182-7188.

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

18. Raskatov, J. A.; Hargrove, A. E.; So, A. Y. Dervan, P. B., Pharmacokinetics of Py-Im polyamides depend on architecture: cyclic versus linear. *J. Am. Chem. Soc.* **2012**, 134, 7995-7999.

19. Meier, J. L.; Montgomery, D. C.; Dervan, P. B., Enhancing the cellular uptake of Py-Im polyamides through next-generation aryl turns. *Nucleic Acids Res.* **2011**, 40, 2345-2356.

20. McGovern, S. L.; Helfand, B. T.; Feng, B.; Shoichet, B. K., A specific mechanism of nonspecific inhibition. *J. Med. Chem.* **2003**, 46, 4265-4272.

21. Feng, B. Y.; Shelat, A.; Doman, T. N.; Guy, R. K.; Shoichet, B. K., High-throughput assays for promiscuous inhibitors. *Nat. Chem. Biol.* **2005**, 1, 146-148.

22. Coan, K. E. D.; Shoichet, B. K., Stability and equilibria of promiscuous aggregates in high protein milieus. *Mol. BioSyst.* **2007**, 3, 208-213.

23. Coan, K. E. D.; Shoichet, B. K., Stoichiometry and physical chemistry of promiscuous aggregate-based inhibitors. *J. Am. Chem. Soc.* **2008**, 130, 9606-9612.

24. Feng, B. Y.; Simeonov, A.; Jadhav, A.; Babaoglu, K.; Inglese, J.; Shoichet, B. K.; Austin, C. P., A high-throughput screen for aggregation-based inhibition in a large compound library. *J. Med. Chem.* **2007**, 50, 2385-2390.

25. Shoichet, B. K.; Seidler, J.; McGovern, S. L.; Doman, T. N., Identification and prediction of promiscuous aggregating inhibitors among known drugs. *J. Med. Chem.* **2003**, 46, 4477-4486.

26. Frenkel, Y. V.; Clark, A. D., Jr.; Das, K.; Wang, Y. H.; Lewi, P. J.; Janssen, P. A.; Arnold, E., Concentration and pH dependent aggregation of hydrophobic drug molecules and relevance to oral bioavailability. *J. Med. Chem.* **2005**, *48*, 1974-1983.
27. Doak, A. K.; Wille, H.; Prusiner, S. B.; Shoichet, B. K., Colloid formation by drugs in simulated intestinal fluid. *J. Med. Chem.* **2010**, *53*, 4259-4265.
28. Chenoweth, D. M.; Harki, D. A.; Phillips, J. W.; Dose, C.; Dervan, P. B., Cyclic pyrrole-imidazole polyamides targeted to the androgen response element. *J. Am. Chem. Soc.* **2009**, *131*, 7182-7188.
29. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J. Am. Chem. Soc.* **2008**, *130*, 6859-6866.
30. Jacobs, C. S.; Dervan, P. B., Modifications at the C-terminus to improve pyrrole-imidazole polyamide activity in cell culture. *J. Med. Chem.* **2009**, *52*, 7380-7388.
31. Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Cycle polyamide motif for recognition of the minor groove of DNA. *J. Am. Chem. Soc.* **1999**, *121*, 1121-1129.
32. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.* **2002**, *10*, 2767-2774.
33. Dynamics Software Package (6.11.13), Wyatt Technologies.
34. Loftsson, T.; Brewster, M. E., Pharmaceutical applications of cyclodextrins: basic science and product development. *J. Pharm. Pharmacol.* **2010**, *62*, 1607-1621.
35. Gould, S.; Scott, R. C., 2-hydroxypropyl-beta-cyclodextrin (HP-beta-CD): A toxicology review. *Food Chem. Toxicol.* **2005**, *43*, 1451-1459.
36. Takagaki, T.; Bando, T.; Kitano, M.; Hashiya, K.; Kashiwazaki, G.; Sugiyama, H., Evaluation of PI polyamide conjugates with eight-base pair recognition and improvement of the aqueous solubility by PEGylation. *Bioorg. Med. Chem.* **2011**, *19*, 5896-5902.
37. Kashiwazaki, G.; Bando, T.; Yoshidome, T.; Masui, S.; Takagaki, T.; Hashiya, K.; Pandian, G. N.; Yasuoka, J.; Akiyoshi, K.; Sugiyama, H., Synthesis and biological properties of highly sequence-specific-alkylating N-methylpyrrole, N-methylimidazole polyamide conjugates. *J. Med. Chem.* **2012**, *55*, 2057-2066.
38. Szejtli, J., Introduction and general overview of cyclodextrin chemistry. *Chem. Rev.* **1998**, *98*, 1743-1753.
39. Loftsson, T.; Brewster, M. E., Pharmaceutical applications of cyclodextrins. 1. Drug solubilization and stabilization. *J. Pharm. Sci.* **1996**, *85*, 1017-1025.

Chapter 5

Synthesis and Evaluation of Polyamide-Isoxazolidine Conjugates as Artificial Transcription Factors

Abstract

Pyrrole-Imidazole hairpin polyamides are a cell-permeable class of DNA-binding small molecules that are able to access chromatin and disrupt the transcription factor-DNA binding interface, thereby downregulating endogenous gene expression in human carcinoma cell lines. Conversely, several polyamide-peptide conjugates have been shown to upregulate transcription *in vitro*, with the polyamide mimicking the DNA-binding domain of a transcription factor and the peptide functioning as the transcriptional activation domain (TAD). Recently, isoxazolidine structures have been used as synthetic TADS, and an isoxazolidine-transcription factor conjugate has yielded an 80-fold induction of gene transcription in HeLa cells. Using a small molecule approach, hairpin polyamides may be utilized to replace the transcription factor-DNA binding domain, yielding polyamide-isoxazolidine conjugates for gene activation studies. Solution-phase polyamide synthesis methods have been developed to generate gram-scale quantities of polyamides, enabling the synthesis of several polyamide-isoxazolidine conjugates. These compounds have been studied using various cell-culture assays for their ability to selectively activate transcription. In all of the studies herein, no significant levels of upregulation were observed.

Introduction

Transcriptional regulation is an important strategy for the modulation of cellular signaling pathways, and molecules that can activate or repress transcription are of therapeutic value.¹⁻² Pyrrole-imidazole (Py-Im) hairpin polyamides are a class of sequence-specific DNA-binding small molecules with affinities comparable to transcription factors.³ Previous studies have shown that polyamides readily enter living cells, traffic to the nucleus, and downregulate endogenous gene expression through disruption of the DNA/transcription factor interface.⁴⁻¹¹ Likewise, polyamide-based upregulation of gene expression is also possible, although the process is somewhat more complicated.

Natural transcription factors can bind to either the wider major groove or narrower minor groove of DNA. These binding events can involve hydrogen bonding, electrostatic interactions, or van der Waals interactions with the base pairs or phosphate backbone of DNA. Proteins that bind to DNA, such as leucine zippers and zinc fingers, can bind as monomers, dimers, or in a combinatorial fashion where multiple DNA-binding proteins are required to regulate a certain gene. In addition to the DNA-binding domain (DBD), natural transcription factors generally contain other structural elements for purposes such as ligand-binding and transcriptional activation. In order to create an artificial transcription factor (ATF) to upregulate transcription, it will be essential to include structural motifs that mimic these domains, particularly the DBD and TAD.

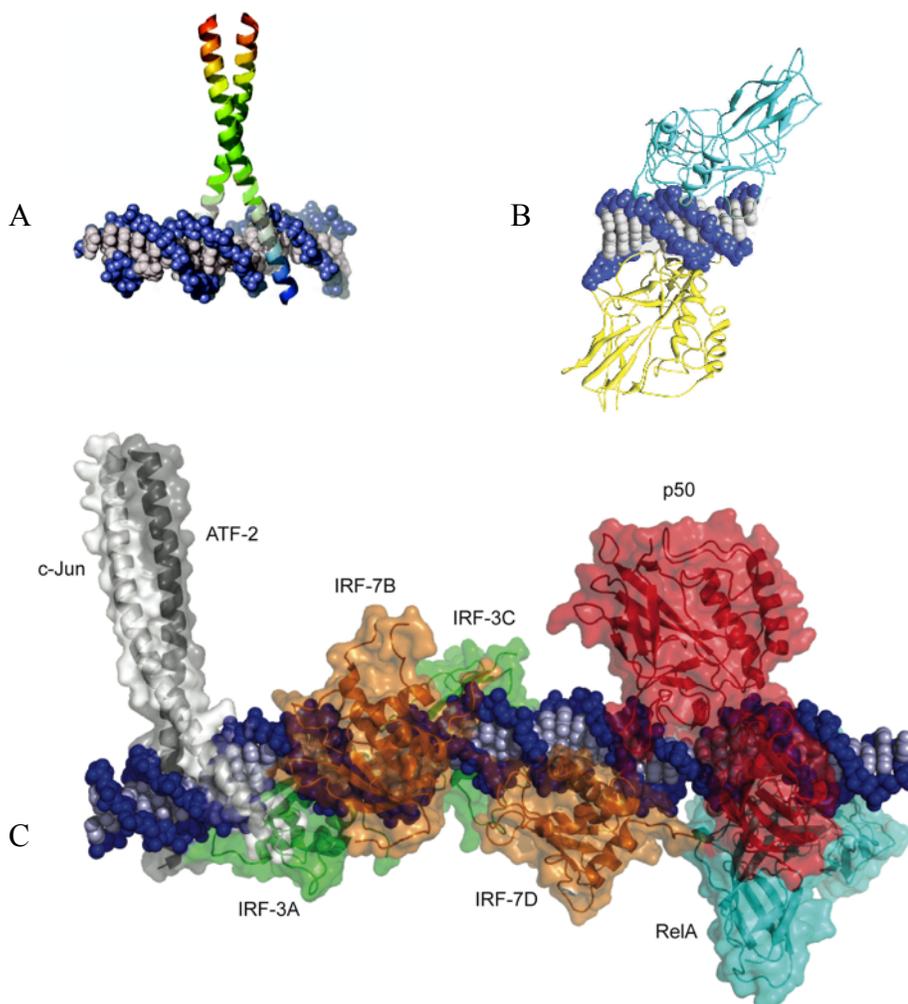


Figure 5.1: Structures of DNA-binding proteins; A) GCN4, a homodimeric leucine zipper; B) NF- κ B, a heterodimer of p50/p65; C) Interferon- β enhanceosome, an example of combinatorial binding.¹²

TADs are structural motifs capable of recruiting and positioning RNA Polymerase II for gene transcription, when localized to specific DNA binding sites. By attaching a TAD to a Py-Im polyamide, it may be possible to selectively activate transcription. In this scenario, the polyamide serves to localize the TAD to a specific sequence in the promoter region of a target gene, thereby directing the assembly of the transcriptional machinery and enabling transcription. In order to allow the activation domain enough conformational flexibility for optimal recruitment of the transcriptional machinery, a stable and flexible linker region is necessary to connect the polyamide to the activation domain.

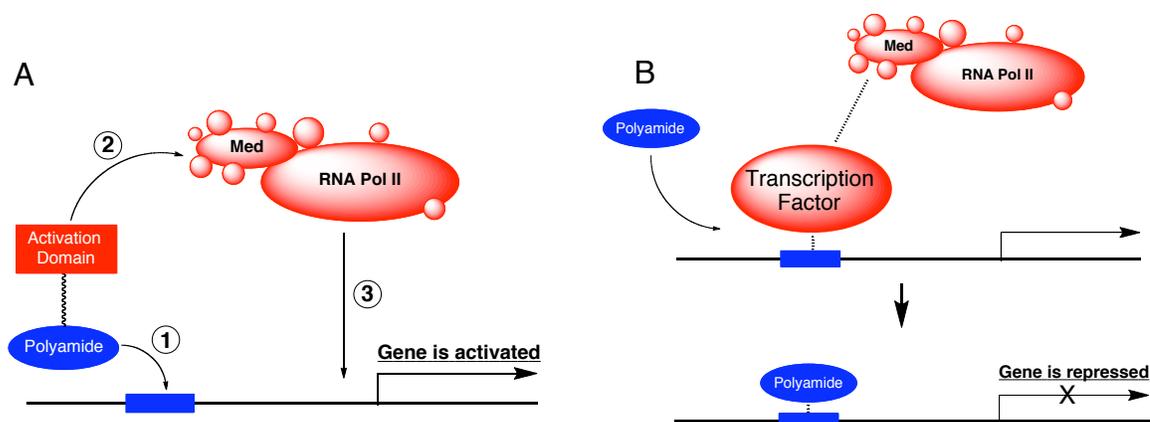


Figure 5.2: Modes of polyamide-based gene regulation. A) Activation: a polyamide attached to a TAD binds to the DNA ①, the TAD recruits the transcriptional machinery ②, which allows assembly of the RNA polymerase II holoenzyme ③, leading to transcription. B) Polyamide can displace natural transcription factors, leading to a downregulation in transcription.

Several polyamide-peptide conjugates have been reported that are capable of upregulating transcription *in vitro*, in which the polyamide mimics the DBD of a transcription factor, and is linked to a peptide known to function as a TAD.¹³⁻¹⁵ A polyamide conjugated to a synthetic activation domain, wrencholol, was also demonstrated to induce transcription *in vitro*.¹⁶ A lack of efficient cellular uptake, likely due to the large molecular weight of these compounds—which often exceeded 4 kDa—prevented *in vivo* applications.^{5,17,18} More recently, a polyamide conjugated to a lower molecular weight synthetic peptoid activation domain showed 5-fold induction of a luciferase reporter in HeLa cells, the first reported ATF that is cell permeable.¹⁹

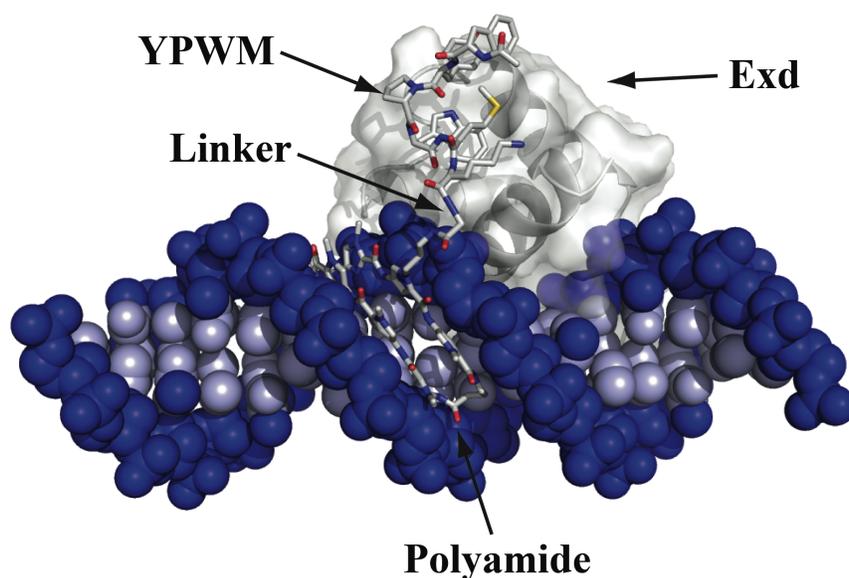


Figure 5.3: An example of a polyamide attached to a peptide activation domain (YPWM) that recruits Exd (a member of the HOX family of transcription factors).

A recent example of a low molecular weight, synthetic TAD is the isoxazolidine structure developed by Mapp and coworkers (Figure 5.4A).²⁰ The transcriptionally active isoxazolidine derivative contains moieties that mimic amino acids (specifically phenylalanine, leucine, and serine) found in natural transcription factors, and is amphipathic, which has been reported as an important feature in TADs.^{1,21}

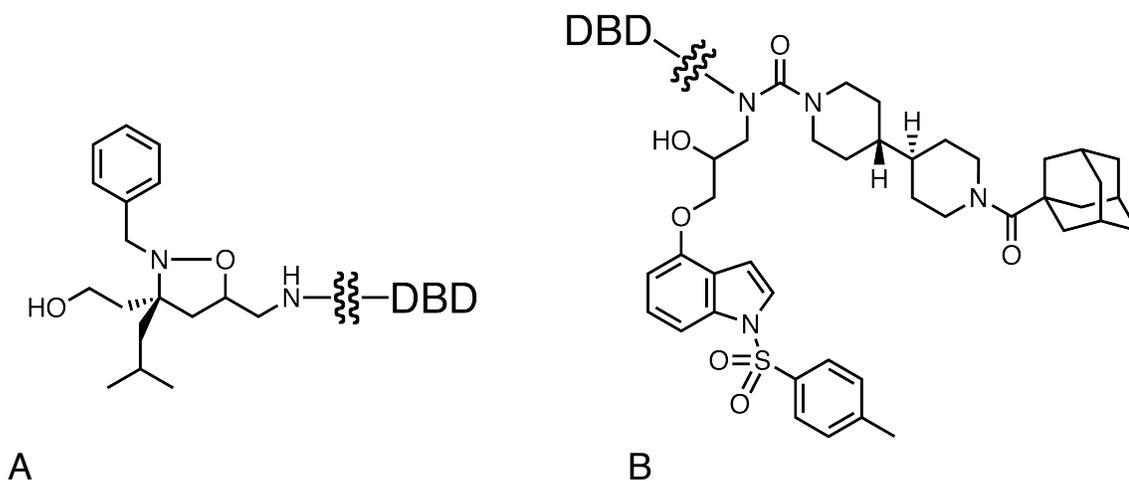


Figure 5.4: Synthetic activation domains. A) Isoxazolidine; B) Wrenchnolol.

Rowe et al. conjugated an isoxazolidine compound to an OxDex steroid—which binds to the glucocorticoid receptor (GR). Gal4, a natural transcription factor, was fused to the functional portion of the GR, which served as the DBD. Transfection of a plasmid containing five Gal4 binding sites upstream of a firefly luciferase reporter gene allowed for the monitoring of transcription *in vivo*.

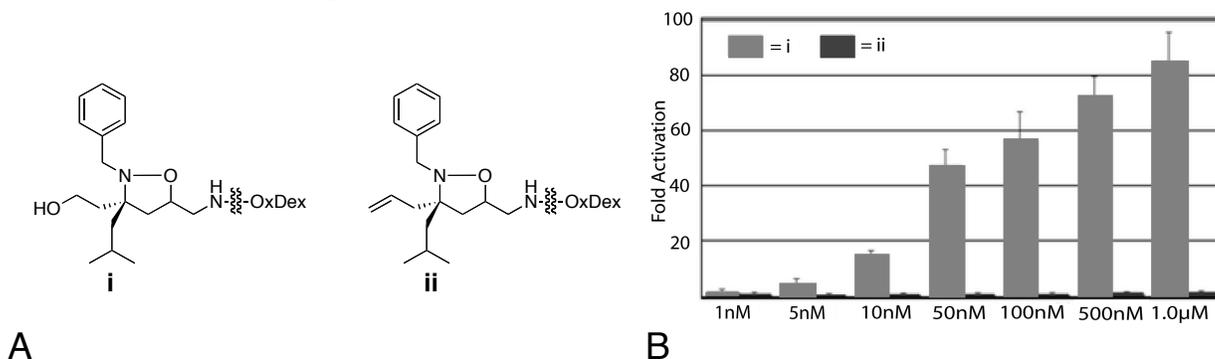


Figure 5.5: a) Structures of active (i) and inactive (ii) isoxazolidine TADs. b) *In vivo* luciferase assay; 80-fold activation at 1.0 μ M. Adapted from Rowe et al.

