

**Biological Activity of a Py-Im Polyamide Androgen Receptor Antagonist**

Thesis by

John W. Phillips

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

California Institute of Technology  
Pasadena, California  
2011

(Defended May 31<sup>st</sup>, 2011)

© 2011

John W. Phillips

All Rights Reserved

*To Valerie.*

*For all those late nights and long weekends.*

**Acknowledgements.**

I would like to thank Peter Dervan for providing a superlative training environment for my six years at Caltech. The mentorship and teaching philosophy you have demonstrated over the years have been invaluable. I would also like to thank the members of my committee, Dennis Dougherty, Bil Clemons, and Judy Campbell for their support and interest in my graduate work.

I have had the privilege of working with supremely talented colleagues in the Dervan lab, for which I am also grateful. I would particularly like to thank my intramural collaborators: Carey Hsu, Jim Puckett, Michelle Farkas, Christian Döse, Dave Chenoweth, Dan Harki, and Ben Li. Thanks also to Nick Nickols, a senior graduate student who gave me the best introduction to graduate-level research that I could have ever hoped for. Jevgenij Raskatov also deserves my gratitude for participating in our many scintillating scientific discussions.

I would also like to thank Kenneth Karanja, my collaborator and coauthor in the Campbell lab. His expertise, keen insight, and unexpected interest helped me close the final chapter of my graduate research, and just in the nick of time, too.

Caltech is home to a number of staff scientists whose fine work has contributed to this thesis. I would like to single out Shelley Diamond for her dedication and professionalism. Her expertise in flow cytometry was instrumental to my work on the mechanism of polyamide cytotoxicity. Her decades of experience and high standards for data quality helped me tremendously.



**Abstract.**

Py-Im polyamides are cell-permeable, programmable, sequence-specific, DNA minor groove-binding small molecules. When designed to bind a DNA sequence that matches the consensus DNA-binding sequence of a transcription factor, they can be used to block the binding of that transcription factor to its response element *in vitro* and in cell culture. We have used this approach to inhibit the genotoxic activity of the endogenous transcription factors HIF1 $\alpha$ , glucocorticoid receptor (GR), and androgen receptor (AR). In this work, we report the completion of a library of hairpin Py-Im polyamides targeted to all possible 5'-WGNNNW-3' (W = A or T) sequences. These compounds bind their target DNA sequences with high affinity. One compound from this set targets the sequence 5'-WGWWCW-3', which matches the DNA binding consensus sequence of GR and AR and has been shown to inhibit the gene regulatory activity of these proteins in cell culture. Herein, we show that a cyclic derivative of this compound maintains its activity against AR-driven gene expression in hormone-sensitive LNCaP prostate cancer cells. As androgen receptor signaling is crucial to prostate cancer growth and metastasis even in its recurrent form, we next examine the activity of the AR/GR antagonist in a tissue culture model of castration-resistant prostate cancer. In this model, the polyamide retains its activity against AR-driven mRNA expression, but it fails to inhibit the binding of AR to its response element. The polyamide-mediated repression is also accompanied by significant cell stress and cytotoxicity, which are explored in the final two chapters of this thesis. The former investigates a role for polyamides as inhibitors of DNA Topoisomerase II. Despite *in vitro* evidence indicating polyamides prevent Topoisomerase II binding, no evidence for this is found in cell culture. The final chapter reveals that polyamide-mediated cytotoxicity is likely due to inhibition of DNA synthesis. This occurs at concentrations similar to those used for transcription factor inhibition, suggesting that S-phase disturbance accompanies efforts to regulate gene expression with polyamides.

**Table of Contents**

List of Figures and Tables.....	vii
Chapter 1: Introduction.....	1
Chapter 2: Completion of a Programmable DNA-Binding Small Molecule Library.....	27
Chapter 3: Cyclic Pyrrole-Imidazole Polyamides Targeted to the Androgen Response Element.....	49
Chapter 4: Characterization of Py-Im Polyamide Androgen Receptor Antagonists in Hormone-Refractory Prostate Cancer Cells .....	72
Chapter 5: Py-Im Polyamides Inhibit DNA Topoisomerase II Activity <i>In Vitro</i> by Disrupting Enzyme Binding.....	92
Chapter 6: Mechanism of Polyamide-Induced Cytotoxicity in Prostate Cancer Cells .....	111

**List of figures and tables.**

## Chapter 1

Figure 1.1.	Chart depicting the genome size and number of protein-coding genes of several eukaryotes .....	2
Figure 1.2.	Signal transduction converges on transcription factors .....	3
Figure 1.3.	Structural diversity of DNA-binding proteins .....	4
Figure 1.4.	Composite model of cooperative assembly of transcription factors mediated by allosteric interactions on the Interferon- $\beta$ enhancer .....	6
Figure 1.5.	AR signaling in prostate cancer .....	8
Figure 1.6.	Two different strategies for the inhibition of the transcription factor HIF1 $\alpha$ .....	9
Figure 1.7.	Engineered zinc finger proteins for control of transcription.....	10
Figure 1.8.	Structure of B-form DNA .....	12
Figure 1.9.	Hydrogen-bonding pattern of the four Watson-Crick base pairs in the major and minor groove.....	13
Figure 1.10.	Sequence-specific, minor groove-binding natural products and their target sequences.....	14
Figure 1.11.	Schematic of an 8-ring hairpin polyamide designed to distinguish all four Watson-Crick base pairs.....	15
Figure 1.12.	Cell permeability and nuclear localization of polyamides in live MCF7 breast cancer cells.....	16
Figure 1.13.	Structural basis for allosteric inhibition of major groove-binding transcription factors by minor groove-binding polyamides....	17
Figure 1.14.	Py-Im polyamide inhibitors of HIF1 $\alpha$ and nuclear hormone receptor signaling.....	19

## Chapter 2

Figure 2.1.	Model for the complex formed between hairpin polyamide <b>24</b> and its match DNA sequence .....	31
Figure 2.2.	Plasmid design for pCFH2, pCFH3, pCFH4, pCFH5, pPh2, and pMFST .....	33
Figure 2.3.	Quantitative DNase I footprint titration experiments .....	34
Table 2.1.	$K_a$ ( $M^{-1}$ ) values reported are the mean values from at least three DNase I footprint titration experiments .....	35
Table 2.2.	Equilibrium association constants $K_a$ ( $M^{-1}$ ) .....	37
Table 2.3.	Equilibrium association constants $K_a$ ( $M^{-1}$ ) .....	39
Figure 2.4.	Quantitative DNase I footprint titration experiments .....	41

## Chapter 3

Figure 3.1.	Chemical structures for cyclic and hairpin polyamides .....	52
Scheme 3.1.	Preparation of <b>10</b> and <b>11</b> .....	53
Scheme 3.2.	Preparation of <b>1</b> , <b>2</b> , and <b>3</b> .....	54
Table 3.1.	$T_m$ values for polyamides <b>1-5</b> .....	55
Figure 3.2.	Targeting the ARE with DNA-binding polyamides .....	56