As seen in Figure 5.5, amphipathic **i** containing an alcohol group was found to function as a TAD, whereas hydrophobic **ii** containing an olefin does not. This system utilizing isoxazolidine construct **i** showed an 80-fold induction in HeLa cells over the control construct **ii**.²² The transcriptional activity of **i** was shown to be comparable to natural peptides used by cells for activation of transcription, making it one of the most active TADs to date. This, combined with its low molecular weight and synthetic accessibility, indicate the great potential for polyamide-isoxazolidine conjugates to function as ATFs.

By directly linking a polyamide to the isoxazolidine TAD, this complicated DBD system can be avoided. The use of a polyamide may be advantageous over DBDs similar to the Gal4/GR/OxDex construct, in that polyamides bind DNA with comparable affinity, and can be programmed to bind many sequences. Therefore, the range of possible targets for

gene activation using polyamides as DBDs is greatly expanded. Py/Im polyamides have also been shown to readily traffic to the nucleus of living cells and modulate endogenous gene expression without the need for lipophilic transfection agents required with other methods.

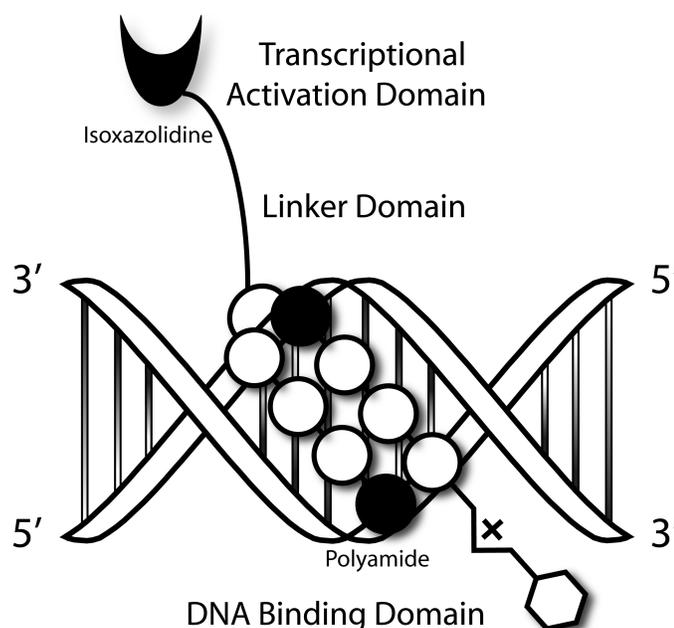


Figure 5.6: General structure of a polyamide-isoxazolidine conjugate, indicating the DBD, linker domain, and TAD.

The polyamide-mediated upregulation of genes has met with limited success in the past, partly due to the small number of molecules that function as general activation domains.¹ In addition, the complication of adding molecular weight to an already large molecule can lead to a decrease in cellular uptake, thereby obstructing *in vivo* applications. Isoxazolidines appear to function as highly active general activation domains, and are relatively low in molecular weight and synthetically accessible. With recent strategies developed in the Dervan lab^{17,23} that have been demonstrated to increase cellular uptake, polyamide-isoxazolidine conjugates would seem to have a great potential for transcriptional activation.

In order to compare results for polyamide-isoxazolidine conjugates to previous isoxazolidine studies, the promoter region containing Gal4 binding sites upstream of the firefly luciferase gene in the plasmid (pG5luc) was targeted. There are five Gal4 binding sites in the promoter, each containing a 5'-WGWWCW-3' sequence, a convenient target site for the polyamides shown in Figure 5.7. The geometry of the designed polyamides, in which the isoxazolidine group is conjugated to an (R)- β -amino- γ -turn, was chosen based on recent work in the Dervan laboratory indicating that substituents at this turn are better tolerated and have a relatively small effect on the binding affinity of the polyamide.²³ The length of the linker will be varied to determine whether this has any significant effect on gene activation.

pG5luc Promoter Region:

```

1  GGTACCGAGT TTCTAGACGG AGTACTGTCC TCCGAGCGGA GTACTTGTCCT
51 CCGACTCGAG CGGGTACTGC TCCTCCGATC GGGTACTGT CCTCCGCGAA
101 TTCCGGAGTA CTGTCCTCCTC AAGACGCTAG CGGGGGGCTA TAAAAGGGGG
151 TGGGGGCGTT CGTCCTCACT CT

```

General Structure:

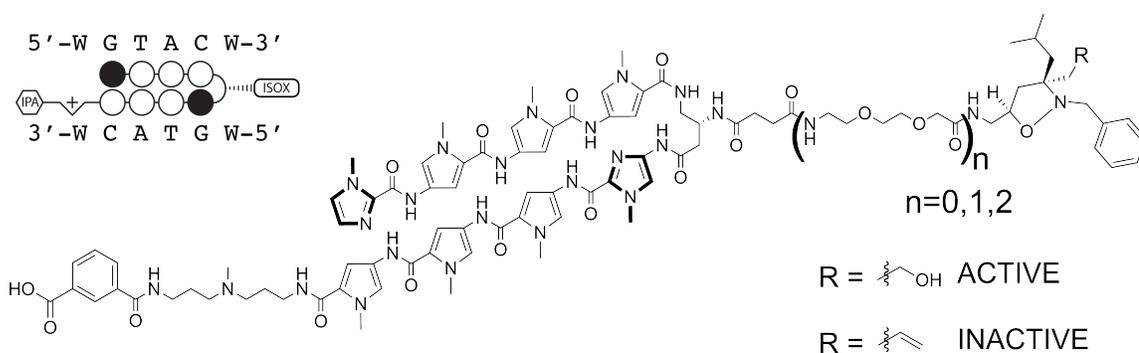


Figure 5.7: Targeting the pG5luc promoter region. There are five binding sites in this promoter, and each contains a 5'-WGWWCW-3' (W=A/T) region (top). The proposed general structure of the polyamide-isoxazolidine conjugates is shown (bottom).

In collaboration with Ryan Casey from Professor Anna Mapp's laboratory (University of Michigan), a library of six polyamide-isoxazolidine derivatives was chosen.

Initially, Mr. Casey began work synthesizing the compounds listed below, which he would later send to Caltech.

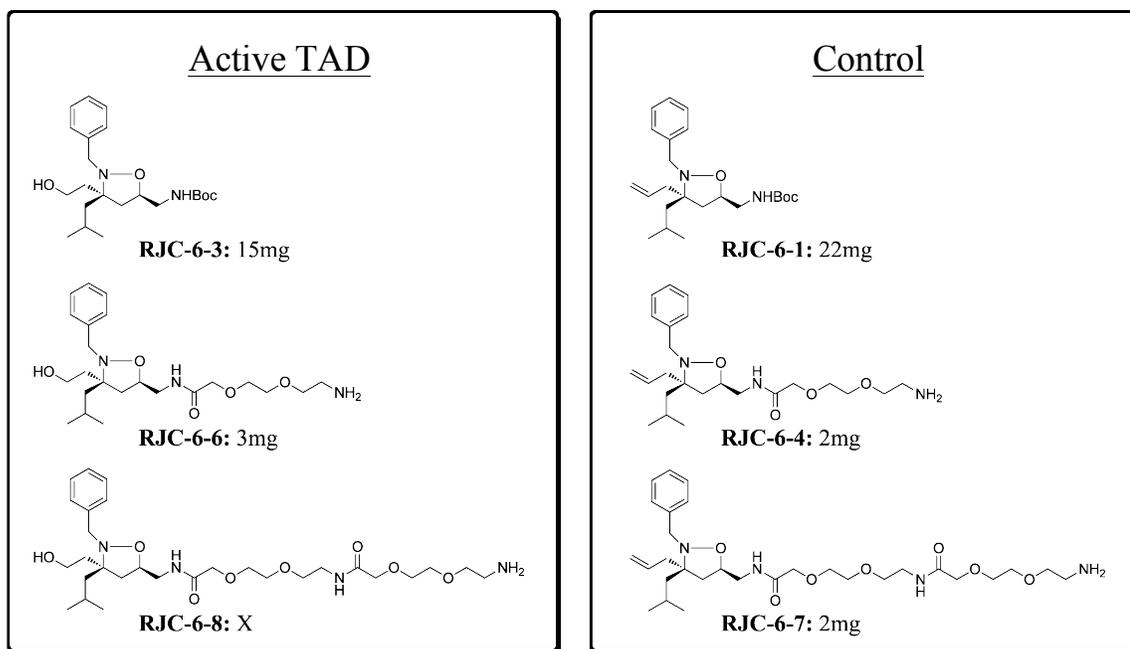


Figure 5.8: Isoxazolidine compounds sent from the Mapp laboratory. Three of these compounds are active (left), and three control compounds are inactive (right).

As seen in Figure 5.8, milligram-scale samples of these compounds were sent to Caltech, including three active TADs, with alcohol moieties, and three inactive control compounds, of varying linker lengths. These compounds can be directly coupled (after *t*-Boc deprotection for RJC-6-1 and RJC-6-3) to a polyamide bearing a carboxylic acid as discussed below. Unfortunately, it was determined that the compound labeled **RJC-6-8** was not the correct compound. The proposed library of polyamide-isoxazolidine conjugates is shown in Figure 5.9.

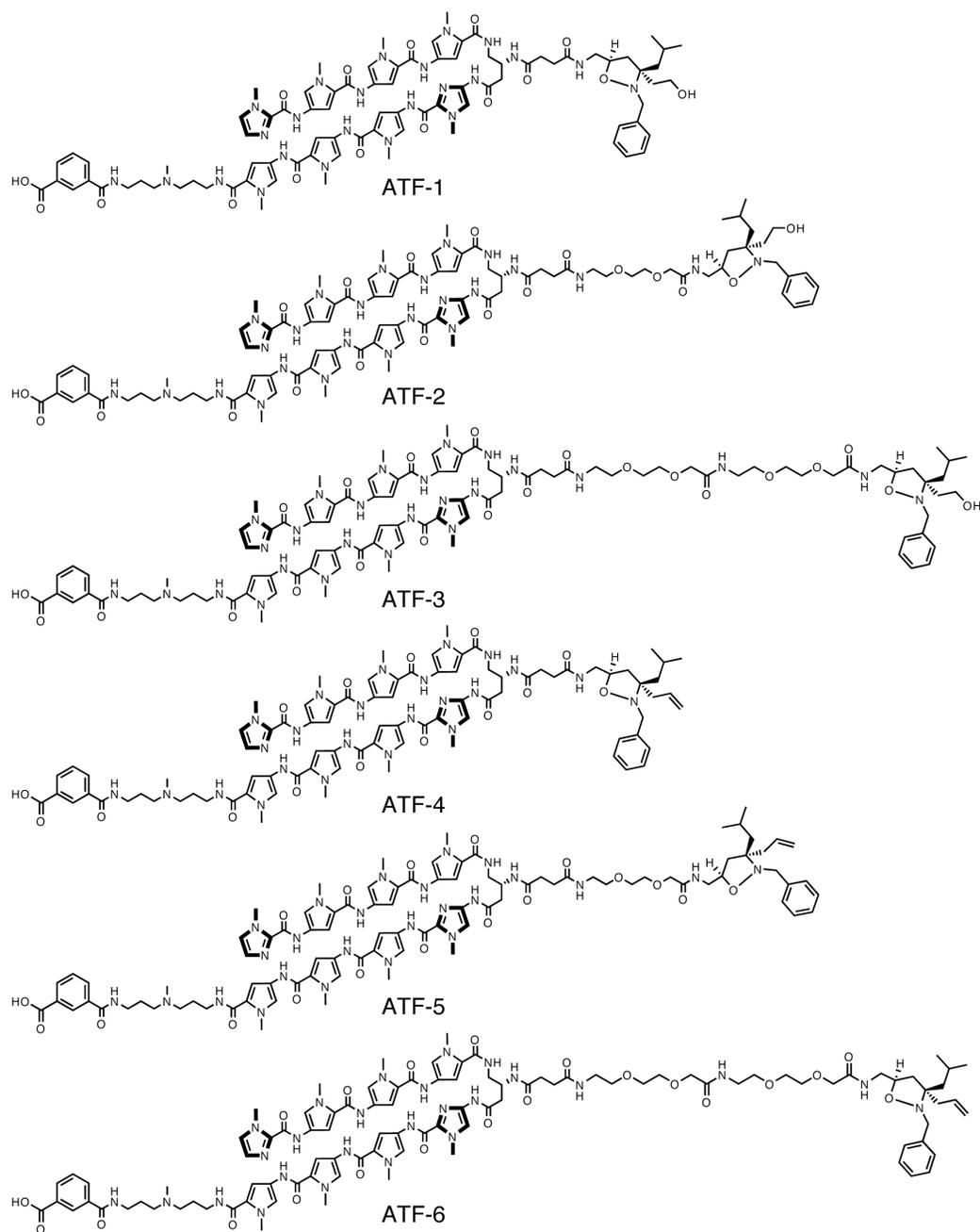


Figure 5.9: Proposed library of polyamide-isoxazolidine conjugates.

Results and Discussion

Solid-phase Polyamide Synthesis

Prior to receiving the isoxazolidine compounds, solid-phase synthesis (SPS)^{24,25} methods were employed to generate sufficient quantities of the unsubstituted polyamide, such that the necessary steps to prepare this polyamide for conjugation with isoxazolidines could be worked out.

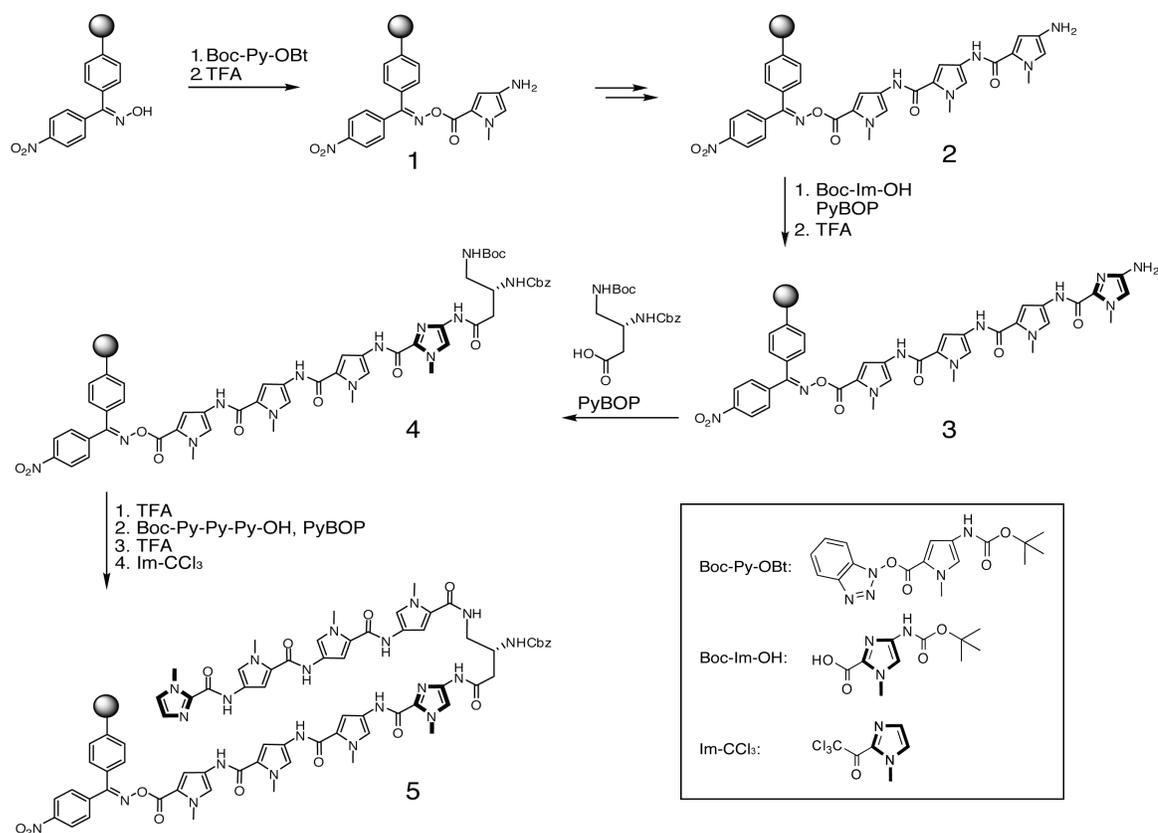


Figure 5.10: *t*-Boc-solid phase synthesis of an eight-ring hairpin polyamide attached to oxime resin, **5**.

This synthesis employed *t*-Boc-SPS with Kaiser Oxime resin, and involved successive PyBOP couplings (or the use of preactivated heterocycles) and TFA deprotections. Each reaction was monitored by the cleavage and analytical HPLC of a

small sample of resin to verify reaction completion. Next, the resin was cleaved with 3,3'-diamino-N-methyl-dipropylamine, as seen in Figure 5.11.

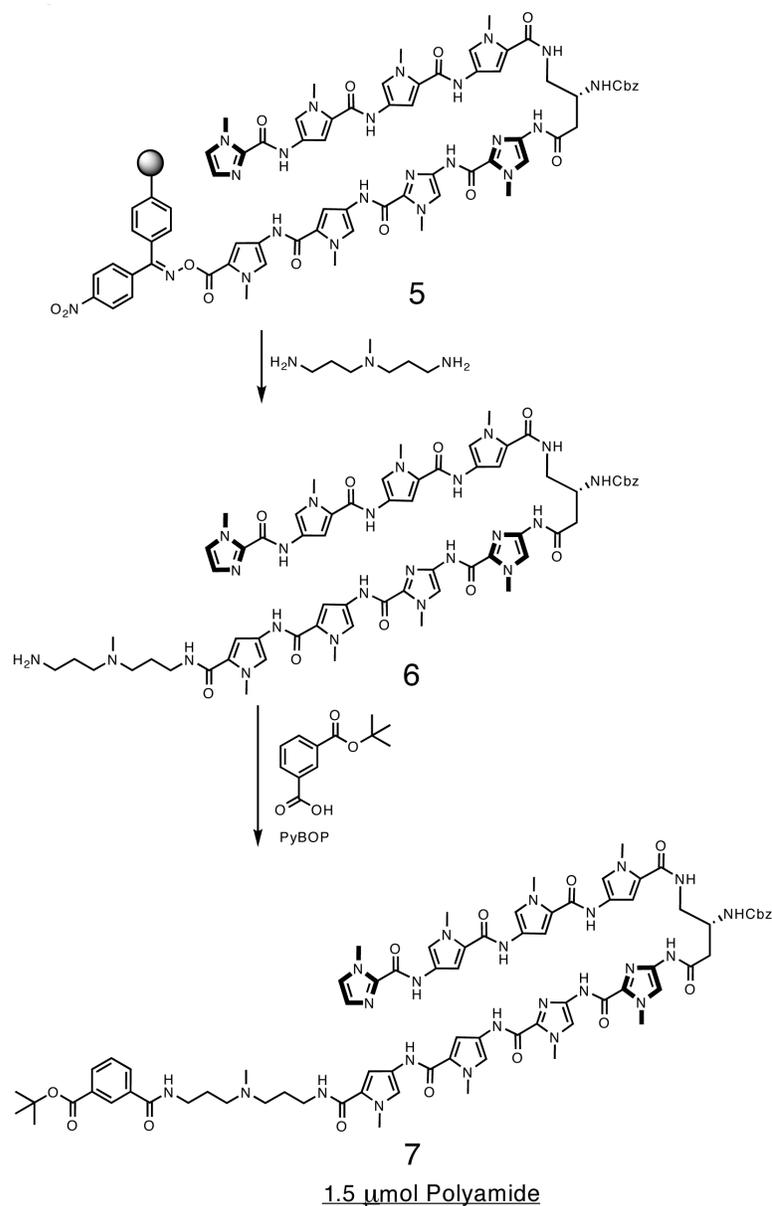


Figure 5.11: Cleavage of polyamide from oxime resin; coupling of mono-*tert*-butyl protected isophthalic acid.