## Chapter 4

Figure 4.1.	Disrupting the AR/ARE interface in HRPC .....	75
Table 4.1.	$IC_{50}$ values for inhibition of PSA mRNA expression .....	77
Figure 4.2.	A Py-Im polyamide antagonist of AR-ARE binding inhibits expression of AR-target gene PSA .....	78
Figure 4.3.	Inhibition of AR occupancy at the FKBP5 intronic enhancer .....	79
Figure 4.4.	Inhibition of prostate cancer cell growth and induction of cytotoxic response following treatment with ARE-targeted polyamide <b>1</b> .....	80
Table 4.2.	Cytotoxicity $IC_{50}$ values in LNCaP and LN-AR cells in	

	response to treatment with <b>1</b> .....	81
Figure 4.5.	Caspase 3/7 activation accompanies PSA downregulation in unstimulated LN-AR cells .....	81
Figure 4.6.	Stabilization of p53 in response to polyamide treatment.....	82
Chapter 5		
Figure 5.1.	Chemical structure and binding preferences of the Py-Im polyamides used in this study .....	96
Figure 5.2.	<i>In vitro</i> DNA relaxation assay demonstrating polyamide-mediated, dose-dependent inhibition of Top2 $\alpha$ -p170 catalytic activity without cleavage complex formation.....	97
Figure 5.3.	Polyamides <b>1</b> and <b>2</b> inhibit Top2 $\alpha$ -p170 binding <i>in vitro</i> .....	98
Figure 5.4.	Dose-dependent induction of cytotoxicity by polyamides <b>1</b> and <b>2</b> in DU145 (wt) and Top2 knockdown cell lines .....	100
Table 5.1.	Cytotoxicity IC <sub>50</sub> values ( $\mu$ M) of compounds <b>1</b> and <b>2</b> in DU145 and DU145-shTop2 cell lines. ....	100
Figure 5.5.	Polyamide treatment causes S-phase arrest in DU145 and DU145-shTop2 $\alpha$ cells.....	101
Chapter 6		
Figure 6.1.	Chemical structure and DNA binding preferences of the Py-Im polyamides used in this study .....	113
Table 6.1.	Summary of cytotoxicity IC <sub>50</sub> values of 5'-WGWWCW-3' (W = A or T) polyamides in AR-expressing, AR-overexpressing, and AR-negative prostate cancer cell lines .....	116
Figure 6.2.	Polyamide treatment induces apoptosis in DU145 cells.....	117
Figure 6.3.	Polyamide treatment causes S-phase arrest .....	119
Figure 6.4.	Polyamide treatment does not induce DNA damage or activate the DNA-damage induced S-phase checkpoint .....	120

## **Chapter 1**

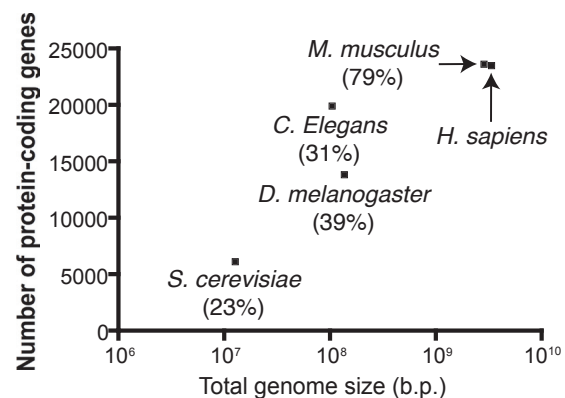
### **Introduction**

## 1.1 A central role for gene regulation

The sequencing of the human genome revealed a huge cache of information: three billion base pairs of genetic information encoded on 23 chromosomes. Despite the large relative size of the human genome, it quickly became apparent that it encodes an astonishingly small number of genes (1). Humans are now estimated to have about 23,000 protein-coding genes, a figure on the same order as the number estimated to belong to the roundworm (*Caenorhabditis elegans*). The homology of human genes to those from other animals is also high; nearly 40% of fruit fly (*Drosophila melanogaster*) genes have human homologs (figure 1.1) (2). When

the homology search is restricted to known human disease genes, this percentage rises to 75% (3). Yet in humans, these protein-coding regions make up less than 2% of the sequence data. The rest is made up of non-coding DNA corresponding to regulatory elements, functional RNAs, introns, and other structural or functional units. These sequences were once described as “junk DNA,” a form of genetic debris accumulated during the evolution of humans from lower

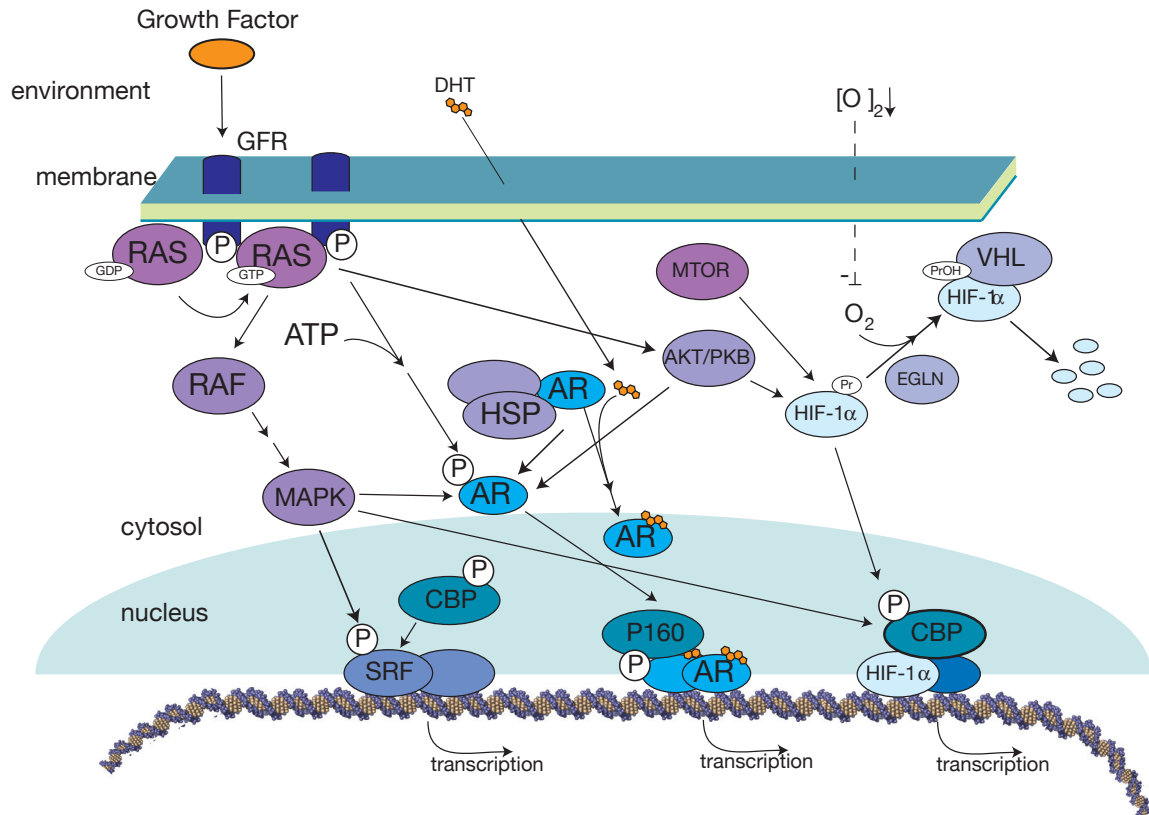
organisms and dismissed by Francis Crick as having “little specificity and convey[ing] little or no selective advantage to the organism” (4). Comparative genomics analyses have shown the opposite to be true: some segments of noncoding DNA are conserved through long evolutionary time periods, providing *prima facie* evidence of their function and selective advantage (5). Given that animals, especially mammals, share the same small set of genes, it would appear that complexity arises more from the regulation of those genes than from their sequence identity.