After preparatory HPLC purification to yield polyamide **6**, PyBOP coupling of mono-*tert*-butyl isophthalic acid afforded 1.5 μmol of polyamide **7**. This *tert*-butyl ester protecting group was used to later allow regioselective amide coupling of isoxazolidines to

a carboxylic acid-functionalized amino turn. Next, palladium-catalyzed transfer hydrogenation afforded free amine **8**, in low yield. Succinylation of the amine then afforded the corresponding carboxylic acid **9** in moderate yield, which could be coupled directly to the amino-substituted isoxazolidine compounds.

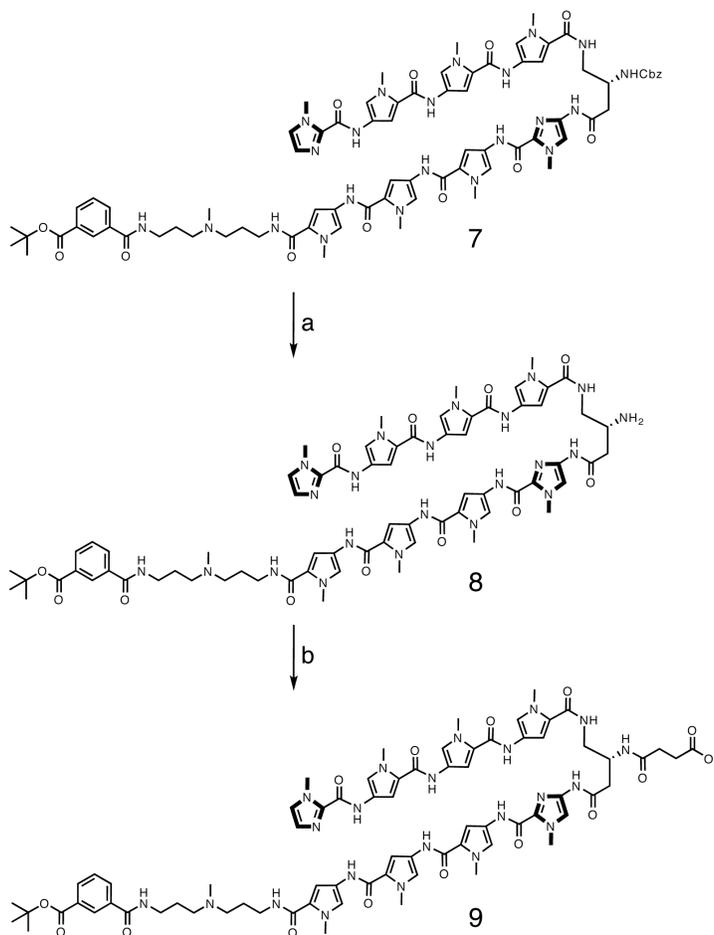


Figure 5.12: Synthesis of carboxylic-acid derivatized polyamide. Conditions: a) Pd(OAc)₂, H₄NHCO₂, DMF/H₂O, r.t., 16h, 27%; b) Succinic Anhydride, DIEA, DMF, r.t., 1 hr, 77%.

However, the yield of these reactions (21% over two steps) was very low, despite attempts at optimization. After the isoxazolidines were received, a nanomole-scale (20nmol) isoxazolidine coupling reaction was attempted, but the miniscule amount of remaining material after work-up and purification made analytical characterization

difficult. These results indicated that further optimization would be necessary in order to preserve valuable material, especially considering two additional steps—an isoxazolidine coupling and subsequent *tert*-butyl ester deprotection—would be required to afford the desired structure. This indicated that a large quantity of polyamide **7** would be necessary to synthesize reasonable quantities (~500 nmol) of the six polyamide-isoxazolidine conjugates.

Solution-phase Polyamide Scale-up

As a way to generate a significant quantity of polyamide for isoxazolidine coupling reactions, a solution-phase route was considered.

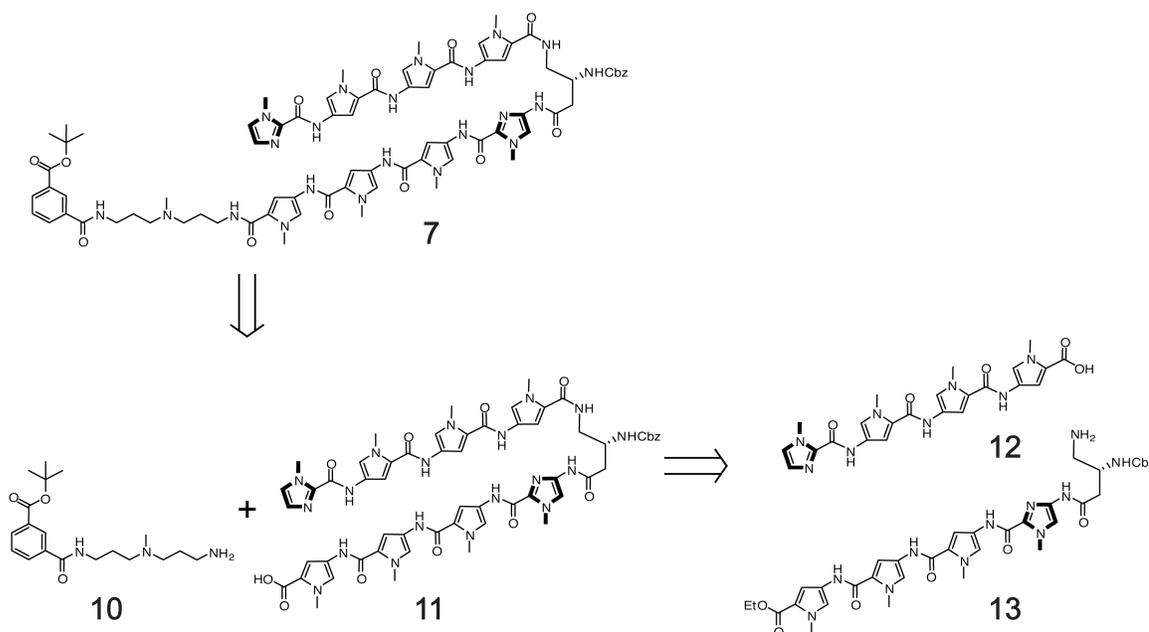


Figure 5.13: Retrosynthesis of polyamide **7**. The tail piece **10** can be coupled to polyamide core **11**, which is generated from tetramer **12**, and pentamer **13** (which contains a hairpin turn unit).

Retrosynthetically, two key amide bond disconnections were envisioned, as seen in Figure 5.13. Polyamide **7** was first disconnected from its tail (**10**) and polyamide core (**11**)

pieces. The polyamide core was further disconnected from its precursors **12** and **13**. Therefore, compounds **10**, **12**, and **13** were the initial synthetic targets.

In order to synthesize the tail region **10**, the amide coupling of commercially available triamine **16** with mono-*tert*-butyl isophthalic acid **15** was proposed. While **15** is commercially available, it is relatively expensive. Other methods to synthesize **15** were attempted, involving the conversion of commercially available isophthalic acid to its mono-*tert*-butyl protected form **15**, without much success. However, a literature search also revealed the synthesis of **15** from commercially available isophthaloyl dichloride **14**.²⁶ This reaction was incorporated to yield **15** in moderate yield, but with extremely inexpensive starting materials. This route allowed the synthesis of large quantities (16 grams) of **15**. Next, DCC-mediated coupling of triamine **16** to carboxylic acid **15** afforded building block **10** in moderate yield.

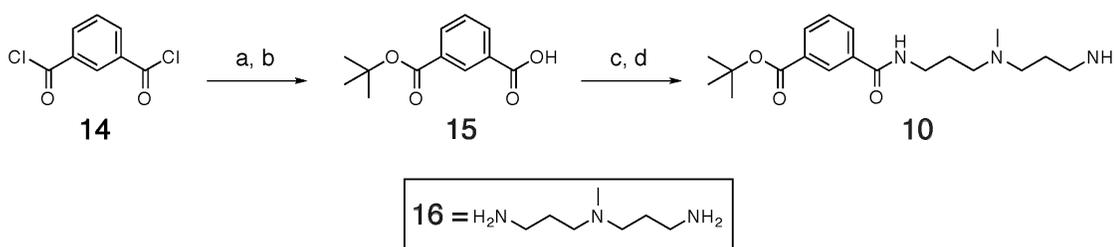


Figure 5.14: Synthesis of **10**. Conditions: a) *tert*-butanol, pyridine, THF, 81°C, 6 hr; b) H₂O, pyridine, THF, 81°C, 6 hr, 52% over two steps; c) DCC, HOBt, DCM, r.t., 1 hr; d) **16**, DCM, r.t., 4 hr, 47% over two steps.

Beginning with trimer **20**, which can be readily synthesized in two steps from Dervan Laboratory stock compounds **17** and **18**, the synthesis of tetramer **12** was completed in two steps by coupling of an activated imidazole unit to **20**, followed by saponification. Tetramer **23** was also synthesized from trimer **20** in two steps through an activated ester coupling of a bifunctional imidazole derivative, followed by acidic

deprotection of the *N*-*tert*-butoxycarbonyl (Boc) protecting group. PyBOP coupling of an (*R*)- β -amino- γ -turn to **23** followed by acidic Boc-deprotection yielded pentamer **13**.

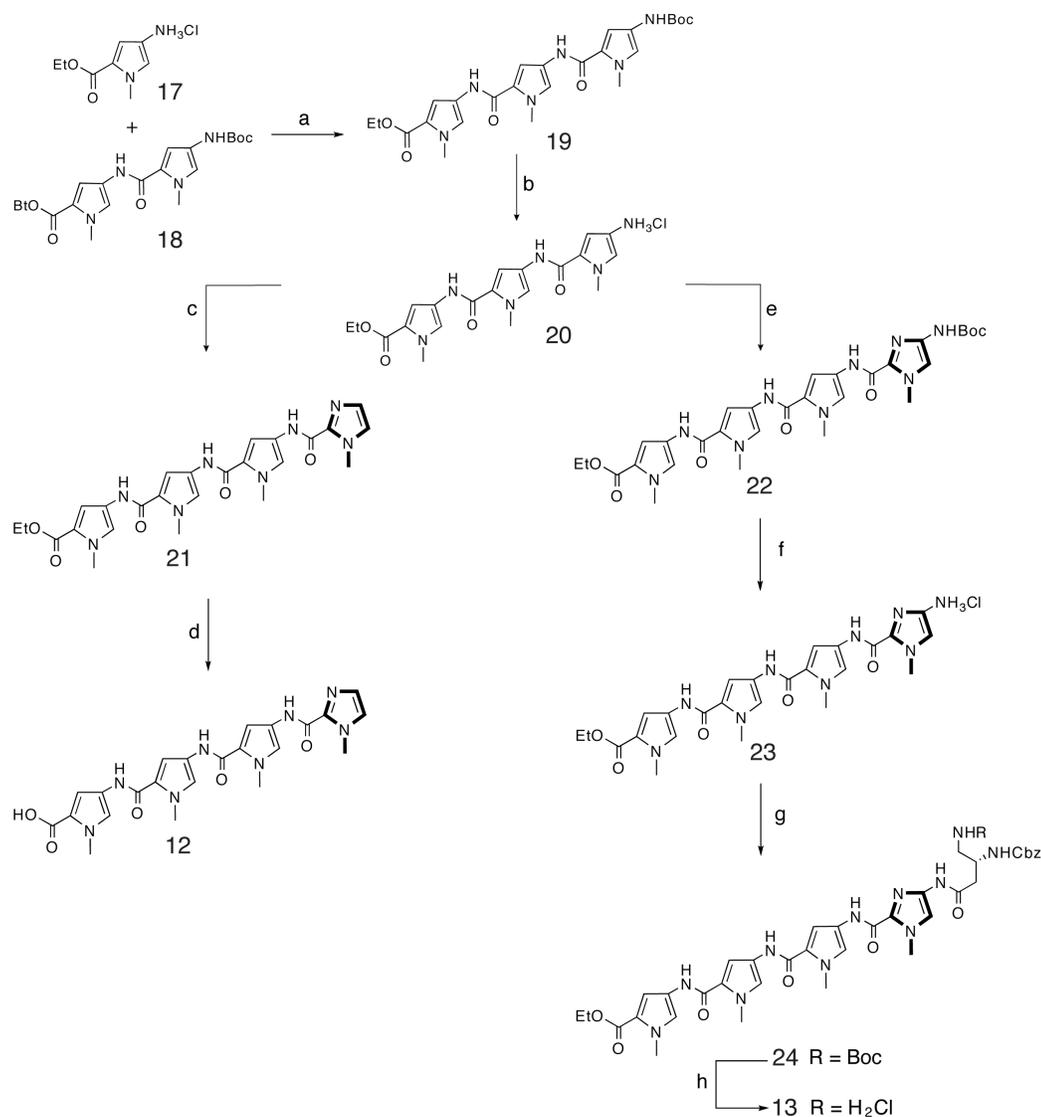


Figure 5.15: Synthesis of **12** and **13** from trimer **20**. Conditions: a) DIEA, DMF, r.t., 16h, 91%; b) 4.0M HCl in 1,4-Dioxane, r.t., 18h, 97%; c) Im-CCl₃, DIEA, DMF, r.t., 16h, 87%; d) 1N NaOH, H₂O, 1,4-Dioxane, 50°, 6h, 98%; e) Boc-Im-COOH, HBTU, DIEA, DMF, r.t., 16h, 81%; f) 2.0M HCl in Et₂O, 1,4-Dioxane, r.t., 3h, 97%; g) Z- β -Dab(Boc)-OH, PyBOP, DIEA, DMF, r.t., 16h, 78%; h) 2.0M HCl in Et₂O, 1,4-Dioxane, r.t., 3h, 99%.

With polyamide precursors **10**, **12**, and **13** synthesized in gram-scale quantities (0.5-1.0 grams of each compound), polyamide **7** was synthesized in 3 steps: PyBOP coupling of **12** and **13** to yield the polyamide core ethyl ester **25**, saponification to afford

carboxylic acid **11**, and PyBOP coupling of **11** with tail unit **10**, as illustrated in Figure 5.16.

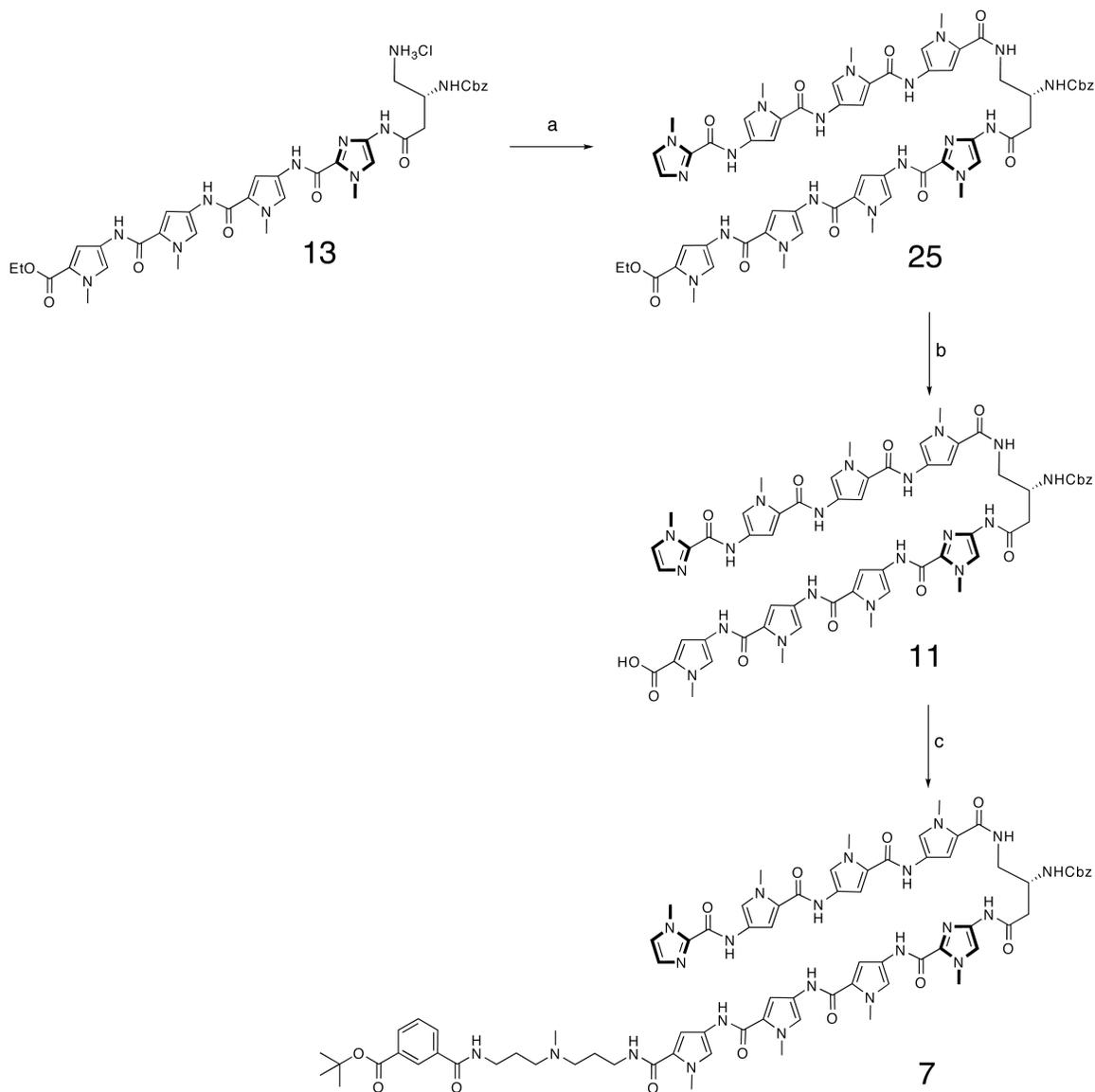


Figure 5.16: Synthesis of Polyamide **7**. Conditions: a) **12**, PyBOP, DIEA, DMF, r.t., 16h, 76%; b) 1N NaOH, H₂O, 1,4-Dioxane, r.t., 15h, 86%; c) **10**, PyBOP, DIEA, DMF, r.t., 2h, 94%.