**Figure 1.1.** Chart depicting the genome size and number of protein coding genes of several eukaryotes. Numbers in parentheses indicate the percentage of that organism’s protein-coding genes that have human homologs. The genome size increases three orders of magnitude from yeast to man but the total number of genes increases only 5-fold and many of those genes are conserved.

## 1.2 Gene regulation by transcription factors

Since the vast array of diverse animal life shares so many common genes, it is perhaps unsurprising that gene regulation should be a complex endeavor. An animal's



**Figure 1.2:** Signal transduction converges on transcription factors. Posttranslational modifications and nuclear trafficking regulate diverse, interconnected signaling pathways. Growth factor signaling is mediated by the activation of receptor tyrosine kinases located on the cell membrane. Growth factor binding triggers a phosphorylation cascade, ultimately activating the SRF transcription factor to promote the transcription of cell cycle related genes. Androgen receptor (AR) is activated directly by binding a membrane-permeable ligand, dihydrotestosterone (DHT), which frees it from the heat shock proteins (HSPs) that keep it out of the nucleus in monomeric form. Liberation leads to dimerization, nuclear trafficking, and regulation of AR target gene expression. The hypoxia inducible factor, HIF1 $\alpha$ , is also modulated by a membrane permeable molecule, molecular oxygen. Under normoxic conditions, HIF1 $\alpha$  is hydroxylated, ubiquitinated, and rapidly degraded. Loss of O<sub>2</sub> rescues the transcription factor from degradation, and the receptor localizes to the nucleus to carry out its transcriptional program. Note that phosphorylation from the growth factor pathway cross-activates the transcriptional activation of the other two.

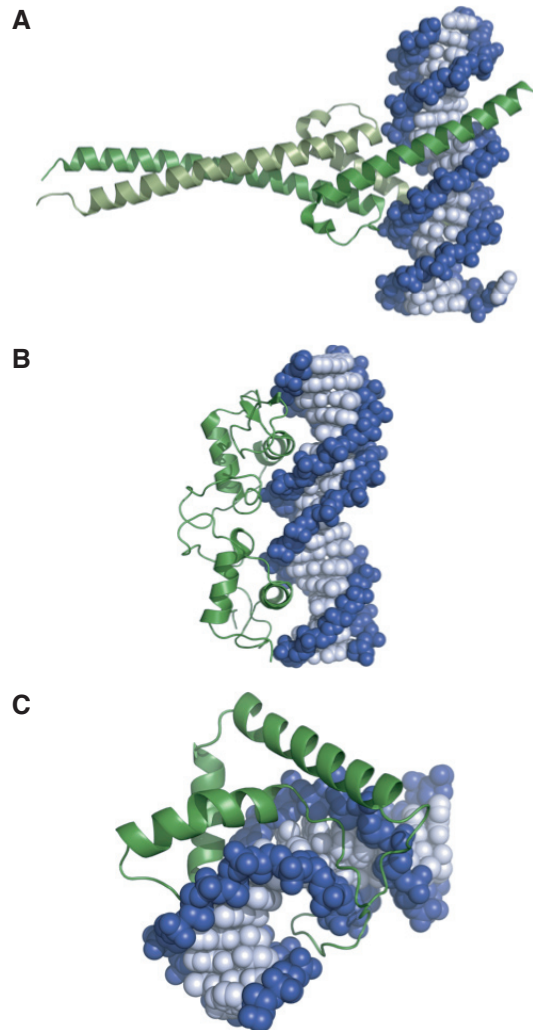
Adapted from N.G. Nickols (2008) "Endogenous gene regulation by DNA binding polyamides," p. 3. (Ph.D.) Dissertation, California Institute of Technology. Used with permission.

proper development and homeostasis depend critically on having the right gene product available at the right time. Cells can respond to a wide array of environmental and



autochthonous stimuli through changes in gene regulation at any level: from chromatin silencing to transcription initiation to posttranslational modification and trafficking. Many of the fastest regulatory responses are conducted through vast networks of proteins that form intertwined signaling cascades that rapidly transmit and amplify signals by using changes in phosphorylation or ubiquitination state, to name just two (figure 1.2). Often, these fast responses converge on proteins that integrate all of the upstream signaling to produce changes in gene expression. These proteins, called transcription factors, are ultimately responsible for the interaction of genes with the environment.

Transcription factors (TF) are defined as sequence-specific DNA-binding proteins. There are approximately 2600 in the human genome (6). Their general structure consists of a DNA-binding domain (DBD) and a transactivation or transrepressor domain (TAD) (7). Each DBD has a consensus DNA binding sequence: a short, loosely defined set of DNA sequence preferences. There are a variety of structural solutions for achieving sequence-specific binding; a few examples of these motifs include the basic helix-loop-helix (bHLH), the zinc finger, and High Mobility Group (HMG) box (figure 1.3) (8-10). Each motif produces a characteristic change in



**Figure 1.3.** Structural diversity of DNA-binding proteins. This figure depicts three different DNA-binding proteins that use different protein folding motifs to achieve sequence-specific recognition of their cognate DNA-binding elements. (A) bHLH proteins Myc and Max heterodimerize to bind in the major groove (PDB 1NKB). (B) A homodimer of truncated androgen receptor (AR) subunits bound to DNA in the major groove. Each subunit contains two zinc finger domains (PDB 1R4I). (C) Minor groove-binding Lef-1 contains an HMG box DNA-binding motif. Protein-binding causes a helical bend (PDB 2LEF).

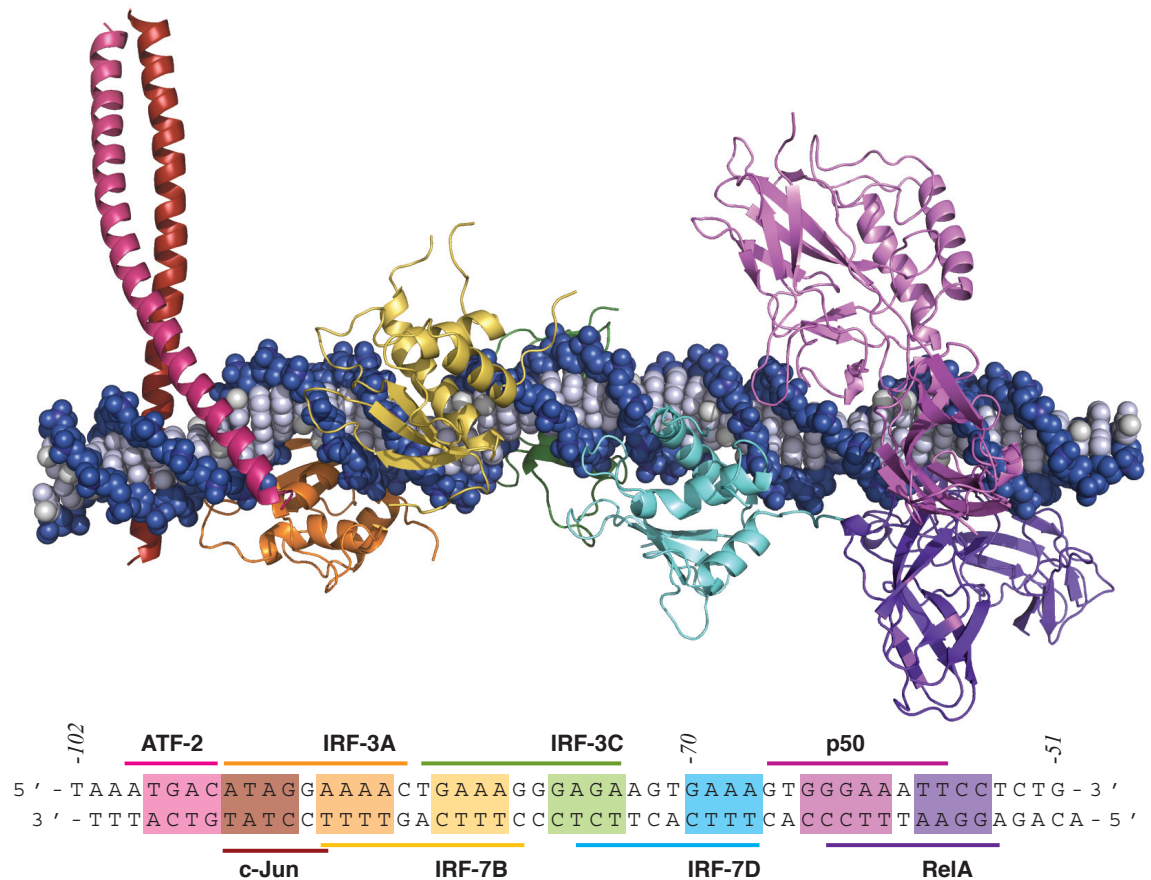
the local structure of its DNA host. They range from relatively small perturbations, like widening of the major groove that occurs upon binding of the zinc finger protein, androgen receptor (AR), to the near 90° helical bends produced by the HMG proteins Lef-1 and Sox2 (11, 12).

Genomic sites that match the consensus and allow TF binding are known as response elements and are found in promoter or enhancer regions within close geometric proximity to the transcription start site of a target gene. The TAD allows TFs to interact with other proteins to recruit coactivators or corepressors to the target locus (13). Coactivators promote histone acetylation and recruit mediator proteins, which form the scaffold for the assembly of the RNA Pol II holoenzyme (14). Corepressors also promote nucleosome remodeling, but instead act as histone deacetylases, which decreases the accessibility of the chromatin architecture to proteins.

The activation of transcription factors can occur through a variety of means. Those TFs that contain a third protein domain, called a signal-sensing domain, can directly respond to cell permeable stimuli through allosteric modulation of protein conformation (e.g., AR) (figure 1.2) (15). Other TFs are rescued from degradation (e.g., HIF) or phosphorylated (e.g., SRF) to increase their transactivation activity (16, 17). There is significant cross-talk between the upstream signaling pathways, such that a single stimulus can modulate the activity of a set of transcription factors, leading to a nuanced transcriptional response that may involve hundreds of genes. Activation of transcription factors allows transcriptional programs to be turned on only in response to the appropriate stimulus.

In addition, TFs often bind cooperatively, acting in concert with several other transcription factors to modulate the DNA topology. This combinatorial binding allows for plenty of diversity in gene expression; the 2600 transcription factors can combine to establish unique control of all 20,000 human genes (18). Combinatorial binding of transcription factors is assisted by the allosteric modulation of the DNA surface that occurs when these proteins bind DNA. The allosteric changes introduced by the binding

of a single transcription factor are transmitted along the DNA double helix and produce a favorable topography for its binding partners (Figure 1.4). In fact, the current model of transcription factor assembly on gene regulatory elements postulates that this allostery, not protein-protein interactions, is the dominant factor driving the highly cooperative binding



**Figure 1.4.** Composite model of cooperative assembly of transcription factors mediated by allosteric interactions on the Interferon- $\beta$  enhancer. DNA topological changes resulting from transcription factor binding enhance binding of other proteins to the same locus. Protein-protein contacts are not observed in either of the crystal structures used to make this composite.

Figure adapted by K.A. Muzikar from D. Panne et al., 2007 (19) using PDB 2O6G and 2O61 and used with her permission (K.A. Muzikar. (2011) "Repression of DNA-binding-dependent glucocorticoid receptor-mediated gene expression," p. 33. PhD Dissertation, California Institute of Technology.)

of multiple proteins to adjacent DNA sequences (19).

### 1.3 The androgen receptor in health and disease

Transcription factor activation can be a powerful phenotypic determinant. Consider the androgen receptor (AR): signaling through this single protein controls the development and maintenance of male sex characteristics (20). A member of the nuclear hormone receptor class of TFs, AR is composed of two N-terminal regulatory domains, a two-zinc finger DNA-binding domain, a hinge region, a ligand-binding domain, and a C-terminal interaction domain (21-25). In the absence of ligand, AR is sequestered in the cytosol by heat shock proteins (HSPs) (26). Steroid binding causes an allosteric conformational shift that liberates the TF from the HSPs, followed by homodimerization, nuclear localization, response element binding, and modulation of target gene transcription (Figure 1.2). In humans, these target genes control the development of primary and secondary sex characteristics, which means AR affects development of the genitalia as well as contributing to height, bone and muscle mass, and hair growth.

AR also plays a central role in the biology of the prostate and its neoplastic derivative, prostate cancer (27). In the normal prostate, AR signaling maintains the prostatic epithelium without producing uncontrolled growth. Castration results in rapid involution of the prostate, largely by apoptosis (28). When cancer emerges, growth inhibition is lost and AR signaling begins to drive the growth and spread of the disease. The molecular biology of prostate cancer oncogenesis is still an active field of research, and the set of AR-responsive genes responsible for the transformation has yet to be fully determined. In many tumors, a chromosomal translocation fusing the ERG oncogene with the AR-responsive TMPRSS2 gene appears to contribute to the uncontrolled growth of these cells (29). Like the normal prostate, prostate cancer will also respond dramatically to surgical or chemical castration, leading to clinical remission in some cases. The average time to relapse is approximately 6 months, and the current standard of care for recurrent disease relies on cytotoxic chemotherapy (30).

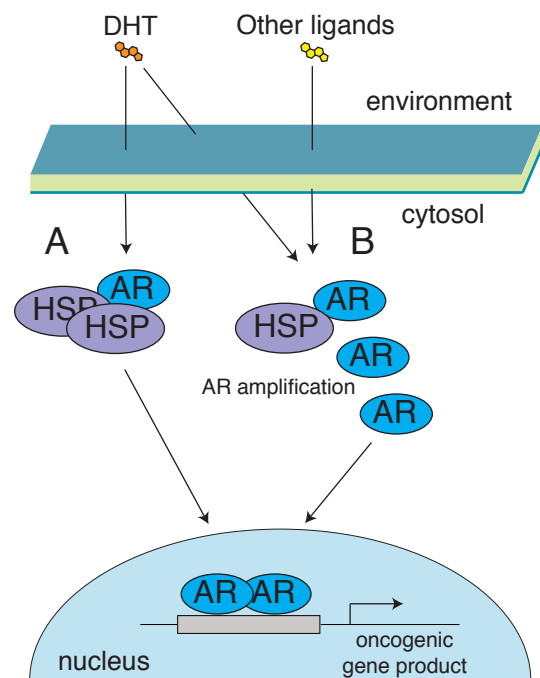
The recurrent form of prostate cancer is often called hormone-refractory, as first-

and second-generation antiandrogens like bicalutamide have failed. Yet this may be somewhat of a misnomer, as recent studies have shown that even in its recurrent form the tumor often relies on AR signaling for its growth: ‘castration-resistant’ may be more appropriate. In mouse models, a simple 5-fold upregulation of AR expression was sufficient to confer resistance to some anti-hormone agents (Figure 1.5) (31). In addition, the overexpression of AR in the human prostate cancer cell line, LNCaP, hypersensitized the pathway, rendering it promiscuous towards other steroid ligands (estrogen, dexamethasone) that would not activate the receptor under normal circumstances. Using these AR-expressing, castration-resistant cells as a model system, researchers have developed a third-generation anti-androgen

(MDV3100) that maintains its potency in recurrent prostate cancer (32). This drug has shown promise in phase I/II clinical trials, demonstrating once again the importance of transcription factors in human biology and medicine (33).