Using this solution-phase synthesis route, it was possible to synthesize over 350mg (over 200 μmol) of **7**, with more synthetic precursors available if necessary. While this route necessitated chromatography steps and required some optimization, it is anticipated that future syntheses of this nature would be more straightforward and less laborious,

allowing synthesis of large quantities of polyamide where previous solid-phase methods would be inadequate. Indeed, Chenoweth et al. from the Dervan laboratory have since developed a high-yielding chromatography-free route for gram-scale solution phase synthesis of a polyamide.²⁷

Synthesis of Polyamide-Isoxazolidine Conjugates

With a large quantity of polyamide **7** available, the next step was derivatization of the polyamide for coupling with the isoxazolidine compounds. Initially, this required deprotection of the Cbz-protected amino turn. Due to the *tert*-butyl protecting group on the polyamide tail, catalytic transfer hydrogenation was chosen as the optimal chemoselective transformation.²⁸⁻³⁰ At first, the use of palladium(II) acetate as a catalyst along with ammonium formate (as a source of hydrogen) was attempted, as mentioned above. However, among the various conditions tested, heating was necessary for full conversion and decomposition was always a major drawback, leading to poor yields. Palladium on carbon (Pd/C) was also attempted as a hydrogenation catalyst, and this reaction afforded amino compound **8** with less decomposition. Despite indications that this was a clean reaction, yields were consistently low after filtration and purification. It was found that with too much Pd/C catalyst, yields were lower. It is speculated that some polyamide becomes bound—perhaps irreversibly—to the catalyst, leading to a loss of product. Additionally, it was found that the standard method of filtering Pd/C, using celite, was also a major reason for lost material. Optimization of reaction conditions and purification methods proved crucial, and eventually a high yielding procedure was developed.

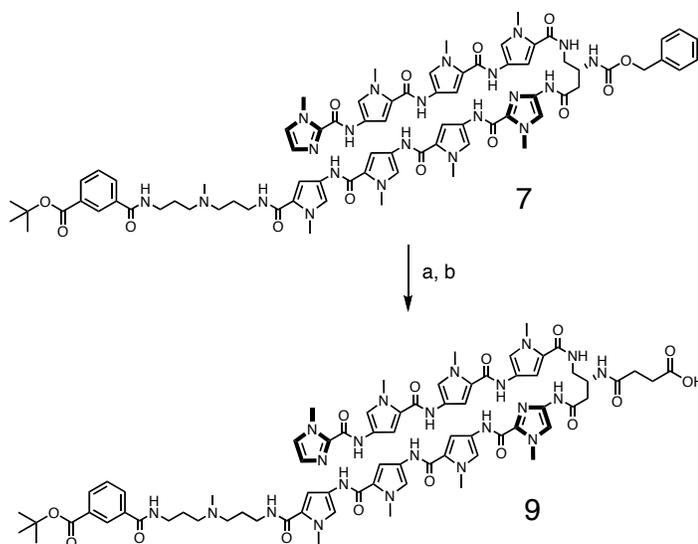


Figure 5.17: Synthesis of acid-functionalized polyamide **9**. Conditions: a) Pd/C, $\text{H}_4\text{NHC}_2\text{O}_4$, H_2O , 1,4-Dioxane, 60° , 12 hours; b) Succinic Anhydride, DMF, DIEA, r.t., 30min, 87% over two steps.

As seen in Figure 5.17, the subsequent addition of succinic anhydride gave polyamide **9** in 87% yield over two steps. The final two transformations were the isoxazolidine-polyamide coupling and the subsequent acid deprotection of the *tert*-butyl group. The coupling step was somewhat problematic, as it is difficult to precisely measure stoichiometric quantities of each compound on such a small scale, and this reaction generally involved the formation of minor side products. A general procedure was developed to maximize product formation, but side-products have not been completely eliminated in this procedure. Finally, the *tert*-butyl group on the tail must be deprotected, and this step was also initially challenging. While different acid-catalyzed conditions (TFA/DCM, HCl/Ether, etc.) were attempted, decomposition of these polyamide-isoxazolidine compounds was pervasive among all methods, and yields were extremely low. In order to limit the loss of invaluable isoxazolidine compounds, this reaction was performed on a very small scale, making it difficult to monitor. Compounding this issue was the need to purify the final conjugates via preparatory HPLC, often leading to an

additional significant reduction in yield. However, a milder deprotection procedure proved to be more successful, employing CeCl_3 and NaI in refluxing acetonitrile. In general, this reaction minimized decomposition and significantly increased the overall yield.

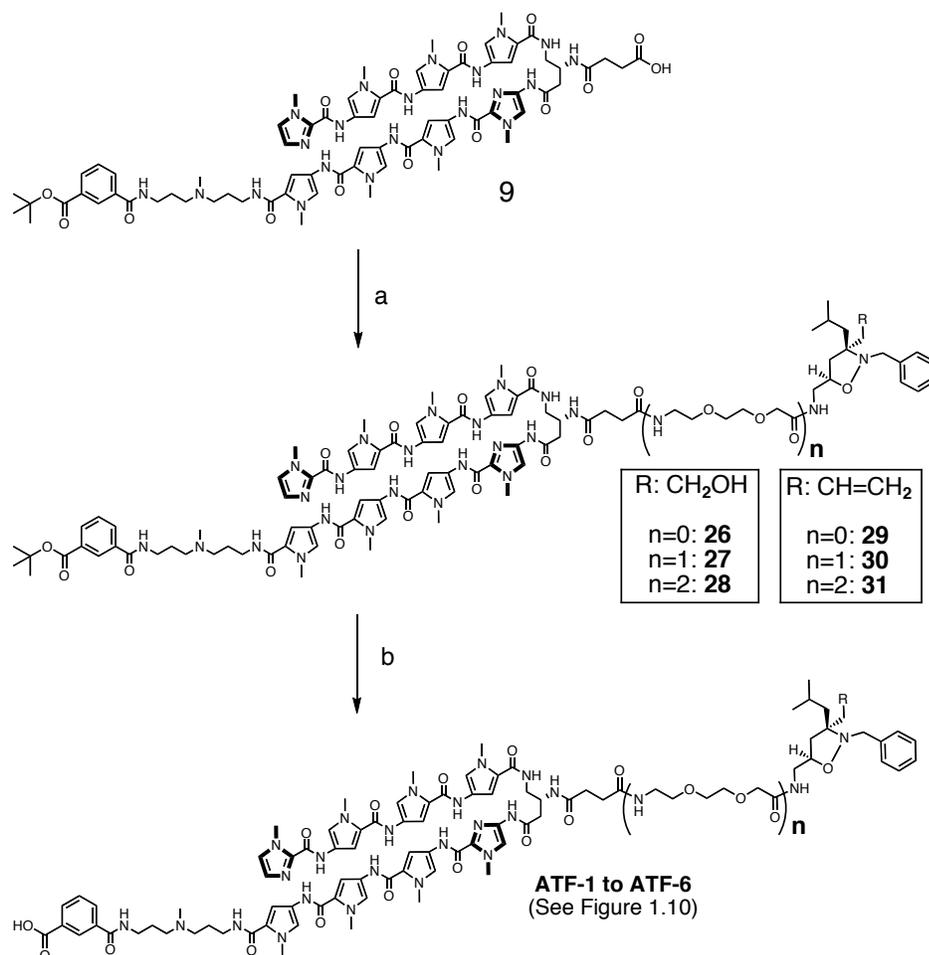


Figure 5.18: Polyamide-isoxazolidine coupling and final *tert*-butyl deprotection. Conditions: a) PyBOP, DIEA, DMF, RJC-6-X (see Figure 5.8), 60° , 3 hours; b) $\text{CeCl}_3 \cdot \text{H}_2\text{O}$, NaI, MeCN, 82°C , 1 hour, 21% yield over two steps.

Using this route, **ATF-1**, **ATF-2**, and **ATF-4** were synthesized. Moderate quantities of **ATF-1** and **ATF-4** were synthesized, while a small amount of **ATF-2** was also synthesized (Figure 5.19). With the excess of isoxazolidine compounds necessary for these coupling reactions, combined with the small amount of available isoxazolidine material, it was possible only to make usable quantities of **ATF-1** and **ATF-4**. However,

these initial stocks allowed preliminary biological investigations to explore the gene activation potential of these conjugates (discussed below).

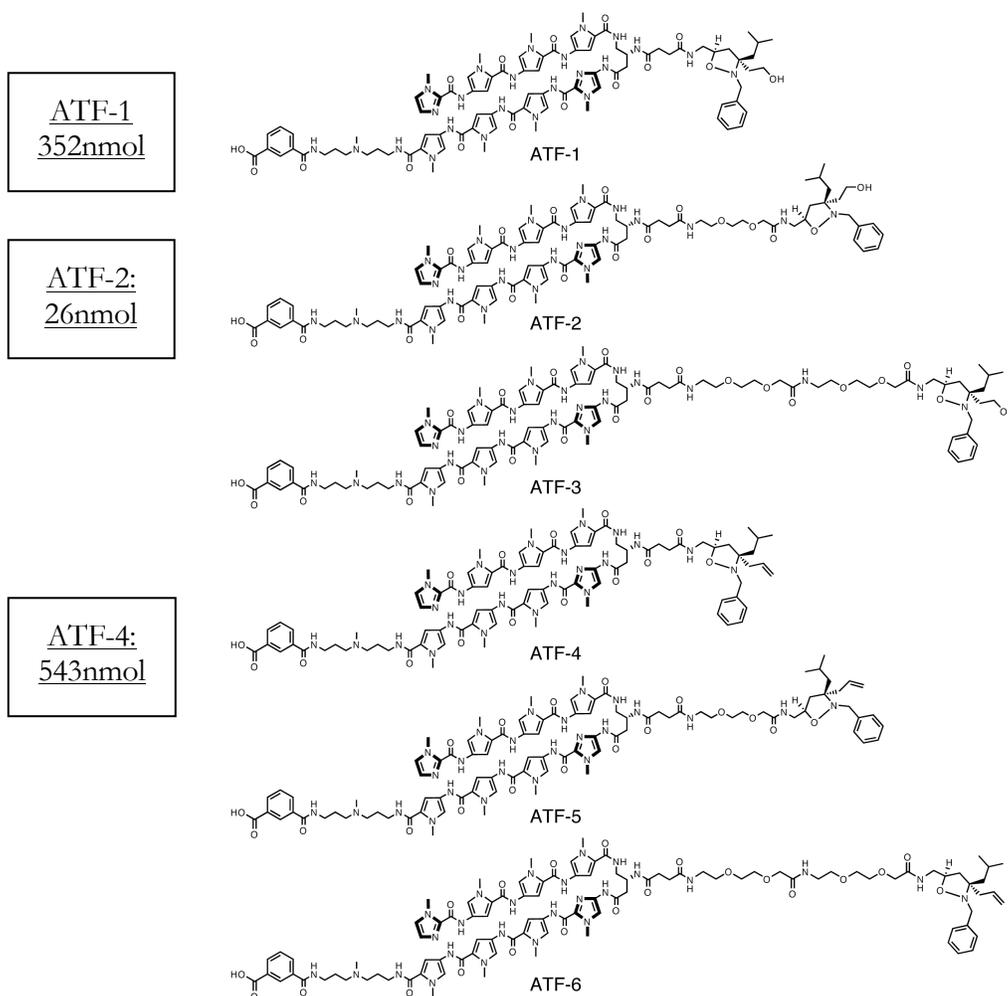


Figure 5.19: Initial synthesized stock of proposed polyamide-isoxazolidine conjugate library.

Linker Synthesis

After a second shipment of isoxazolidine compounds arrived from Mr. Casey in the Mapp laboratory, in somewhat larger quantities than before (Figure 5.20), optimization of the linker couplings was attempted at Caltech, due to low reported yields in previous procedures, which limited the amount of isoxazolidine that was sent.

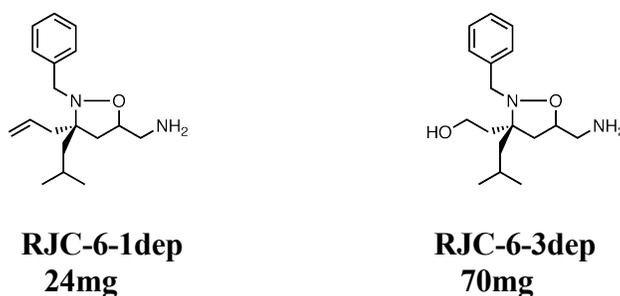


Figure 5.20: Second batch of isoxazolidines received from Mapp laboratory.

In considering the optimal route for attachment of a longer length linker, it was deemed that previous methods involving sequential attachment of two short PEG linkers were less than desirable, due to compounding low yield steps. In addition, the previous Fmoc-protected linkers yielded deprotection byproducts, which necessitated preparatory HPLC purification, further decreasing the yield.

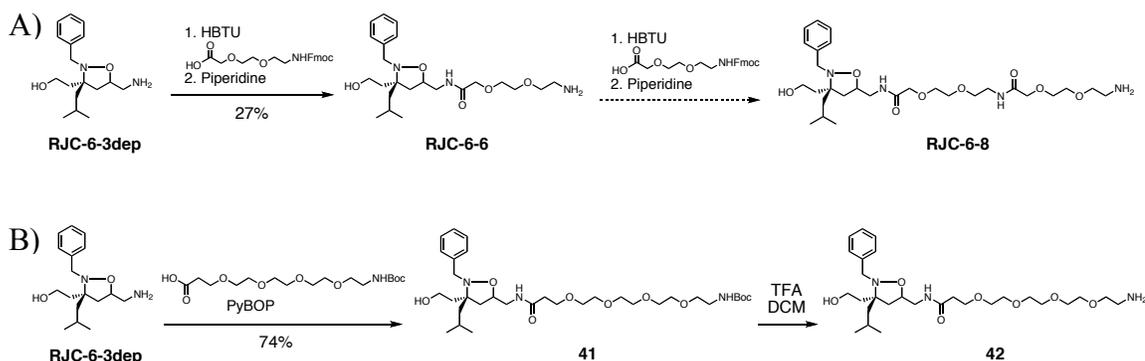


Figure 5.21: Routes for linker attachment (long linker variant). A) Previous strategy, involving sequential attachment of two short PEG linkers. B) Strategy utilizing a single longer linker, with a Boc protecting group.

As seen in Figure 5.21, another envisioned route was to use a single longer linker with a Boc protecting group. This procedure involved a single coupling, providing a relatively non-polar product, which can be purified by preparatory TLC. The linker-isoxazolidine coupling reaction proceeded with relatively high yield. After Boc deprotection, this compound can be directly coupled to a polyamide (Figure 5.22). It was found that the previous CeCl_3/NaI t-butyl deprotection procedure caused significant

decomposition in this route, so an improved 95% TFA/5% H₂O protocol was developed allowing significantly improved yields.

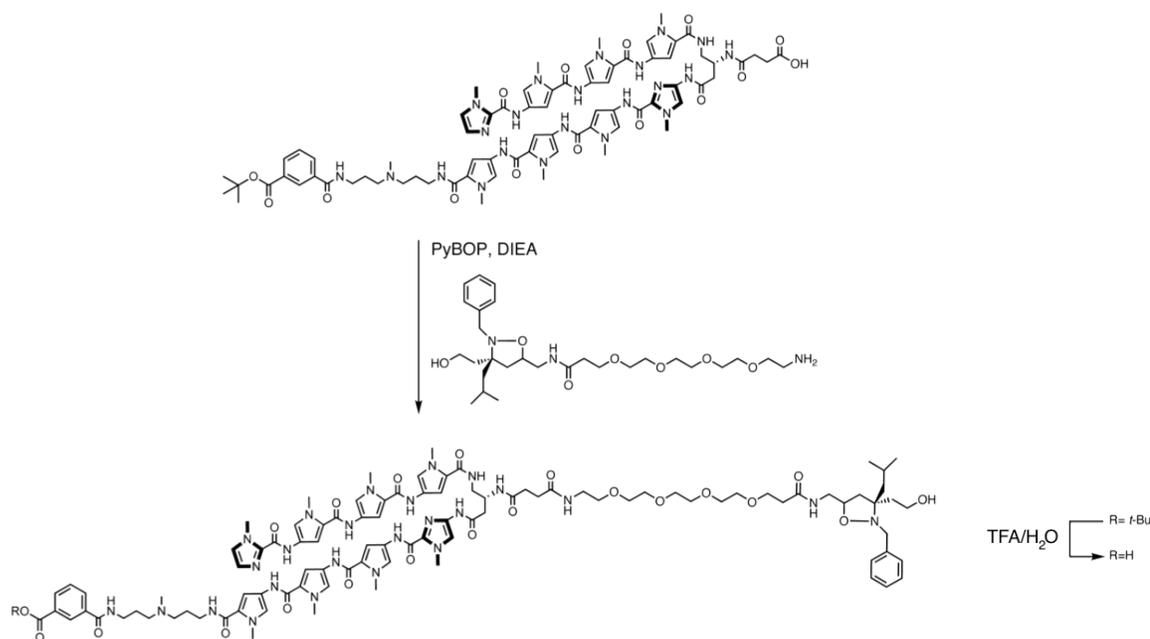


Figure 5.22: Synthesis of polyamide-isoxazolidine with longer linker.

In Vivo Activation Studies

Concurrent with the synthesis of a polyamide-isoxazolidine with a longer linker, the initial completed compounds underwent pilot studies to assay for upregulation of transcription. Samples of **ATF-1**, **ATF-2**, and **ATF-4** were sent to Michigan, along with parent polyamide **43** (an additional control) to be added to HeLa cells transfected with the pG5luc plasmid, the same system previously used in the Mapp laboratory.

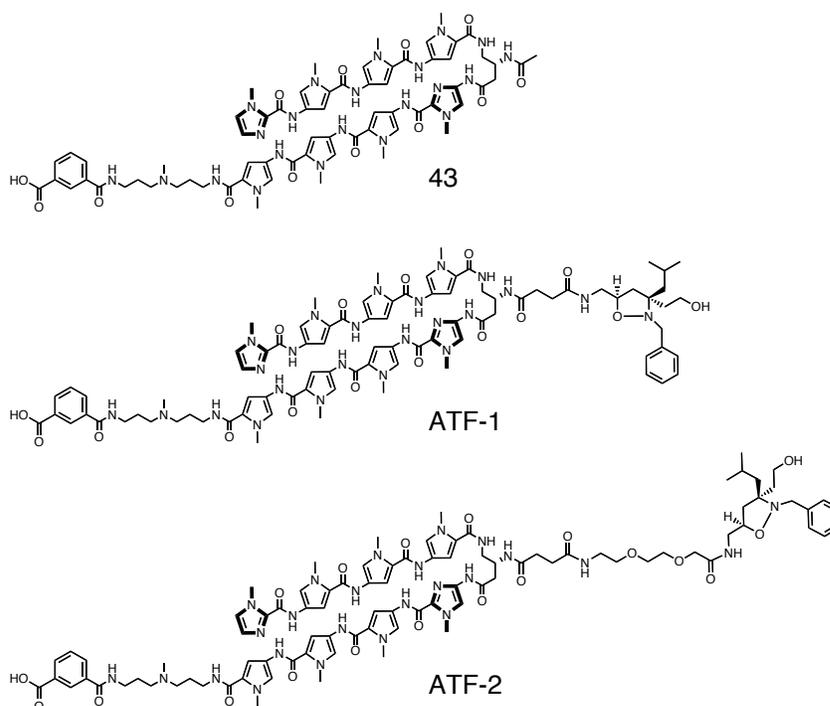


Figure 5.23: Compounds used in initial luciferase assay by Mapp laboratory.