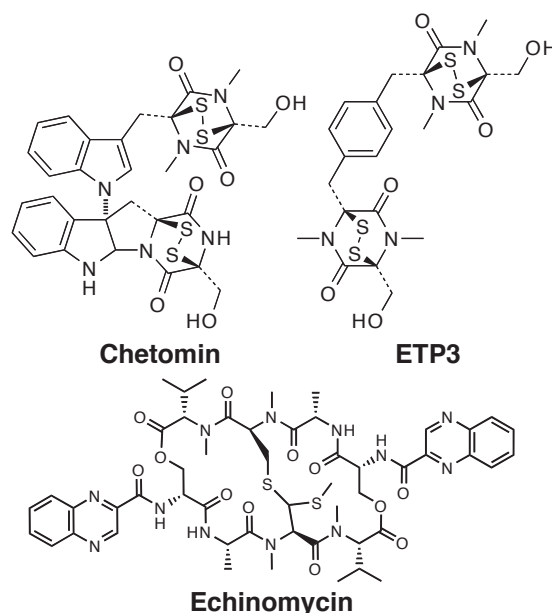
#### 1.4 Transcription factors as drug targets: beyond nuclear hormone receptors

In a widely cited review published in 2002, J. E. Darnell made a strong case for targeting transcription factors in cancer therapy (34). Instead of targeting the large number of oncogenic gene products implicated in pathogenesis, he advocated targeting



**Figure 1.5.** AR signaling in prostate cancer. (A) Prostate cancer arises when growth inhibition of AR-mediated mitogenic signaling is lost. Treatment-naïve disease responds to therapies directed at reducing androgen production (e.g., dihydrotestosterone, DHT). (B) Upregulation of AR expression is a common mechanism of castration resistance. Increased receptor expression sensitizes the AR pathway to low levels of residual androgens as well as other steroid ligands (e.g., glucocorticoids). Growth of castration-resistant tumors is still dependent on AR signaling, so high affinity antagonists (3rd generation anti-androgens) like MDV3100 retain their activity.

the relatively few dysregulated, overactive transcription factors common to many tumors. The transcription factors STAT3, NF $\kappa$ B, and  $\beta$ -catenin are noted as prominent examples. But unlike the nuclear hormone receptors, signaling cascades, not small molecules, activate these proteins. This makes them difficult to approach using traditional pharmacologic methods that target enzyme-substrate or receptor-ligand interactions. Darnell challenged chemists to respond with compounds that inhibit the two types of binding most important to transcription factor inhibition: protein-protein and protein-DNA interactions.



**Figure 1.6.** Two different strategies for the inhibition of the transcription factor HIF1 $\alpha$ . Chetomin and ETP3 inhibit the association of the transcription factor with its co-regulatory partner, p300. The intercalator echinomycin prevents HIF1 $\alpha$  binding to its DNA response element.

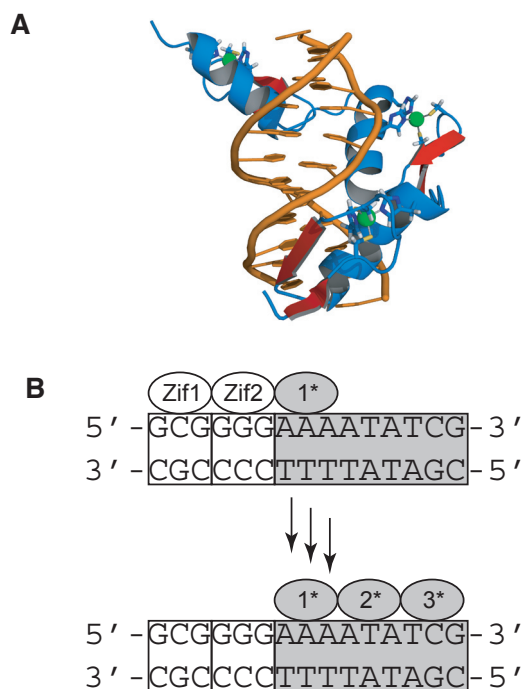
In the years since 2002, a number of small-molecule transcription factor inhibitors have been identified (35). As an example, the transcription factor HIF1 $\alpha$  has been successfully targeted *in vitro* and in mouse models. One of the more successful HIF1 $\alpha$  protein-protein interaction inhibitors is ETP3, a dimeric compound based on the natural product chetomin that disrupts the protein-protein interaction of HIF1 $\alpha$  with its binding partner, p300, and has submicromolar IC<sub>50</sub> against HIF1 $\alpha$  promoter activity (Figure 1.6) (36, 37). Another natural product, echinomycin, was identified in a high-throughput screen and shown to inhibit the protein-DNA interactions of the transcription factor at nanomolar concentrations (38). Yet each of these solutions is idiosyncratic: the first relied on intimate structural knowledge of the transcription factor, and the second required well-established knowledge of the protein's function as a prerequisite for designing the screen. Nearly all of the small molecule transcription factor antagonists developed over the last few years



have the same story; each solution is unique, and it is difficult to apply the lessons learned from one system to the next target. A general chemical solution to this problem would be invaluable to human medicine.

Most of the general solutions to transcription factor inhibition involve the use of large, biomolecular constructs. There are both nucleotide- and protein-based approaches. RNAi has been used extensively as a research tool for transcription factor inhibition by protein depletion (39). The specificity of this technique is exquisite, and *in vitro*, very effective. But to reach the clinic, RNAi approaches must solve a difficult delivery problem: how to achieve systemic distribution of large, polyanionic molecules in complex animals. The solution has remained elusive for years. Other nucleotide-based approaches, like antisense and decoy oligonucleotides, have similar delivery issues.

The first peptidic approach to transcription factor inhibition employed engineered zinc finger proteins to compete with an endogenous transcription factor for its binding site (40). A zinc finger motif consists of an  $\alpha$ -helix and an antiparallel  $\beta$ -sheet that together coordinate a single zinc ion and specify a 3-base pair DNA sequence (Figure 1.7) (41). Arrays of multiple zinc fingers can be constructed in a modular fashion or selected using directed evolution to target all



**Figure 1.7.** Engineered zinc finger proteins for control of transcription. (A) Crystal structure of zinc finger protein Zif268 containing three zinc finger motifs. Coordinated zinc ions are shown as green spheres. Image created by T. Splettstoesser from PDB 1A1L (41) and used with permission under Creative Commons license. (B) Schematic diagram depicting the iterative selection of zinc finger motifs to bind a 9-base pair DNA sequence. Two zinc finger motifs (Zif1 & Zif2) of known sequence selectivity are used to position a third (1\*) whose binding to the target 5'-AAA-3' sequence is optimized by affinity-based selection. The optimized 1\* motif is then paired with Zif2 and a new protein sequence (2\*) is optimized to bind 5'-ATA-3'. A third iteration with optimized 1\* & 2\* is used to generate the full length protein to recognize the 5'-AAAATATCG-3' target sequence.

possible DNA sequences up to 18 base pairs in length (42, 43). In a similar fashion to RNAi, this technology can achieve high specificity, as some 18-base pair sequences are unique in the human genome (44). Unfortunately, like RNAi, engineered zinc fingers also have significant delivery problems and rely on lentiviral gene delivery vectors for use in animals (45).