Additionally, a second plasmid containing a Renilla luciferase construct was used, thereby allowing a dual luciferase assay to be performed. Cells were plated, incubated for 24 hours, and dosed with polyamide-isoxazolidine conjugates for a 40 hour period.

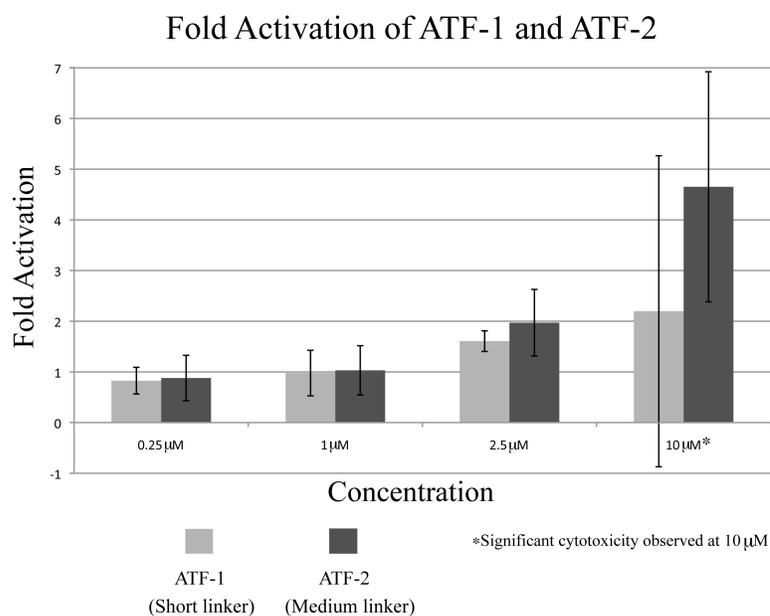


Figure 5.24: Initial results from Mapp laboratory: dual luciferase reporter assay in HeLa cells.

In addition to the Renilla luciferase normalization, the fold activation is also normalized for wells containing no synthetic conjugates, as well as for luciferase signal elicited by the parent polyamide, **43**. At 10 μ M, significant cytotoxicity was observed, contributing to the large error. These initial results showed some promise, as there was a modest (~2-fold) increase in luciferase signal for cells treated with ATF-1 and ATF-2 at 2.5 μ M. However, previous studies in the Mapp laboratory exhibited an 80-fold activation with a Gal4-based DBD.

In a later experiment, polyamide **44**, containing a longer linker, was assayed by the Mapp laboratory in a similar dual luciferase reporter assay.

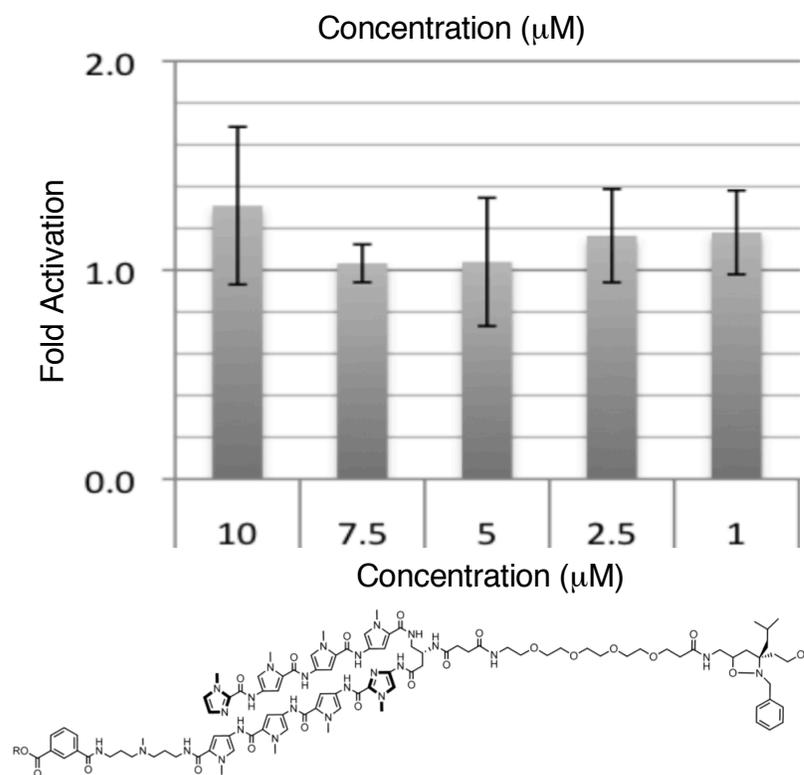


Figure 5.25: Dual luciferase reporter assay for compound **44** in HeLa cells.

As for the results of ATF-1 and ATF-2, no significant activation of luciferase activity was observed in this assay.

In addition to the experiments in HeLa cells run by the Mapp laboratory, experiments were performed at Caltech using the LNAR+ cell line. This cell line was developed by the Charles Sawyer laboratory at Sloan-Kettering. It consists of transformed LNCaP cells that have been stably transfected with a firefly luciferase reporter gene controlled by an Androgen Response Element (ARE) in the promoter region (which likely contains a GWWC sequence). Unfortunately, the sequence of this construct isn't available. However, it was envisioned that this experiment would still be informative as to whether or not these polyamide-isoxazolidine conjugates are activating transcription.

In this experiment, polyamides at varying concentrations were added to cells in quadruplet, on a 96-well plate. Cells were plated at 50,000 cells/mL, incubated for 24 hours, and treated with both active and inactive polyamide-isoxazolidine conjugates, as well as parent polyamide compounds as further controls. Positive and negative controls were set up by the presence and absence of induction by dihydrotestosterone (DHT), respectively. DHT is a potent agonist for androgen receptor activation, so cells induced by DHT should exhibit an increased luciferase signal. In addition to the luminescence assay that indicates firefly luciferase activity, a WST-1 cell proliferation assay was also performed to determine any cytotoxicity caused by the polyamides along the range of concentrations studied. This data can also be used to normalize the luciferase data, in that WST-1 assays provide an estimate of cell viability by measuring mitochondrial activity.

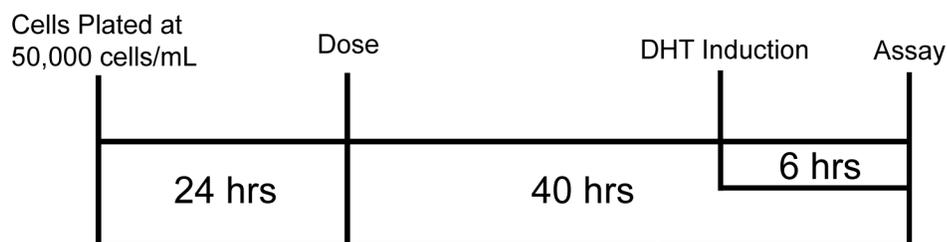


Figure 5.26: Experimental timeline. Cells are plated, incubated 24 hours, dosed with compounds, and incubated for 40 hours. A six-hour DHT induction was also performed in wells with no other added compounds.

After initial pilot studies were performed to optimize the assay (e.g. optimization of incubation time for the WST-1 assay, examination of a wide range of polyamide concentrations, etc.), an experiment consisting of four compounds, and ranging from 3nM to 10 μ M in concentration was conducted.

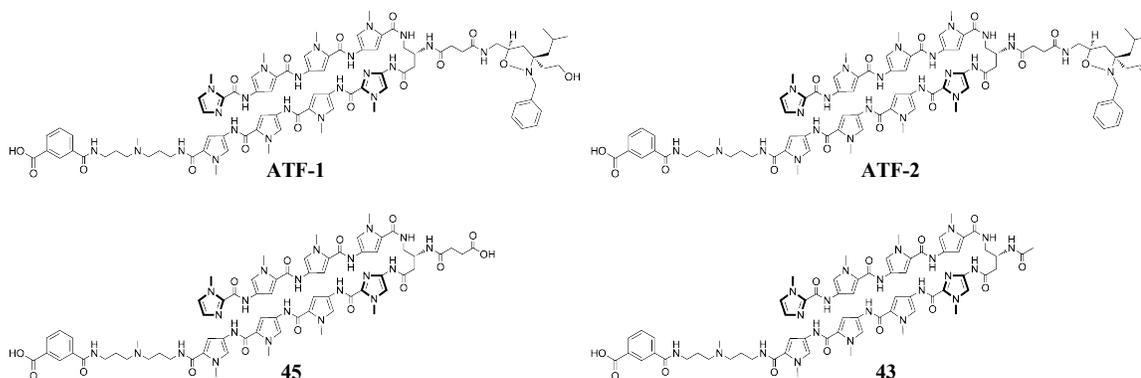
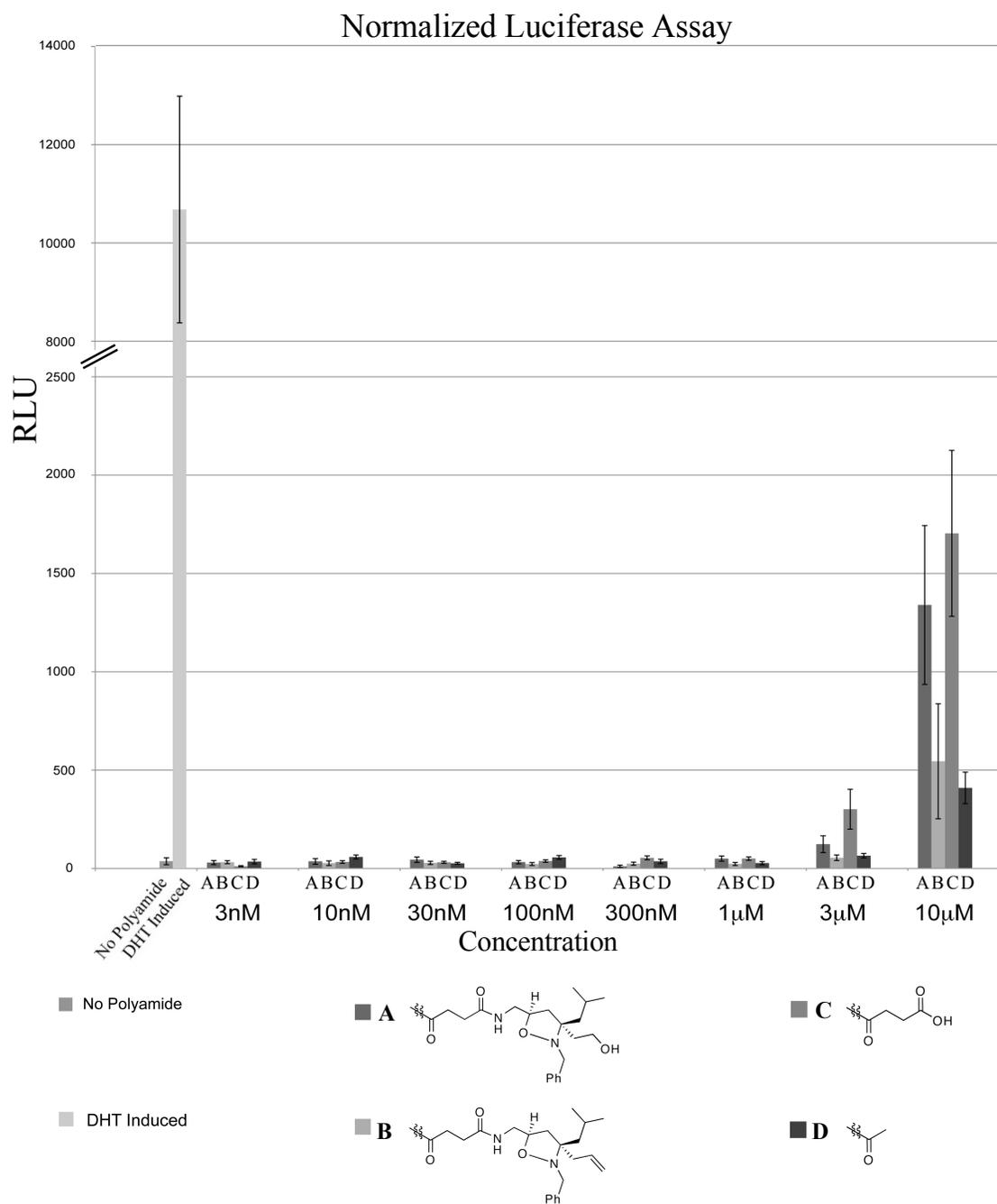


Figure 5.27: Compounds used in LNAR+ assay.

Compounds **43** and **45** can be easily synthesized from intermediates in the solution-phase route detailed above. After the luciferase assay was performed, the data was collected and normalized by subtracting out the background values of wells without any cells. As mentioned above, it was also normalized for WST-1 values, and the final values reported in Relative Luminescence Units (RLU).



General Structure:

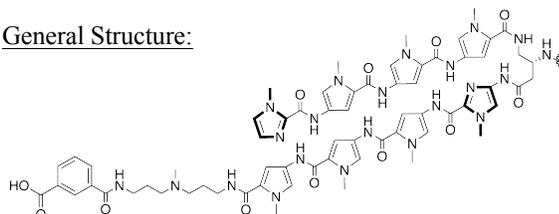


Figure 5.28: LNAR+ assay. The structures of each of the polyamides is indicated below. Concentration is varied from 3nM to 10µM. Note: the y-axis is split.

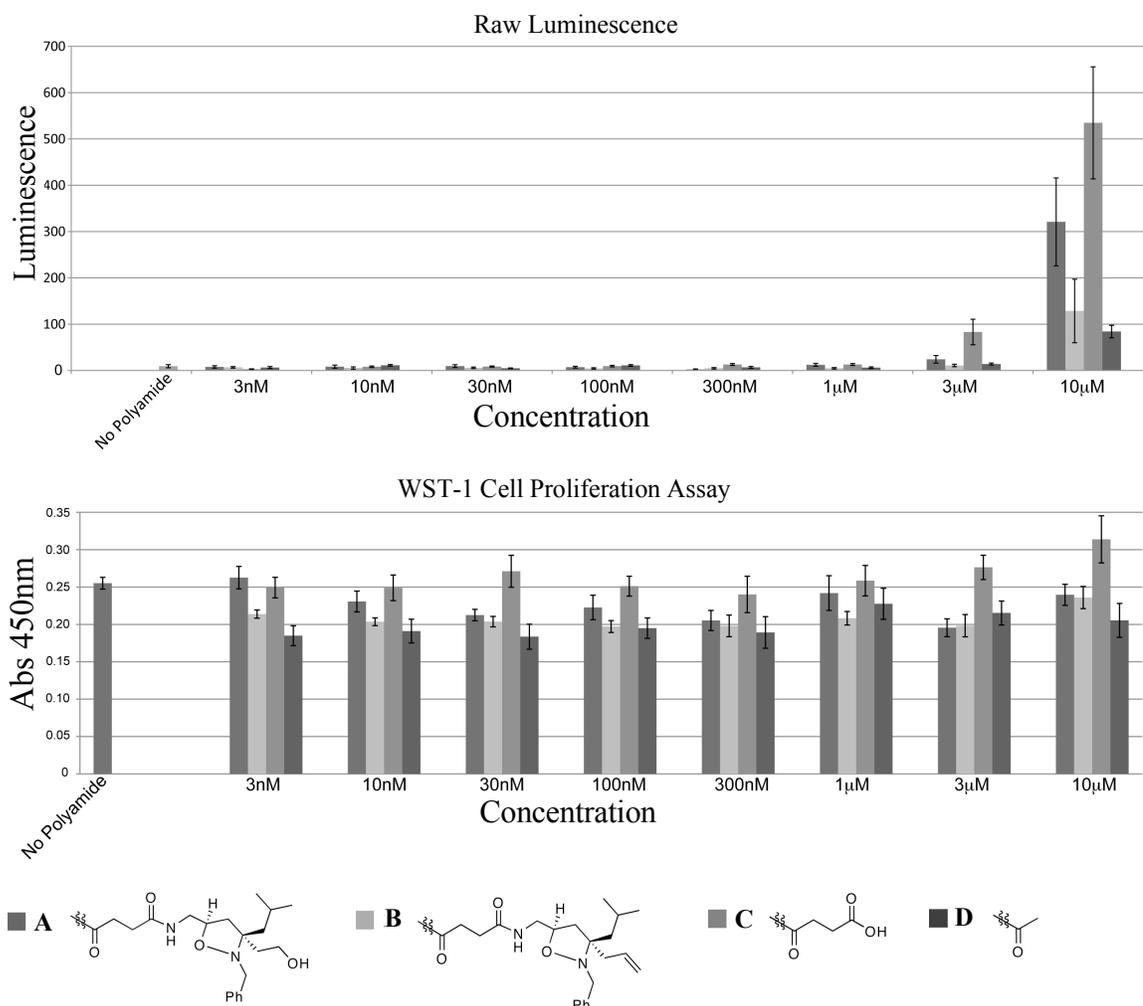


Figure 5.29: LNAR+ assay raw data. It is notable that there is no significant decrease in WST-1 signal for any of the compounds.

The data seem to indicate that all four compounds cause an increase in luciferase signal to various degrees, which certainly was not anticipated. From this data, the succinylated polyamide **44** (listed as compound C in this study) actually seems to have the most drastic effect, followed by the active polyamide isoxazolidine, **ATF-1** (compound A). At 10 μ M, however, all four compounds appear to elicit a significant increase in luciferase signal.

To investigate this system further, the concentration could be increased, to determine if the luciferase signal continues to increase, especially with no observable

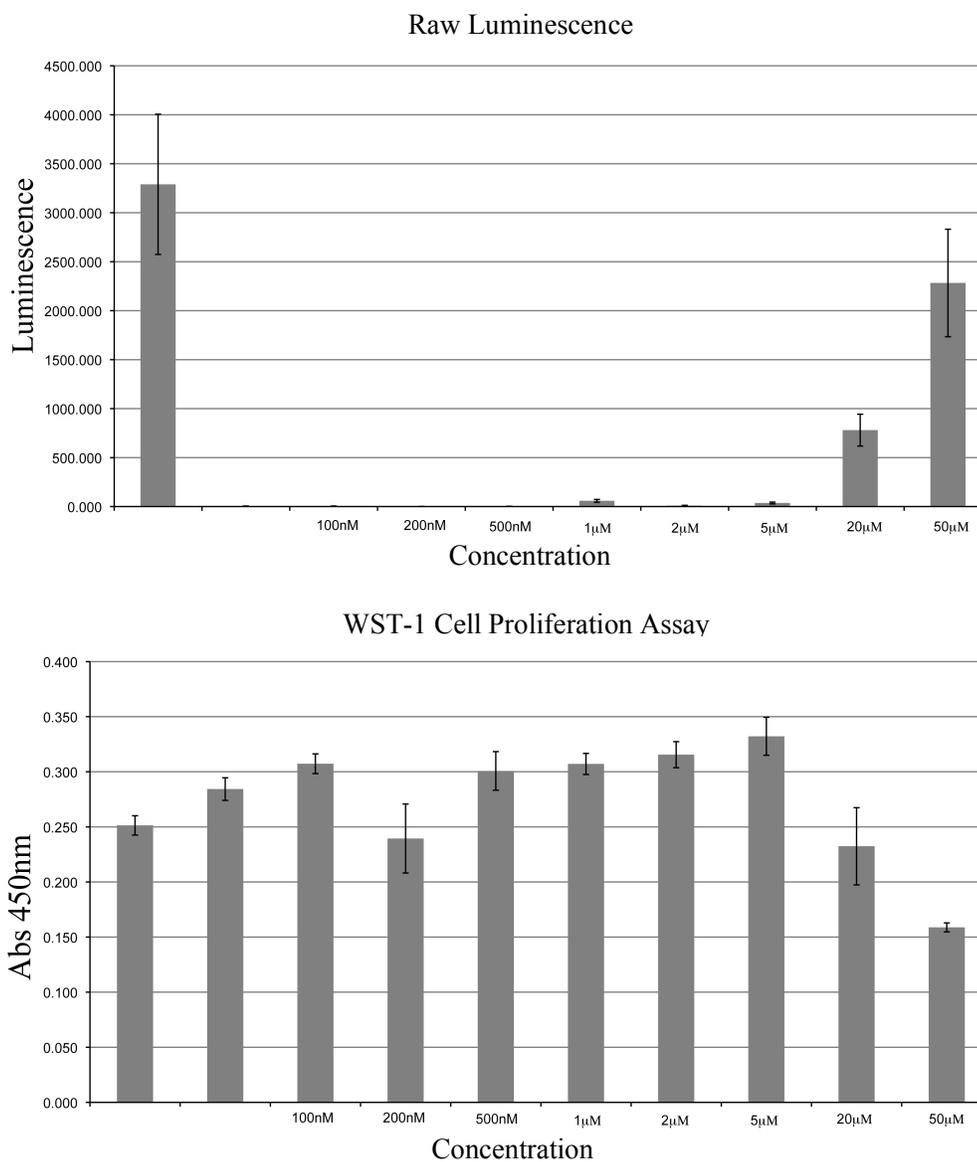


Figure 5.30: LNAR+ assay results for **43**.

In this experiment, **43** elicited an extremely high luciferase signal, comparable to that of DHT at 50µM. In addition, there was observable cytotoxicity at this concentration. However, this polyamide has been previously shown to downregulate AR transcription,¹¹ so it seems unlikely that it would be directly activating transcription of the luciferase gene at the ARE. It is speculated that perhaps at these high concentrations there is some other global effect, causing downstream events that lead to activation of transcription at the ARE. This is speculation however, as there is no known process by which this occurs. Regardless

of the reason for this effect, it was determined that it was likely not possible to use this system to determine whether or not a polyamide-isoxazolidine conjugate is directly upregulating transcription, if a luciferase signal is elicited for all polyamide compounds tested.

In addition to the luciferase experiments outlined above, an alternative method to assay for transcriptional activation in an endogenous system was envisioned. In previous work by Muzikar et al.,³¹ Py-Im polyamides targeted to the glucocorticoid response element (GRE) have shown the ability to disrupt the binding of glucocorticoid receptor (GR) and inhibit expression of GR target genes such as glucocorticoid-induced leucine zipper (GILZ).

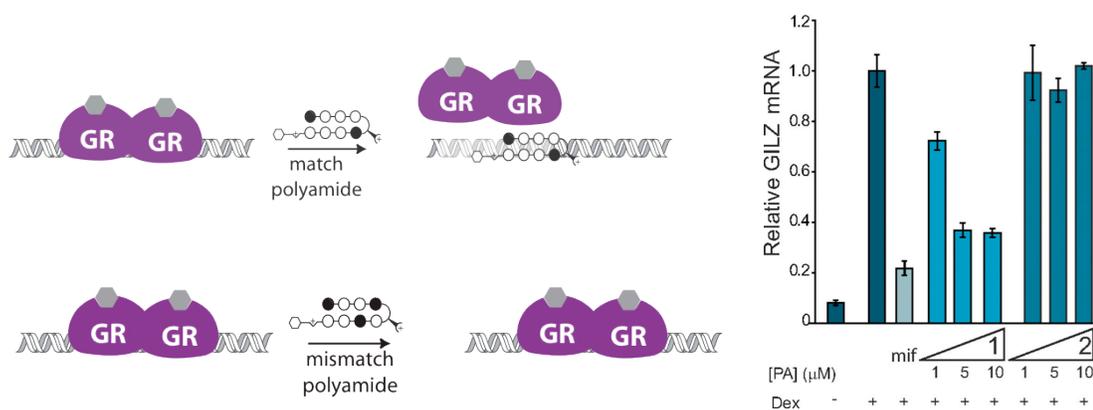


Figure 5.31: Polyamide disruption of glucocorticoid receptor binding, and downregulation of dexamethasone-induced GILZ transcription in A549 lung carcinoma cells.

Conversely, by using a polyamide-isoxazolidine conjugate targeting the same sequence of the GRE, it was envisioned that it might be possible to upregulate the transcription of GILZ. This experiment was performed under similar conditions to those used previously to assess inhibition of GR-mediated expression of GILZ, with 48 hour incubation of A549 cells with polyamide-isoxazolidine conjugates. The levels of GILZ mRNA for both dexamethasone induced and non-induced cells were determined using

quantitative real-time reverse transcriptase PCR relative to β -glucuronidase, an endogenous control, performed in collaboration with Katy Muzikar.

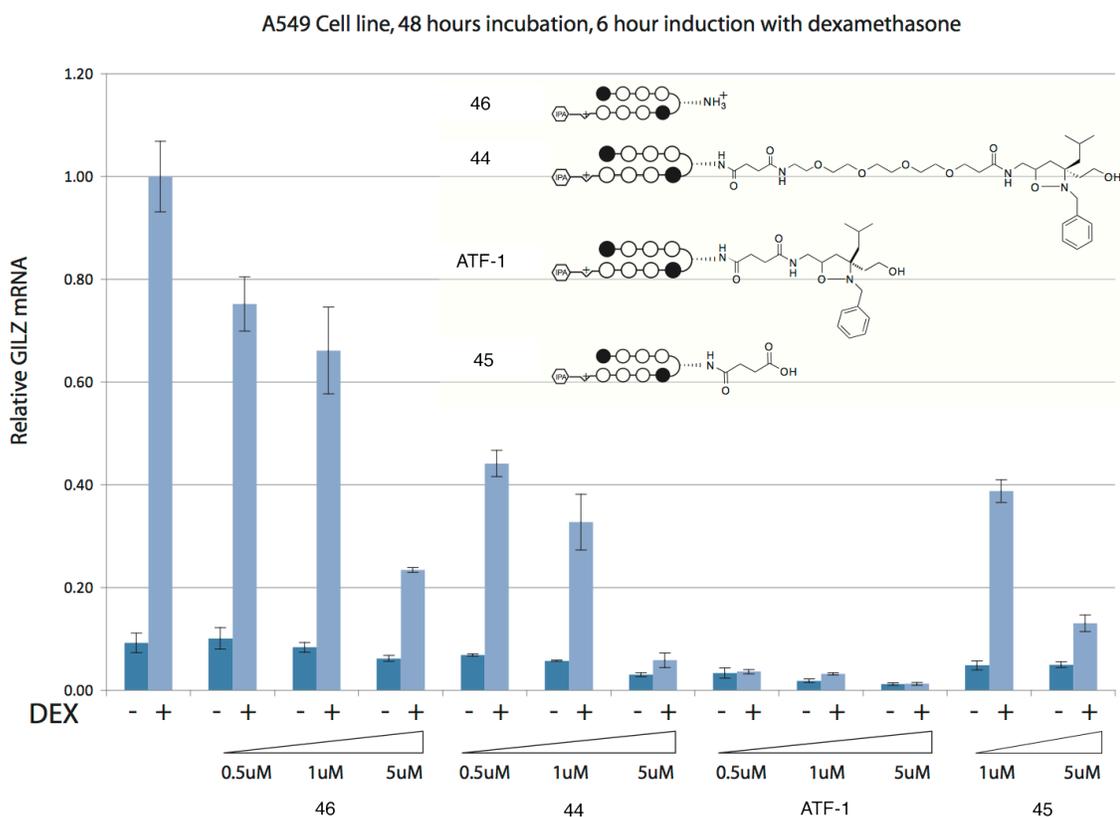


Figure 5.32: Quantitative real-time reverse transcriptase PCR of A549 cells dosed with polyamide-isoxazolidine compounds.

Based on these data, it appears that there is no significant increase in levels of GILZ mRNA in any of the compounds studied, including the polyamide-isoxazolidine conjugates. For the non-induced systems, low levels of GILZ mRNA were seen across the panel of compounds, and for the dexamethasone-induced cells, GILZ mRNA levels were either low or decreased with increasing concentration of polyamide, consistent with polyamide-based disruption of the GR-DNA interface and concentration dependent repression of GILZ expression.

Finally, a cell-free assay was envisioned as another route to determine if polyamide-isoxazolidine conjugates were able to activate transcription *in vitro*. Based on previous studies,¹³⁻¹⁵ the plasmid pMLΔ53, containing a g-less cassette reporter gene upstream of a TATA box and known ligation sites, was chosen and obtained via a generous gift of the laboratory of Cheng-Ming Chiang at the University of Texas Southwestern Medical Center. Following previous protocols, construction of a plasmid for gene activation experiments was attempted via insertion of a synthetic oligonucleotide containing the desired WGWWCW binding sites into a Bgl2 restriction site. After initial cloning attempts failed, subsequent sequencing results indicated there was no Bgl2 restriction site in this plasmid, despite it bearing the same name as a previous plasmid with a known Bgl2 site. Another strategy was attempted via insertion of a synthetic oligonucleotide between EcoRI and XmaI restriction sites, which were identified via sequencing, and were upstream of the TATA box and g-less cassette. However, in all further attempts to assay *in vitro* transcription, even basal levels of transcription of the g-less cassette were not observable as expected based on previous studies with the original pMLΔ53 plasmid. Further attempts to obtain the original plasmid with a Bgl2 restriction site were unsuccessful.

Conclusions

A general method for the synthesis of large quantities of polyamides functionalized for coupling to isoxazolidines has been accomplished. The entire solution-phase synthetic route is mapped out, and is likely applicable to other synthetic modulations for future studies (e.g. different polyamide core, alternative TADs, etc.). Conditions for attachment of various linkers and conjugation of several isoxazolidines moieties to polyamides have been optimized. However, based on multiple cell-culture assays, no significant levels of transcriptional upregulation have been observed to date.

While it is possible that the chosen geometry of the polyamide-isoxazolidine compounds is ineffective at activating transcription, a series of tail-conjugated polyamide-isoxazolidine compounds have been synthesized and studied by Dr. Daniel Gubler. In similar cell culture studies, these compounds also failed to show any levels of transcriptional upregulation. It is currently not fully understood exactly why polyamide-isoxazolidine conjugates were unsuccessful as artificial transcription factors, though it is speculated that the large size of these conjugates may be a negative determinant for cellular uptake, and this could mean that the levels of intracellular concentration are never high enough for significant effects to be observed. Other effects such as geometric constraints or protein binding may also be inhibiting the functionality of these compounds, and further studies will be necessary to deconvolute these various factors.

Section 5B: Synthesis of WM Conjugates

Background

In addition to studying polyamide-isoxazolidine conjugates, another project involving transcriptional activation was also undertaken: a collaboration with Professor Aseem Ansari's laboratory (University of Wisconsin). This involved the synthesis of a "WM" compound (so called because it contains tryptophan and methionine moieties) conjugated to a polyamide. These conjugates can facilitate Exd-DNA dimerization, analogous to the compound shown in Figure 5.3.

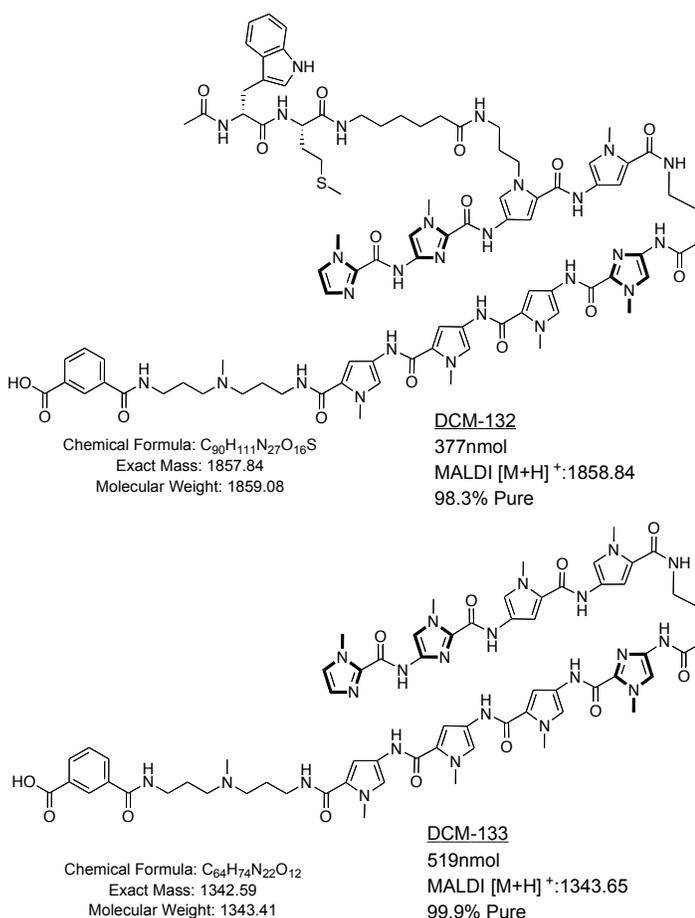


Figure 5.33: The WM compound **DCM-132** and control **DCM-133** that were synthesized. **DCM-132** contains a natural methionine and a tryptophan with D stereochemistry.

These compounds were previously synthesized by Dr. Ryan Stafford of the Dervan group, and were shipped to the Ansari laboratory. Using **RLSiv-279A** (Dr. Stafford's version of **DCM-132**) in Jurkat cells, a reporter luciferase assay was used to assay for gene activation. This method indicated a 2-fold activation of transcription at 1 μ M. To determine whether this effect was concentration dependent, more material was necessary to allow higher concentrations to be assayed. However, the previously shipped material had nearly all been used in these initial experiments. Leslie Donato of Professor Ansari's group asked if the Dervan laboratory could synthesize more of this compound, to be shipped to the Ansari laboratory.

Following Dr. Stafford's protocol, these compounds were synthesized using Boc-solid-phase polyamide synthesis, as mentioned previously. The synthesis of control compound **DCM-133** was relatively straightforward: after cleavage from resin and preparatory HPLC purification, PyBOP coupling of isophthalic acid yielded the final product, which was again purified by preparatory HPLC. In the case of **DCM-132**, a phthalimide-protected amino group linked to a pyrrole was used as one of the monomers during the synthesis, to provide the handle for peptide conjugation as seen in Figure 5.34. In addition, after cleavage from resin, and preparatory HPLC purification, three more synthetic steps were required: coupling of the peptide fragment, deprotection of the Boc group on the tail, and isophthalic acid coupling. The final products were then purified via preparatory HPLC.

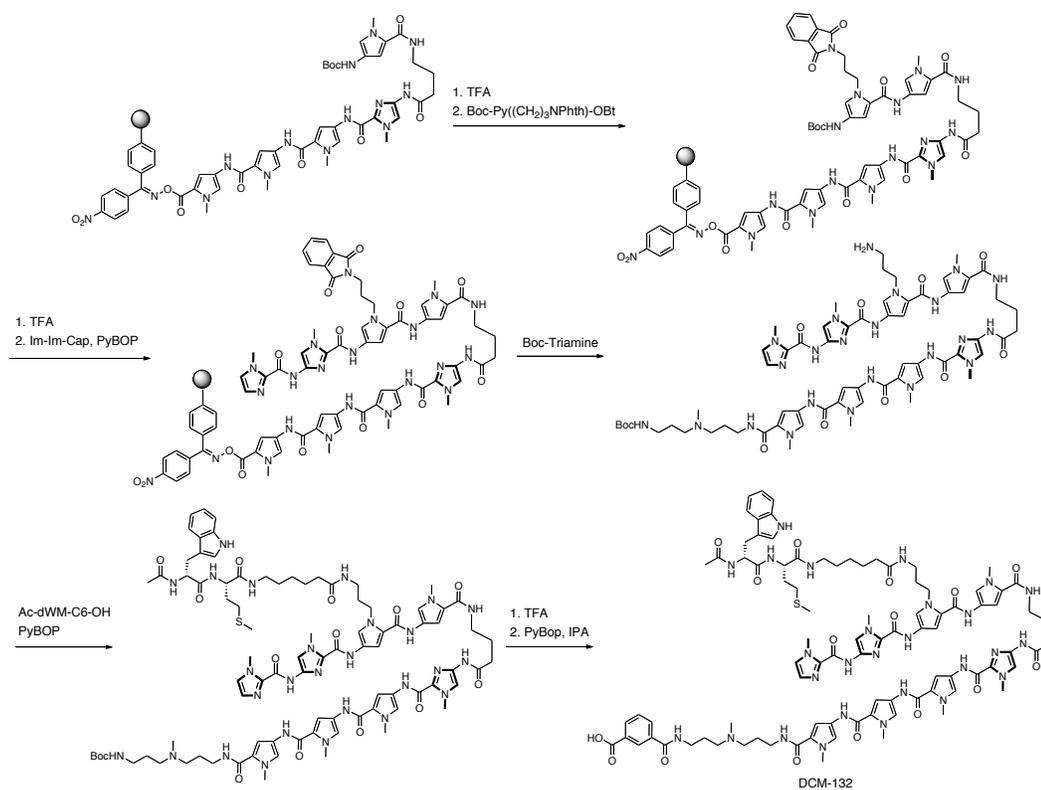


Figure 5.34: Synthesis of WM conjugate.

These compounds were synthesized in reasonable quantities (377nmol and 519nmol for **DCM-132** and **DCM-133** respectively), with comparable yields to those previously reported. After their synthesis, these compounds were sent to the University of Wisconsin. Unfortunately, further luciferase assays employing increased concentrations did not increase the luciferase signal substantially.

Materials and Methods

Polyamides were synthesized using Kaiser oxime resin (0.56 mmol/g) and benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) from Novabiochem. Z- β -Dab(Boc)-OH was purchased from Peptides International. Monomers and dimers for solid phase synthesis were synthesized using published protocols.^{24,32} *N,N*-dimethylformamide (DMF), Dichloromethane (DCM), methanol (MeOH), trifluoroacetic acid (TFA), acetic anhydride, *N,N*-diisopropylethylamine (DIEA), 2.0M HCl in Ether, succinic anhydride, 1,4-dioxane, and diethyl ether were purchased from Aldrich. 1M aqueous NaOH, 1M aqueous HCl, solid sodium sulfate (Na₂SO₄), and solid sodium bicarbonate (NaHCO₃) were purchased from EMD Chemicals.