More recently, a peptide-based approach has been used to target protein-protein interactions of transcription factors. This strategy employs stabilized, ‘stapled’  $\alpha$ -helical oligopeptide fragments of dominant negative transcription factor binding partners (46). These peptides are apparently small enough to enter cells through endocytic mechanisms and achieve nuclear localization without the use of delivery agents. As a recent example, a stapled helix targeted to the NOTCH transcription factor disrupted the pathway *in vitro* and in mouse models following systemic administration (47). This technology has promise, but it is not as mature as RNAi. Little biodistribution data are available.

### **1.5 DNA-binding Py-Im polyamides for the control of transcription**

We have chosen to pursue the development of a general chemical methodology for the inhibition of transcription factor signaling. We have targeted the DNA side of the protein-DNA interface, in part to take advantage of the relative simplicity of the structure of the DNA double helix. For this purpose, we employ sequence-specific DNA-binding Py-Im polyamides as part of a general strategy for developing transcription factor inhibitors. Candidate compounds must meet several stringent criteria. In order to be general, the compounds must be able to bind a wide array of DNA sequences in a modular fashion with high affinity and specificity. In order to be effective in biological systems, they must be cell permeable, traffic to the nucleus, and bind genomic DNA in its native chromatin conformation. Having arrived at their target site, they must inhibit the transcription factor, either through direct competition for the binding site surface, or by allosteric modulation of the local DNA topography. Finally, to be useful as therapeutics, candidate compounds



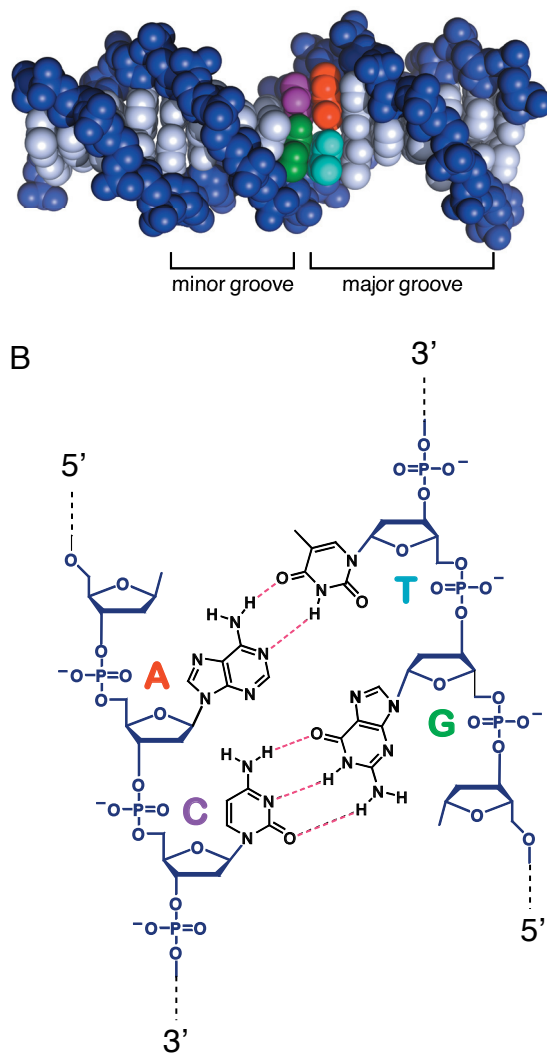
would ideally obtain systemic distribution in animals without the need for delivery agents. The following section details how the Py-Im polyamide has been engineered to meet all of these requirements.

## 1.6 Molecular recognition of the DNA A

### minor groove

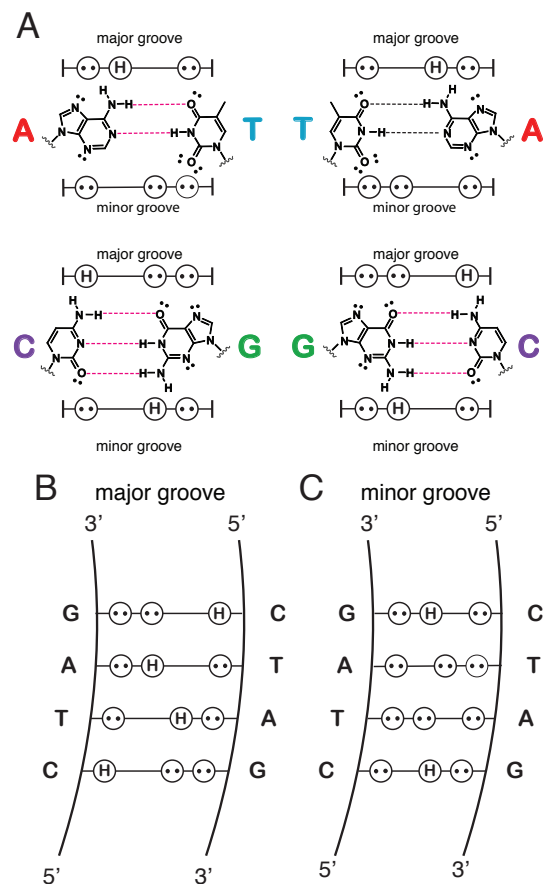
B-form DNA is an asymmetric double helix composed of two antiparallel strands of nucleotides (48). The asymmetry causes the helix to form a wide and shallow major groove as well as a narrow and deep minor groove (Figure 1.8). Small molecules that bind DNA can do so in either groove; they can also intercalate between the base pairs or bind the phosphate backbone. Sequence-specific small molecules interact with the particular hydrogen-bonding and steric patterns that accompany the DNA base pairs (49). Each base pair presents a unique surface of hydrogen bond donors and acceptors that can be recognized by a complimentary surface on a small molecule (Figure 1.9). Examples of some sequence-specific DNA-binding molecules are shown in Figure 1.10.

The natural product distamycin A is the original scaffold from which Py-Im polyamides developed (50). Distamycin A



**Figure 1.8.** Structure of B-form DNA. (A) Crystal structure of B-form DNA double helix showing the location of the narrow minor groove and wide major groove. The phosphate backbone and ribose residues are colored dark blue. The colored nucleobases represent adenine (magenta), thymine (teal), guanine (green), and cytosine (purple). (B) Chemical structure of the Watson-Crick base pairs showing the hydrogen bonding pattern. Figure adapted from K.A. Muzikar (2011).

binds to the DNA minor groove with 1:1 or 2:1 stoichiometry. The compound consists of three *N*-methylpyrrole (Py) rings linked by amide bonds that hydrogen bond with the N3 of purines or the O2 of pyrimidines. The 2:1 antiparallel binding mode prefers A-T tracts due to steric constraints conferred by the exocyclic amine of guanine in the minor groove (Figure 1.10) (51, 52). Substituting *N*-methylimidazole (Im) for one of the *N*-methylpyrroles relieves this steric hindrance and produces a hydrogen bond acceptor for the exocyclic amine of guanine. In the 2:1 antiparallel configuration, this compound pairs Im across from Py to specify 5'-WGWCW-3'; the parent compound binds 5'-WWWWW-3' (W = A or T) (53, 54). This simple substitution marked the beginning of systematic DNA base pair recognition in the minor groove by small molecules.