All NMR spectroscopy data was obtained on a 500MHz Varian instrument. Analytical high-pressure liquid chromatography (HPLC) was performed using a Beckman Gold system with a Phenomenex Gemini C18 column, using 0.1% TFA/water (A) and acetonitrile (B) solvent system. Preparatory HPLC purification was performed using the same solvents on an Agilent Technologies 1200 Series system using a Phenomenex Gemini C18 column. Matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF) was performed on an Applied Biosystems Voyager DE Pro spectrometer. UV-vis were recorded in water and a minimum of acetonitrile using an Agilent 8453 spectrophotometer.

Polyamide synthesis was performed as previously reported.^{24,25} Briefly, oxime resin was swelled by flushing with DMF in a peptide synthesis vessel containing a cap, stopcock, and frit. Initial pyrrole loading of the resin occurred overnight (16 hrs) at 37° C. In general, couplings lasted 2 hours. Boc-protecting groups were removed with 20% TFA

in DCM for 25 min at r.t. Carboxylic acids (3 eq) were preactivated with PyBOP (3.5 eq) and DIEA (5 eq) for 30 min in DMF. Cleavage of the polyamide from resin was performed using 1 mL of triamine for 200 mg of resin, allowing the reaction to proceed at 55°C for 12 hours. Preparatory HPLC was used to purify polyamides after cleavage. Mono-*tert*-butyl isophthalic acid or isophthalic acid (10eq) were coupled to polyamide using PyBOP (3eq) and DIEA (5eq).

Solution Phase Polyamide Synthesis

(R)-*tert*-butyl 3-(3-((3-(4-(4-(4-(4-(3-(benzyloxycarbonylamino)-4-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)propyl)-(methylamino)propylcarbamoyl)benzoate (7). To carboxylic acid **11** (310 mg, 0.255mmol) dissolved in 1mL of DMF was added DIEA (178 μ L, 1.02mmol) and PyBOP (265mg, 0.510mmol) and allowed to stir at room temperature for 30 minutes. To this mixture was next transferred a solution of **10** (178mg, 0.510mmol) in 200 μ L DMF and DIEA (8 μ L, 2eq) and allowed to stir for 2 hours. Reaction was monitored by HPLC. Upon completion, the reaction mixture was poured into 20mL of H₂O, transferred to Eppendorf tubes, and centrifuged at 13,000rpm for 10 minutes. The tan solid was then washed twice more and centrifuged, providing 368mg of polyamide **7** (94% yield), which was 91% pure by HPLC; MALDI-TOF: Calc. Mass: 1546.70 [M+H]⁺:1547.71, [M+Na]⁺:1570.53

(R)-4-(4-(2-(5-(5-(5-(3-((3-(3-(*tert*-butoxycarbonyl)benzamido)propyl)(methylamino)-propylcarbamoyl)-1-methyl-1H-pyrrol-3-ylcarbamoyl)-1-methyl-1H-pyrrol-3-ylcarbamoyl)-1-methyl-1H-pyrrol-3-ylcarbamoyl)-1-methyl-1H-imidazol-4-ylamino)-1-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-4-oxobutan-2-ylamino)-4-oxobutanoic acid (9). To 4.9mg of Pd/C (1eq) in a glass vial fitted with a teflon septum was added Cbz-protected **7** (5 μ mol, 1 eq) dissolved in 850 μ L of 1,4-Dioxane and 800 μ L of water.

To this mixture was added ammonium formate (31.5mg, 200eq), and the solution was stirred overnight at 60°. Reaction was monitored by analytical HPLC, and transferred to Eppendorf tubes. After centrifugation at 13,000rpm, supernatant was collected. Pellet was washed twice with 4:1 1,4-dioxane:water, these samples were centrifuged and supernatants combined. After a final centrifuge spin, supernatant was poured into a scintillation vial, and excess solvent was removed via rotary evaporation, then placed on the high vacuum line for 4 hours. Next, 500µL of DMF were added along with DIEA (43.5µL, 50eq). To this mixture was added succinic anhydride (2.5mg, 5eq), and stirred at room temperature for 1 hour. Reaction was monitored via HPLC. Preparatory HPLC purification yielded 3.0 mg of **9** (87%), which was 98% pure by HPLC; MALDI-TOF: Calc. Mass:1512.68 [M+H]⁺:1513.69.

Tert-butyl 3-(3-((3-aminopropyl)(methyl)amino)propylcarbamoyl)benzoate (10). To a solution of **15** in 15mL DCM was added HOBt (334mg, 1.1eq), then DCC (510mg, 1.1eq), and this solution was stirred for 1 hr. Solution filtered to remove urea byproduct, then diluted by the addition of 250mL DCM, and added to an addition funnel. This was dripped into a solution of triamine **16** (7.25mL, 20eq) and 5mL DCM over 3.5 hours. Rotary evaporation, and extraction to remove excess **16**, yielded 362mg of **10**. ¹H NMR [499.8 MHz, DMSO-d₆]: δ 8.68 (t, 1H), 8.31 (s, 1H), 8.01 (m, 2H), 7.58 (dd, 1H), 3.28 (m, 2H), 2.32 (m, 4H), 2.05 (s, 3H), 1.67 (m, 2H), 1.52 (s, 9H), 1.50 (m, 2H), 1.46 (m, 2H). ESI-MS: Calc. Mass: 349.2 [M+H]⁺:350.3.

(R)-4-(4-(4-(4-(3-(benzyloxycarbonylamino)-4-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxylic acid (11). To 50mg of **25** in a scintillation vial was added 1mL 1,4-dioxane, then 1mL 1N aqueous NaOH. Solution was stirred at r.t. for 15 hours, then poured into Falcon tube. 1N aqueous HCl was added until pH<4, and solid **11** precipitates to form a tan brown gel. Centrifugation was performed, the supernatant poured off, and the solid was washed twice with cold water, each time centrifuging and removing supernatant. Product was next lyophilized to yield 42 mg of **11** (86%). 91% pure by HPLC. MALDI-TOF: Calculated Mass: 1215.47 [M+H]⁺:1216.65 [M+Na]⁺:1238.63 [M+K]⁺:1254.59.

1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxylic acid (12). See procedure for compound **11**. 98%

yield. ^1H NMR [499.8 MHz, DMSO- d_6]: δ 10.47 (s, 1H), 9.97 (s, 1H), 9.92 (s, 1H), 7.45 (s, 1H), 7.40 (s, 1H), 7.27 (s, 1H), 7.22 (s, 1H), 7.18 (s, 1H), 7.06 (s, 1H), 7.04 (s, 1H), 6.91 (s, 1H), 3.99 (s, 3H), 3.81 (s, 3H), 3.79 (s, 3H). ESI-MS: Calc. Mass: 492.2 $[\text{M}+\text{H}]^+$:493.2.

((R)-ethyl 4-(4-(4-(4-(3-(benzyloxycarbonylamino)-4-(tert-butoxycarbonylamino)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxylate (24). To a solution of Z- β -Dab(Boc)-OH (350mg, 1.33eq) in 252 μL DIEA (2eq) and 10mL DMF was added PyBOP (517mg, 1.33eq), and this solution was stirred for 30 minutes. A separate solution of **23** mixed with 252 μL DIEA (2eq) and 5mL DMF was then added to the first solution. Mixture was stirred at room temperature for 16 hours. Rotary evaporation removed as much DMF as possible, and this viscous oil was purified using silica gel chromatography, eluted with 9:1 EtOAc:Hexanes, yielding 506mg of **24** (78%). ^1H NMR [499.8 MHz, DMSO- d_6]: δ 10.19 (s, 1H), 10.02 (s, 1H), 9.97 (s, 1H), 9.92 (s, 1H), 7.45 (s, 1H), 7.43 (s, 1H), 7.35-7.23 (m, 6H), 7.13 (s, 1H), 7.09-7.02 (m, 2H), 6.90 (s, 1H), 4.98 (s, 2H), 4.20 (q, 2H), 3.96 (s, 3H), 3.89-3.80 (m, 9H), 3.05 (m, 2H), 2.47 (m, 2H), 1.37 (s, 9H), 1.26 (t, 3H). ESI-MS: Calc. Mass: 869.4 $[\text{M}+\text{H}]^+$:870.3.

(R)-ethyl 4-(4-(4-(4-(4-amino-3-(benzyloxycarbonylamino)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxylate (13). To a stirred solution of **24** (274mg, 1eq) in DCM (5mL) was added 2M HCl in Et₂O (5mL). The resulting suspension was stirred for 7 hours at room temperature. Reaction mixture was concentrated via rotary evaporation, affording 251mg of **13** (99% yield). ^1H NMR [499.8 MHz, DMSO- d_6]: δ 10.47 (s, 1H), 9.97 (s, 1H), 9.92 (s, 1H), 7.45 (s, 1H), 7.40 (s, 1H), 7.27 (s, 1H), 7.22 (s, 1H), 7.18 (s, 1H), 7.06 (s, 1H), 7.04 (s, 1H), 6.91 (s, 1H), 3.99 (s, 3H), 3.81 (s, 3H), 3.79 (s, 3H). ESI-MS: Calc. Mass: 769.3 $[\text{M}+\text{H}]^+$:770.1.

3-(tert-butoxycarbonyl)benzoic acid (15) Compound was synthesized as reported in the literature.²⁶ Briefly, to a solution of isophthaloylchloride (30g, 0.15mol) in 250mL of THF was added *tert*-butanol (13.8mL, 0.15mol) and pyridine (26.3mL, 0.33mol), and this mixture was refluxed for 6 hours. Next, D.I. H₂O was added, and the mixture was refluxed again for 6 hours. After quenching, extraction, and concentration, recrystallization from toluene yielded 16.6g of large needle like crystals. (52% yield). ^1H NMR [499.8 MHz,

DMSO-d₆]: δ 8.39 (s, 1H), 8.12 (m, 2H), 7.62 (dd, 1H), 1.54 (s, 9H). ESI-MS: Calc. Mass: 222.1 [M+H]⁺:223.1.

Ethyl 1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxylate (21). To trimer **20** (400mg, 1eq) in 5mL DMF and 304 μ L of DIEA (2eq) was added Im-CCl₃ (228mg, 0.95eq) dissolved in 5mL DMF and 304 μ L of DIEA (2eq). This solution was stirred 16 hours at room temperature. Reaction mixture was added to 3.6mL 1M aqueous HCl (4eq) in 200mL of water, creating a dark brown precipitate. Solid was filtered and washed with copious amounts of water. Solid was transferred to a separatory funnel along with 80mL of 1%MeOH in EtOAc. Organic layer was washed with 20mL brine, dried over sodium sulfate, filtered, and the solvent removed via rotary evaporation. ¹H NMR [499.8 MHz, DMSO-d₆]: δ 10.47 (s, 1H), 9.97 (s, 1H), 9.92 (s, 1H), 7.45 (s, 1H), 7.40 (s, 1H), 7.27 (s, 1H), 7.22 (s, 1H), 7.18 (s, 1H), 7.06 (s, 1H), 7.04 (s, 1H), 6.91 (s, 1H), 4.39 (q, 2H), 3.99 (s, 3H), 3.81 (s, 3H), 3.79 (s, 3H), 1.27 (t, 3H). ESI-MS: Calc. Mass: 520.2 [M+H]⁺:521.4.

Ethyl 4-(4-(4-(4-(tert-butoxycarbonylamino)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxylate (22). To a solution of Boc-Im-COOH (102.7mg, 0.95eq) dissolved in 4.5mL DMF and 152 μ L of DIEA (2eq) was added HBTU (162mg, 0.95eq). Solution stirred at room temperature for 20 minutes. This solution was added to trimer **20** (200mg, 1eq) in 3mL DMF and 152 μ L of DIEA (2eq). This mixture was stirred 16 hours at room temperature. Reaction mixture was added to 1.8mL 1M aqueous HCl (4eq) in 93mL of water, creating a dark brown precipitate. Solid was filtered and washed with copious amounts of water. Purified via silica gel chromatography, to yield 220mg of **22** (81%). ¹H NMR [499.8 MHz, DMSO-d₆]: δ 9.95 (s, 3H), 9.92 (s, 1H), 7.43 (s, 1H), 7.27 (s, 1H), 7.24 (m, 2H), 7.14 (s, 1H), 7.08 (s, 1H), 6.94 (s, 1H), 4.20 (q, 2H), 3.93 (s, 1H), 3.82 (m, 3H), 1.46 (s, 9H), 1.27 (t, 3H). ESI-MS: Calc. Mass: 635.3 [M+H]⁺:636.3.

Ethyl 4-(4-(4-(4-amino-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxylate (20). To Boc-protected compound **22** (431mg, 0.678mmol) was added 2.5mL of 1,4-Dioxane. This mixture was added to 40mL of 2.0M HCl/Et₂O, forming a suspension, which was stirred for 3 hours. Solid was collected via vacuum filtration and left under high vacuum to yield 376mg of **23** (97%). ¹H NMR [499.8 MHz, DMSO-d₆]: δ 9.95 (s, 3H), 7.43

(s, 1H), 7.27 (s, 1H), 7.24 (m, 2H), 7.14 (s, 1H), 7.08 (s, 1H), 6.94 (s, 1H), 4.20 (q, 2H), 3.93 (s, 1H), 3.82 (m, 3H), 1.27 (t, 3H). ESI-MS: Calc. Mass: 448.2 [M+H]⁺:449.1.

(R)-ethyl 4-(4-(4-(4-(3-(benzyloxycarbonylamino)-4-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxylate (25). To carboxylic acid **12** (35mg, 1eq) was added 500μL DMF, 25μL DIEA (4eq), and PyBOP (37mg, 1eq), and the mixture was stirred for 20 minutes. To this a solution of **13** (58mg, 1eq) along with 500μL DMF and 25μL DIEA were added. Stirred at room temperature for 16 hours. Reaction mixture poured into a solution of 0.58mL 1M HCl (8eq) in 20mL water. A suspension formed, and the solid was centrifuged and lyophilized to yield 68mg of **25** (76% yield). Compound was 85% pure by HPLC. MALDI-TOF: Calculated Mass: 1243.51, [M+H]⁺:1244.51, [M+Na]⁺:1266.49, [M+K]⁺:1282.46.

Isioxazolidine Coupling Reactions

Tert-butyl 3-(3-((3-(4-(4-(4-(4-((R)-3-(4-(((3S,5R)-3-allyl-2-benzyl-3-isobutylisoxazolidin-5-yl)methylamino)-4-oxobutanamido)-4-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)propyl)(methyl)amino)propylcarbamoyl)benzoate (29). To isioxazolidine **RJC 6-1** (5.6mg, 5eq) was added 100μL of DCM. 100μL of TFA was next added dropwise, and the resulting mixture was stirred for 5 minutes. After concentration via rotary evaporation, 100μL of DMF was added, followed by DIEA (2.25μL, 5eq). In a separate vial, 2.59μmol of polyamide **9** was dissolved in 200μL DMF, and DIEA (0.9μL, 2 eq) followed by PyBOP (2.7mg, 2eq) were added, and the mixture was stirred for 2 minutes. To this solution was added the solution of deprotected isioxazolidine, and the resulting mixture was stirred for 16 hours at 23°C. After determination of completion by analytical HPLC, the solution was poured into 5mL H₂O. This solution was poured into a Waters C18 Sep-Pak, and flushed with 10% acetonitrile in H₂O with 0.1% TFA to remove any DMF and other highly polar contaminants. A 1:1 acetonitrile: H₂O w/ 0.1% TFA wash

was then used to elute the desired product. The resulting solution was next lyophilized to dryness directly into a ½ dram glass vial, and the crude material was used directly in the following step. Crude **29** was 64% pure by HPLC, and was directly used in the following step.

Tert-butyl 3-(3-((3-(4-(4-(4-(4-((R)-3-(4-(((3S,5R)-2-benzyl-3-(2-hydroxyethyl)-3-isobutylisoxazolidin-5-yl)methylamino)-4-oxobutanamido)-4-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)propyl)(methyl)amino)propylcarbamoyl)benzoate (**26**). See procedure for compound **29**.

2.1 μmol of **9** used, with 2.9 mg (3.5eq) of **RJC 6-3** used. Crude material carried forward. 55% pure by HPLC.

3-(3-((3-(4-(4-(4-(4-((R)-3-(4-(((3S,5R)-3-allyl-2-benzyl-3-isobutylisoxazolidin-5-yl)methylamino)-4-oxobutanamido)-4-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)propyl)(methyl)amino)propylcarbamoyl)benzoic acid (**ATF-4**).

To crude **29** lyophilized in a glass vial was added acetonitrile (500 μL) and the resulting suspension was stirred. To this suspension was added NaI (5.75mg, 15eq), followed by CeCl₃•7H₂O (29mg, 30eq). This solution was refluxed for 1 hour, at which point it was poured into 3mL H₂O, and partially purified via Waters C18 Sep-Pak as in previous step. The resulting solution was concentrated to remove most of the acetonitrile, and taken up in 1:4 acetonitrile:H₂O with 0.1% TFA, and purified via preparatory HPLC, yielding 543nmol (21% yield from **9**) of ATF-4. MALDI-TOF: Calc. Mass: 1726.83, [M+H]⁺:1727.47. 99% pure by HPLC.

3-(3-((3-(4-(4-(4-(4-((R)-3-(4-(((3S,5R)-2-benzyl-3-(2-hydroxyethyl)-3-isobutylisoxazolidin-5-yl)methylamino)-4-oxobutanamido)-4-(1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)butanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)propyl)(methyl)amino)propylcarbamoyl)benzoic acid (**ATF-1**). See procedure for synthesis of

ATF-4. Isolated 352nmol (17% yield from **9**) MALDI-TOF: Calc. Mass: 1730.82, $[M+H]^+$:1731.65. 99% pure by HPLC

Tert-butyl 3-(3-((3-(4-(4-(4-(4-((R)-1-((3S,5R)-2-benzyl-3-(2-hydroxyethyl)-3-isobutylisoxazolidin-5-yl)-17-((1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)methyl)-3,12,15-trioxo-5,8-dioxo-2,11,16-triazanonadecanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)propyl)(methyl)amino)propylcarbamoyl)benzoate (**27**). See procedure for synthesis of **29**.

Used 482nmol of **9**, and 0.32mg isoxazolidine (2eq). Crude material carried forward. 37% pure by HPLC.

3-(3-((3-(4-(4-(4-(4-((R)-1-((3S,5R)-2-benzyl-3-(2-hydroxyethyl)-3-isobutylisoxazolidin-5-yl)-17-((1-methyl-4-(1-methyl-4-(1-methyl-4-(1-methyl-1H-imidazole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)-1H-pyrrole-2-carboxamido)methyl)-3,12,15-trioxo-5,8-dioxo-2,11,16-triazanonadecanamido)-1-methyl-1H-imidazole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)-1-methyl-1H-pyrrole-2-carboxamido)propyl)(methyl)amino)propylcarbamoyl)benzoic acid (**ATF-2**). To lyophilized **27** was added 100 μ L of 1:1 DCM:TFA. This solution was agitated for 10 minutes, then concentrated via rotary evaporation. To the resulting residue was added 1:4 acetonitrile:H₂O with 0.1% TFA, and preparatory HPLC chromatography afforded 26nmol of **ATF-2** (6% yield from **9**). MALDI-TOF: Calc. Mass:1875.90, $[M+H]^+$:1877.17. 97% pure by HPLC.