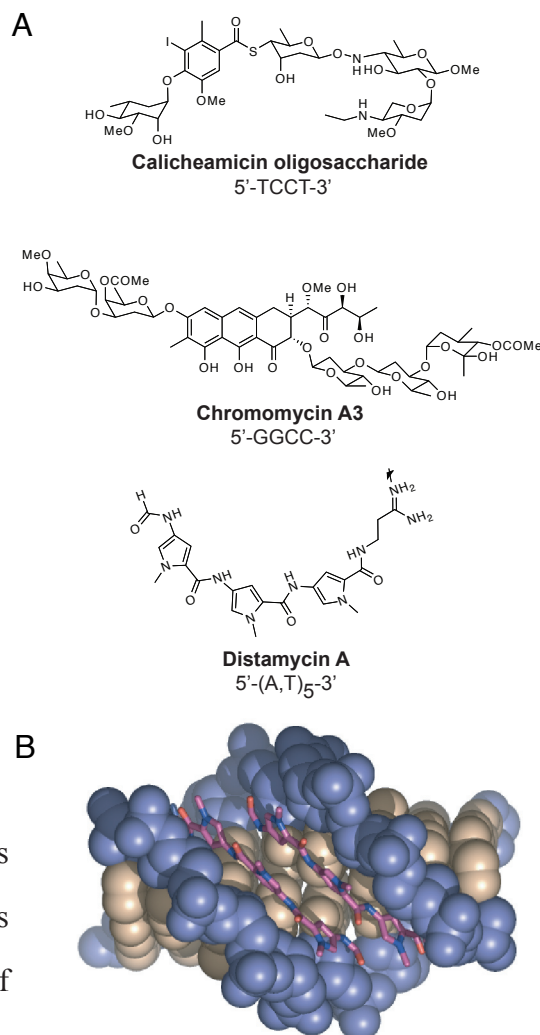
**Figure 1.9.** Hydrogen-bonding pattern of the four Watson-Crick base pairs in the major and minor groove. (A) The hydrogen-bonding patterns of each of the base pairs along the edges of the nucleobase is unique. Formal representation of the hydrogen bonding surface available for sequence-specific interactions with DNA-binding molecules in the major (B) and minor (C) grooves. Figure adapted from K.A. Muzikar (2011).

Py-Im polyamides, although originally based on the distamycin structure, have been extensively optimized over many years for DNA binding affinity and sequence specificity. The Im/Py pairing has been shown to be a general solution for specifying for G-C base-pair, and the reverse, Py/Im, specifies for C-G (55). A Py/Py pair is degenerate for A or T (W). Each ring pair specifies for a DNA base pair; polyamides specifying DNA sequences of 16 base pairs in length have been characterized *in vitro* (56). The composition of the individual amino acid units has also been extended beyond *N*-methylpyrrole and *N*-

methylimidazole to include other heterocyclic units (57). The most successful of these was the development of the hydroxypyrrole (Hp) monomer, such that Hp/Py specifies T-A and Py/Hp specifies A-T (Figure 1.11) (58). Thus polyamides can distinguish all four Watson-Crick base pairs through interactions with the DNA minor groove.

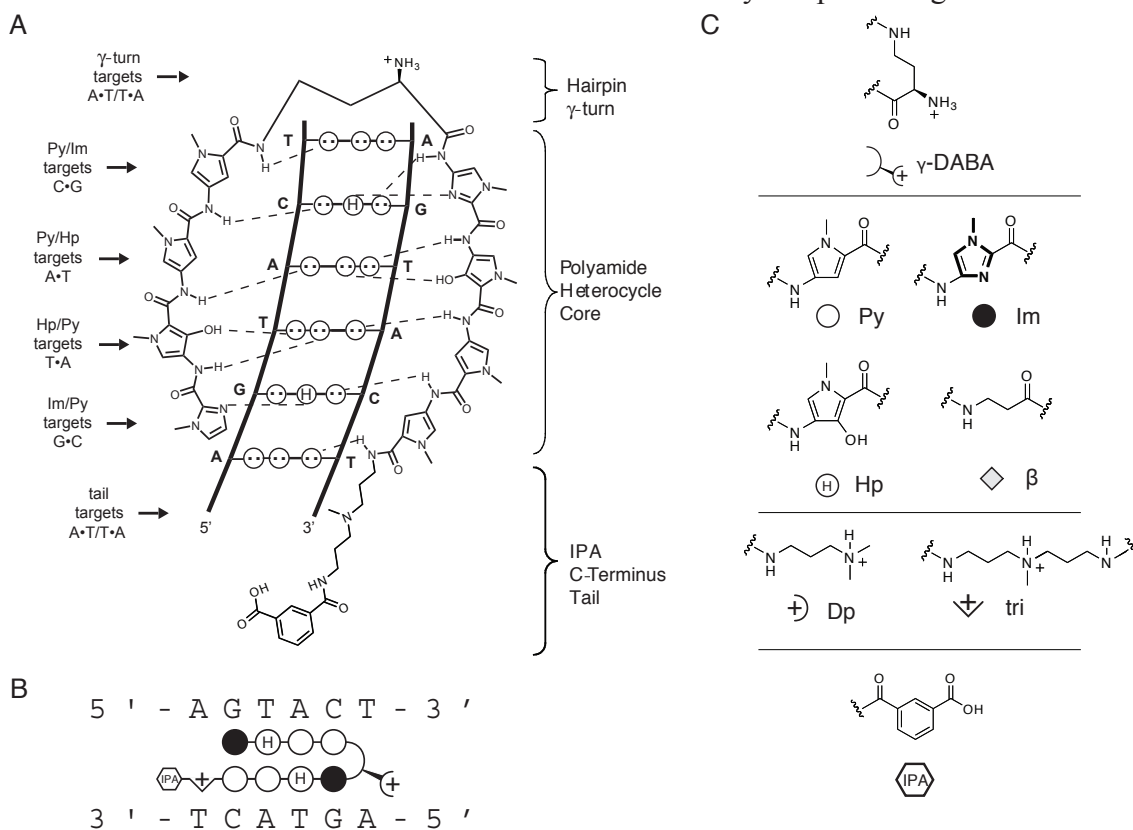
The configuration of the heterocycle ring pairs has also been optimized by the Dervan group. Development of a covalent linkage between the two antiparallel strands of pyrroles and imidazoles led to a 100 to 3600-fold increase in binding affinity as well as the ability to target non-palindromic sequences (59). This led to the ‘hairpin’ configuration of antiparallel strands linked at one end that has been most well studied in biological systems. The addition of a chiral amino group on this hairpin turn was found to reinforce binding orientation, further increasing their specificity and affinity (60). The combination of the hairpin configuration, chiral turn, and Py/Im pairing rules have produced polyamides with nanomolar affinities to rival those of endogenous transcription factors and 10 to 100 fold specificity for a single base pair mismatch.