Isoxazolidine Scale-up

tert-butyl hypochlorite (**32**). A solution of commercial household bleach (500mL) was stirred at 0°C in a large Erlenmeyer flask. The flask was covered with aluminum foil and lights in the vicinity were turned off, while a solution of *tert*-butanol (37mL, .39mol) and glacial acetic acid (24.5mL, .43mol) was added in a single portion. This solution was stirred for 3 minutes, then washed with 10% NaHCO₃ followed by water, and the organic layer was collected and dried over CaCl₂, and used directly in subsequent steps without further purification, affording 30g of *tert*-butyl hypochlorite (71% yield). Prepared according to the method of Mintz and Walling.³³

(E)-3-methylbutanal oxime (34). Followed procedure published by Kubanov et al.³⁴ Briefly, aldehyde **33** (20g, 0.23mol) was refluxed in 200mL Benzene, with hydroxylamine hydrochloride (16.9g, 0.244mol) and Na₂CO₃ (25.819g, 0.244mol) for 5 hours. After extraction, the residue was purified via fractional distillation to yield 11.2g of oxime **34** (48% yield). ¹H NMR [499.8 MHz, CDCl₃]: δ 7.45, 6.76 (t, 1H), 2.10 (dd, 2H), 1.92-1.81 (m, 1H), 0.96 (m, 6H).

(R)-(3-isobutyl-4,5-dihydroisoxazol-5-yl)methanol (35). To a stirring solution of oxime **34** (3g, 29.6 mmol) in 60mL of toluene at -78°C was added **32** (3.35mL, 29.6mmol) drop wise and this mixture was stirred for 2 hours. In a separate flask, allyl alcohol (2.62mL, 38.5mmol) was stirred at 0°C in 120mL of toluene. To this mixture was added *tert*-BuOH (9.27 mL, 97.7 mmol) followed by drop wise addition of a 3.0M solution of EtMgBr in Et₂O (29.6 mL, 88.8mmol), and this mixture was stirred for 1 h. The solution of hydroximinoyl chloride was then transferred via cannula to the allylic alcohol solution and the mixture allowed to warm to room temperature and stirred for 15 h. After quenching with NH₄Cl and H₂O, The organic layer was washed with DCM (3X) and the combined organic extracts were combined, dried, concentrated, and purified via silica gel chromatography. This provided 3.58g of isoxazoline **35** (77% yield). ¹H NMR [499.8 MHz, CDCl₃]: δ 4.62-4.68 (m, 1H), 3.76 (dd, 1H), 3.54 (dd, 1H), 2.93 (dd, 1H), 2.80 (dd, 1H), 2.21-2.19 (m, 2H), 1.92-1.86 (m, 2H), 0.95 (d, 3H), 0.93 (d, 3H).

((3S,5R)-3-allyl-3-isobutylisoxazolidin-5-yl)methanol (36). To a solution of isoxazoline **35** (200mg, 1.27mmol) in 10mL of toluene cooled to -78°C was added BF₃•OEt₂ (484μL, 3.82mmol) drop wise, and the solution was stirred for 30 minutes. A 2.0M solution of allylmagnesium chloride (4.45mL, 8.91mmol) was next added drop wise. After stirring for 4 hours, saturated NaHCO₃ was added, and the solution was extracted with EtOAc. The organic extracts were combined, washed with H₂O and brine, dried over Na₂SO₄, concentrated, and purified via flash column chromatography, yielding 230mg of isoxazolidine **36** (91% yield). ¹H NMR [499.8 MHz, CDCl₃]: δ 5.87-5.80 (m, 1H), 5.12-5.08 (m, 2H), 4.21-4.17 (m, 1H), 3.69 (dd, 1H), 3.53 (dd, 1H), 2.36 (dd, 1H), 2.24-2.19 (m, 2H), 1.84-1.77 (m, 1H), 1.62 (dd, 1H), 1.44 (dd, 1H), 1.38 (dd, 1H), 0.94 (d, 3H), 0.92 (d, 3H). ESI-MS: Calc. Mass: 199.2 [M+Na]⁺:222.2.

(3S,5R)-3-allyl-5-((tert-butylidimethylsilyloxy)methyl)-3-isobutylisoxazolidine (37). To a solution of isoxazolidine **36** (170mg, 0.85mmol) in 50mL of THF at 0°C were added DMAP (5.3mg, 0.0427mmol) and Et₃N (166 μL, 1.19mmol), followed by drop wise addition of TBSOTf (274μL, 1.19mmol). The mixture was

next warmed to room temperature, and stirred for 2 hours. It was next cooled to 0°C and quenched with saturated NH₄Cl, and extracted with Et₂O. The combined organic extracts were next washed with brine, dried, and concentrated via rotary evaporation. Silica gel chromatography yielded 251mg of **37** (94% yield). ¹H NMR [499.8 MHz, CDCl₃]: δ 4.63–4.59 (m, 1H), 3.69–3.62 (m, 2H), 3.56 (dd, 1H), 2.88 (d, 2H), 2.37–2.25 (m, 2H), 1.14 (t, 3H), 0.88 (s, 9H), 0.85 (s, 9H), 0.07 (s, 3H), 0.06 (s, 3H), 0.04 (s, 3H), 0.04 (s, 3H). ESI-MS: Calc. Mass: 313.2 [M+H]⁺:314.2.

((3S,5R)-3-allyl-2-benzyl-3-isobutylisoxazolidin-5-yl)methanol (38). To a solution of **37** (220mg, 0.702mmol) in 2 mL of DMF was added DIEA (367μL, 2.10mmol) followed by BnBr (417μL, 3.508mmol). The solution was next irradiated in a 1000 W microwave (5 x 15 s) @ 30% power with mixing after each cycle. The reaction was then diluted with H₂O, and extracted with Et₂O (3X). The organic extracts were combined and washed with H₂O and brine, dried over Na₂SO₄, filtered, and concentrated via rotary evaporation. The crude residue was next dissolved and stirred in 1mL THF, and a 1.0M solution of TBAF (1.05mL, 1.05mmol) was added drop wise. This solution was stirred for 4 hours, at which point H₂O was added, and it was extracted with EtOAc (3X). The combined organic extracts were washed with brine, dried, and concentrated. Silica gel chromatography provided pure 166mg of pure **38** (82% yield over two steps). ¹H NMR [499.8 MHz, CDCl₃]: δ 7.40-7.20 (m, 5H), 5.98-5.88 (m, 1H), 5.09-5.13 (m, 2H), 4.02-4.10 (m, 1H), 3.90 (d, 1H), 3.83 (d, 1H), 3.61-3.54 (m, 2H), 2.45 (dd, 1H), 2.28 (dd, 1H), 2.18-2.24 (br s, 1H), 2.01-2.08 (m, 1H), 1.87-1.93 (m, 1H), 1.62 (dd, 1H), 1.39 (dd, 1H), 0.98 (d, 3H), 0.97 (d, 3H). ESI-MS: Calc. Mass: 289.2 [M+H]⁺:290.1.

Linker Synthesis

2-(2-(2-aminoethoxy)ethoxy)-N-(((3S)-2-benzyl-3-(2-hydroxyethyl)-3-isobutylisoxazolidin-5-yl)methyl)acetamide (RJC-6-6). To a solution of Fmoc-AEEA-COOH (2.6mg, 6.7μmol) in 100μL DMF was added DIEA (3.4μL, 20μmol), followed by PyBOP (3.56mg, 6.8μmol), and this mixture was stirred for 30 minutes at room temperature. Next, a solution of RJC-6-3dep (1mg, 3.4μmol) with DIEA (3.4μL, 20μmol) was added, and this mixture allowed to stir for 16 hours. This mixture was extracted with ethyl acetate, and organic fractions were concentrated. To this crude residue was added 300μL of 20% piperidine in DMF, and

the resulting solution stirred for 30 minutes. Next, the mixture was diluted with methanol, extracted with pentane, and concentrated, followed by preparatory HPLC purification, providing 0.4mg of **RJC-6-6** (27% yield). ¹H NMR [499.8 MHz, CD₃OD]: 7.38-7.21 (m, 5H), 4.27-4.17 (m, 1H), 4.00-4.90 (m, 4H), 3.79-3.72 (m, 3H), 3.61-3.43 (m, 6H), 3.37 (dd, 1H), 3.00 (t, 2H), 2.45 (dd, 1H), 2.07-1.98 (m, 1H), 1.90 (t, 2H), 1.85-1.66 (m, 2H), 1.45 (dd, 1H), 1.0 (d, 3H), 0.96 (d, 3H). ESI-MS: Calc. Mass: 437.3 [M+H]⁺:438.3.

tert-butyl 1-((3S)-2-benzyl-3-(2-hydroxyethyl)-3-isobutylisoxazolidin-5-yl)-3-oxo-6,9,12,15-tetraoxa-2-azaheptadecan-17-ylcarbamate (41). To a solution of 2,2-dimethyl-4-oxo-3,8,11,14,17-pentaoxa-5-azaicosan-20-oic acid (12.5mg, 34.3μmol) in 300μL DMF was added DIEA (11.91μL, 68.4μmol) and PyBOP (18.2mg, 35.0μmol). This solution was stirred at 45°C for 30 minutes. Subsequently, a solution of RJC-6-3dep (5mg, 17μmol) with DIEA (5.96μL, 34.2μmol) was added to the activated PEG linker solution. This mixture was allowed to stir overnight. Water and brine were added, and the mixture was extracted with EtOAc (3X), dried, and concentrated. Purification via preparatory TLC afforded 8.1mg of the desired compound **41** (74% yield). ¹H NMR [499.8 MHz, CDCl₃]: 7.38-7.24 (m, 5H), 4.12-4.05 (s, 1H), 3.94-3.81 (m, 2H), 3.72-3.44 (m, 8H), 3.41-3.30 (m, 2H), 3.28-3.20 (m, 1H), 2.46-2.19 (m, 4H), 1.93-1.80 (m, 2H), 1.68-1.59 (m, 1H), 1.44 (s, 9H), 1.41-1.32 (m, 1H), 1.24 (s, 2H), 0.98 (t, 6H). ESI-MS: Calc. Mass: 639.4 [M+H]⁺:640.3

Cell Culture

All cell lines were purchased from ATCC (Manassas, VA) and maintained in the following media: A549 cells (F-12K); LNAR+ (RPMI 1640); HCT-116 (McCoy's 5a Medium Modified); MCF-7 (Eagle's Minimum Essential Medium). All media were supplemented with 10% FBS and cultured at 37°C under 5% CO₂.

LNAR+ Luciferase Experiments

LNAR+ cells were plated in 96-well plates at a density of 50,000 cells per mL (5,000 cells/well) in RPMI 1640 medium (ATCC) supplemented with 10% FBS (Irvine Scientific, Santa Ana, CA) and 4 mM penicillin/streptomycin. After 24 h, the medium

was replaced with RPMI 1640 medium containing 10% charcoal-stripped FBS, 4 mM penicillin/streptomycin, and polyamides at varying concentrations in quadruplet. Cells were grown for an additional 34 h before treatment of induced wells with 1nM dihydrotestosterone (DHT) for 6 h. After 40 hours total polyamide treatment, 10 μ L WST-1 reagent was added, and cells were incubated at 37°C with gentle shaking for 30 min, before measuring the absorbance at 450 nm. Next, cells were lysed by addition of 100 μ L of Glo Lysis Buffer (Promega), incubated 5 minutes, then 100 μ L of each well was added to a separate opaque-walled 96-well plate containing 100 μ L Bright-Glo™ reagent, mixed briefly and gently, and immediately imaged for luminescence.

Measurement of dexamethasone-induced mRNA.

RNA isolation: A549 cells (ATCC) were plated in 24-well plates at a density of 20-25 x 10³ cells per well (40-50 x10³ cells per mL) in F12-K medium (ATCC) supplemented with 10% FBS (Irvine Scientific, Santa Ana, CA) and 4 mM penicillin/streptomycin. After 24 h, the medium was replaced with F12-K containing 10% charcoal-stripped FBS, 4 mM penicillin/streptomycin, and polyamides or mifepristone at the designated concentrations. Cells were grown for an additional 48 h and then treated with 100 nM dexamethasone for 6 hours. The medium was removed, cells were washed with ice-cold PBS and immediately lysed with RLT buffer from an RNeasy kit (Qiagen). Further RNA isolation was carried out with the RNeasy kit as described in the manufacturer's manual. The isolated total RNA was quantified. The yields were 12-15 μ g per well.

Reverse transcription: A 2.5 µg sample of total RNA was used to reverse transcribe cDNA using Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's protocol. Random hexamers were used as primers. The total volume for each RT reaction was 20 µL.

Real-time quantitative RT-PCR: Analysis was performed using the GILZ gene primers described below, purchased from Integrated DNA Technologies. Quantitative real-time RT-PCR was performed with SYBR 66 Green PCR Master Mix (Applied Biosystems, Foster City, CA) following the manufacturer's suggested protocol. Temperature cycling and detection of the SYBR Green emission were performed with an ABI 7300 real-time instrument using Applied Biosystems Sequence Detection System version 1.2. Statistical analysis was performed on three independent experiments. mRNA of the genes of interest were measured relative to β -glucuronidase as an endogenous control. Primer sequences were designed using Primer3.

To amplify the 97-bp fragment from the 3'-translated region of GILZ:

Forward primer: 5'- CTCCCCGTTTGTTTTTCTCA -3'

Reverse primer: 5'- TGCTCCTTCAGGATCTCCAC -3'

To amplify the β -glucuronidase gene as an endogenous control:

Forward primer: 5'- CTCATT TGGAATTTTGCCGATT -3'

Reverse primer: 5'- CCGAGTGAAGATCCCCTTTTTA -3'

Acknowledgements

I would like to thank David Chenoweth for assistance with the solution phase scale up. I would also like to thank Fei Yang and Katy Muzikar for a great deal of assistance with cell culture assays, and furthermore would like to thank Fei for assistance with plasmid work.

References

1. Mapp, A. K.; Ansari, A. Z., A TAD further: Exogenous control of gene activation. *ACS Chem. Biol.* **2007**, *2*, 62-75.
2. Darnell, J. E., Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2002**, *2*, 740-749.
3. Dervan, P. B., Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.* **2001**, *9*, 2215-2235.
4. Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B., Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.* **2004**, *32*, 2802-2818.
5. Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B., Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 12063-12068.
6. Edayathumangalam, R. S.; Weyermann, P.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., Molecular recognition of the nucleosomal "super groove". *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 6864-6869.
7. Dudouet, B.; Burnett, R.; Dickinson, L. A.; Wood, M. R.; Melander, C.; Belitsky, J. M.; Edelson, B.; Wurtz, N.; Briehn, C.; Dervan, P. B.; Gottesfeld, J. M., Accessibility of nuclear chromatin by DNA binding polyamides. *Chem. Biol.* **2003**, *10*, 859-867.
8. Suto, R. K.; Edayathumangalam, R. S.; White, C. L.; Melander, C.; Gottesfeld, J. M.; Dervan, P. B.; Luger, K., Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.* **2003**, *326*, 371-380.
9. Gottesfeld, J. M.; Turner, J. M.; Dervan, P. B., Chemical approaches to control gene expression. *Gene Expression* **2000**, *9*, 77-91.
10. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.* **2007**, *2*, 561-571.
11. Nickols, N. G.; Dervan, P. B., Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 10418-10423.
12. Panne, D.; Maniatis, T.; Harrison, S. C., An atomic model of the interferon-beta enhanceosome. *Cell* **2007**, *129*, 1111-1123.
13. Mapp, A. K.; Ansari, A. Z.; Ptashne, M.; Dervan, P. B., Activation of gene expression by small molecule transcription factors. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3930-3935.
14. Ansari, A. Z.; Mapp, A. K.; Nguyen, D. H.; Dervan, P. B.; Ptashne, M., Towards a minimal motif for artificial transcriptional activators. *Chem. Biol.* **2001**, *8*, 583-592.

15. Arora, P. S.; Ansari, A. Z.; Best, T. P.; Ptashne, M.; Dervan, P. B., Design of artificial transcriptional activators with rigid poly-L-proline linkers. *J. Am. Chem. Soc.* **2002**, *124*, 13067-13071.
16. Kwonj, Y.; Arndt, H. D.; Qian, M.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M., Small molecule transcription factor mimic. *J. Am. Chem. Soc.* **2004**, *126*, 15940-15941.
17. Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B., Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res.* **2007**, *35*, 363-370.
18. Dervan, P. B.; Edelson, B. S., Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299.
19. Xiao, X. S.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T., A cell-permeable synthetic transcription factor mimic. *Angew. Chem. Int. Ed.* **2007**, *46*, 2865-2868.
20. Minter, A. R.; Brennan, B. B.; Mapp, A. K., A small molecule transcriptional activation domain. *J. Am. Chem. Soc.* **2004**, *126*, 10504-10505.
21. Casey, R. J.; Desaulniers, J. P.; Hojfeldt, J. W.; Mapp, A. K., Expanding the repertoire of small molecule transcriptional activation domains. *Bioorg. Med. Chem.* **2009**, *17*, 1034-1043.
22. Rowe, S. P.; Casey, R. J.; Brennan, B. B.; Buhrlage, S. J.; Mapp, A. K., Transcriptional up-regulation in cells mediated by a small molecule. *J. Am. Chem. Soc.* **2007**, *129*, 10654-10655.
23. Dose, C.; Farkas, M. E.; Chenoweth, D. M.; Dervan, P. B., Next generation hairpin polyamides with (R)-3,4-diaminobutyric acid turn unit. *J. Am. Chem. Soc.* **2008**, *130*, 6859-6866.
24. Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B., Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.* **2002**, *10*, 2767-2774.
25. Baird, E. E.; Dervan, P. B., Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146.
26. Maeda, K.; Yamamoto, N.; Okamoto, Y., Helicity induction of poly(3-carboxyphenyl isocyanate) by chiral acid-base interaction. *Macromolecules* **1998**, *31*, 5924-5926.
27. Chenoweth, D. M.; Harki, D. A.; Dervan, P. B., Solution-Phase Synthesis of Pyrrole-Imidazole Polyamides. *J. Am. Chem. Soc.* **2009**, *131*, 7175-7181.
28. Herman, D. M.; Turner, J. M.; Baird, E. E.; Dervan, P. B., Cycle polyamide motif for recognition of the minor groove of DNA. *J. Am. Chem. Soc.* **1999**, *121*, 1121-1129.
29. Melander, C.; Herman, D. M.; Dervan, P. B., Discrimination of A/T sequences in the minor groove of DNA within a cyclic polyamide motif. *Chem. Eur. J.* **2000**, *6*, 4487-4497.
30. Gowda, D. C.; Abiraj, K., Heterogeneous catalytic transfer hydrogenation in peptide synthesis. *Lett. Pept. Sci.* **2002**, *9*, 153-165.

31. Muzikar, K. A.; Nickols, N. G.; Dervan, P. B., Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc. Natl. Acad. Sci. USA* **2009**, 106, 16598-16603.
32. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., Shape selective recognition of T·A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues. *Bioorg. Med. Chem.* **2003**, 11, 4333-4340.
33. Mintz, M. J.; Walling, C., t-Butyl hypochlorite. *Org. Synth. Coll.* **1973**, 5, 183.
34. Kurbanov, S.; Sirit, A.; Sen, N., Investigation of the alkylation of aldoximes and ketoximes with alpha-epoxides. *Org. Prep. Proced. Int.* **1999**, 31, 681-688.