The DNA-binding capabilities of Py-Im polyamides are not restricted to the canonical B-form helix. They have been engineered to bind alternative DNA structures, including



**Figure 1.10.** Sequence-specific, minor-groove binding natural products and their target sequences. (A) Calicheamicin binds as a monomer, and chromomycin binds as a dimer. Distamycin A can bind in a 1:1 and a 2:1 conformation. (B) Crystal structure of Distamycin A bound to the DNA sequence 5'-GTATATAC-3' in a 2:1 conformation. (PDB 378D)

the interwoven strands of the DNA double-crossover array for patterning nanostructures



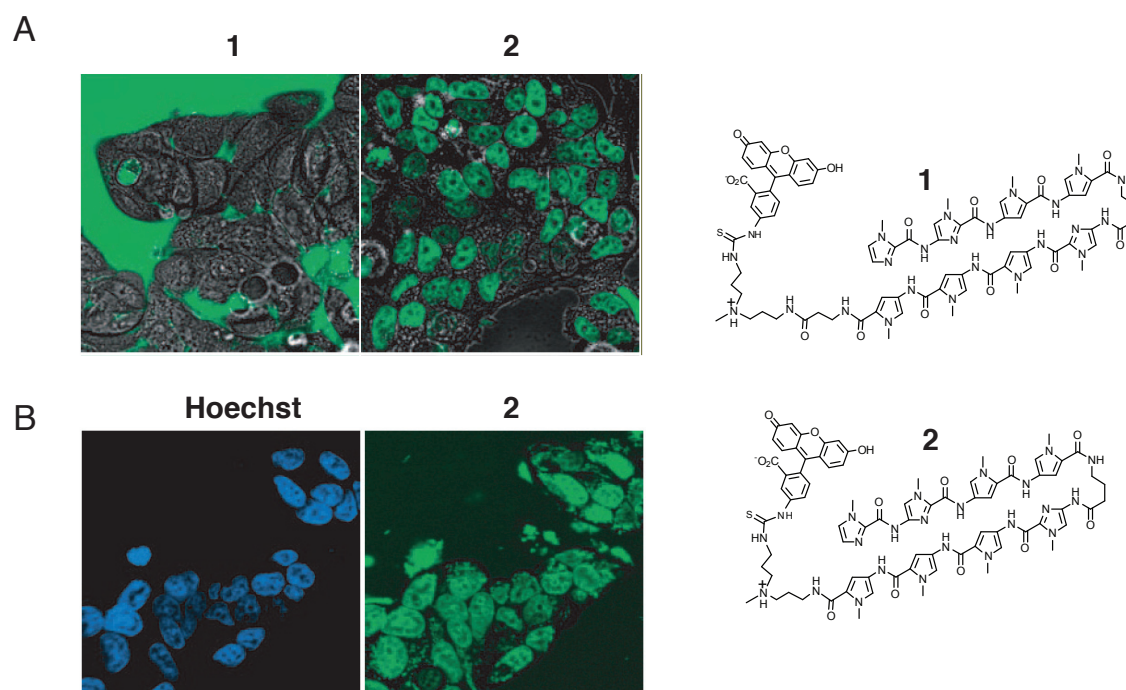
**Figure 1.11.** Schematic of an 8-ring hairpin polyamide designed to distinguish all four Watson-Crick base pairs. (A) Interaction of the polyamide with the hydrogen-bonding pattern in the minor groove of a 5'-AGTACT-3' DNA sequence. Each base pair is interpreted by a pair of heterocyclic rings positioned across from each other in antiparallel fashion by the  $\gamma$ -turn. The turn and tail have a strong preference for A•T or T•A, so the formal sequence targeted by the polyamide is 5'-WGTACW-3' (W = A or T). (B) Ball-and-stick representation of the polyamide. Filled circles represent imidazoles, open circles represent pyrroles, and 'H' represents hydroxypyrrrole. (C) Key to ball-and-stick representation of hairpin polyamides.

(61). Biologically relevant DNA conformations have also been targeted in a sequence-specific fashion, including DNA in its histone-bound form as part of the nucleosome core particle (62).

### 1.7 Cell permeability

In order to act as transcriptional regulators in biological systems, polyamides must be cell permeable, ideally without using any delivery agents. Cell permeability of Py-Im polyamides has been directly observed in live cells by using fluorophore conjugates of these

compounds. Over one hundred polyamide-fluorophore conjugates have been studied in over a dozen cell lines, revealing them to be broadly, though not universally cell permeable (63, 64). No delivery agents were used, and the cells were not fixed prior to examination by confocal microscopy (Figure 1.12). Pyrrole-imidazole sequence content, dye choice and position, dye linker composition, and overall molecular weight all contribute to the biocompatibility of these compounds. In addition, some cell lines are more promiscuous toward polyamide uptake than others. Like the stapled  $\alpha$ -helices discussed above, polyamide uptake appears to be energy dependent and likely to be endocytotic, but this has not been established as a general property of the compounds as a class. We infer cell permeability

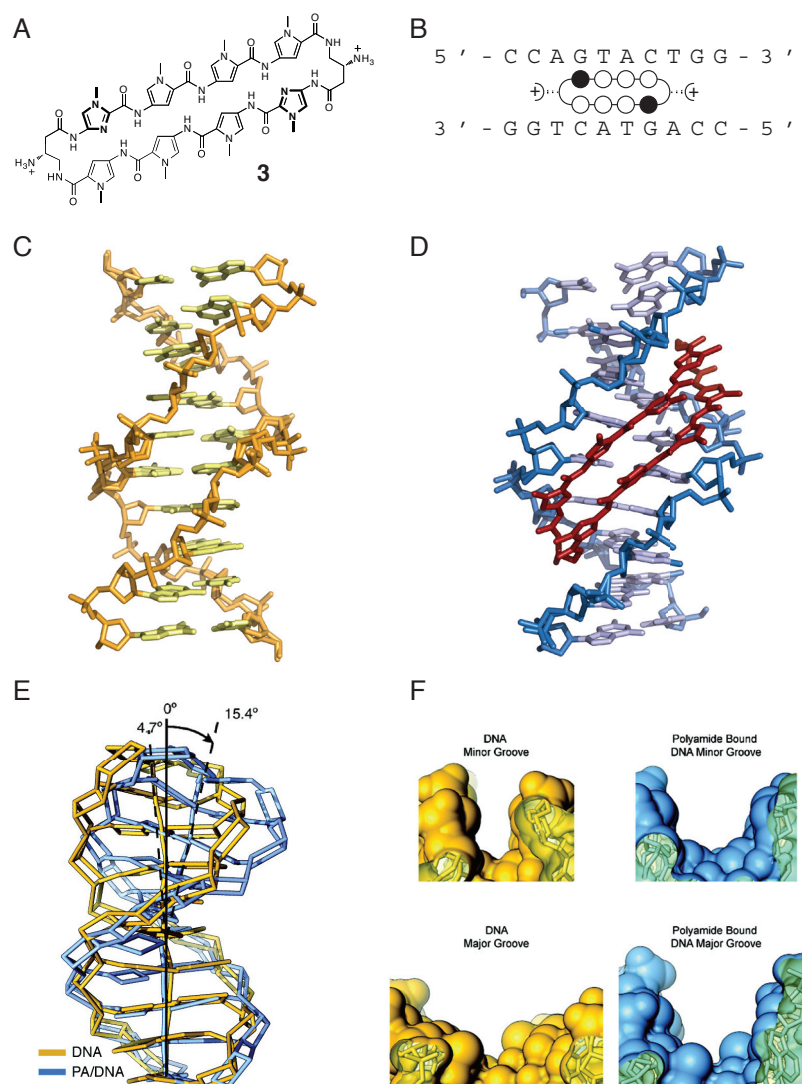


**Figure 1.12.** Cell permeability and nuclear localization of polyamides in live MCF7 breast cancer cells. (A) Laser confocal photomicrograph of fluorescein-conjugated polyamides **1** and **2** after 10-14 h. treatment. Compound **1** is excluded from the cells; compound **2** is cell permeable. (B) Laser confocal photomicrograph of compound **2**, showing colocalization with Hoechst, a known nuclear staining agent.

of the unlabeled parent compounds based on the permeability of their fluorophore-labeled derivatives as well as their observed effects on transcription (see below).

### 1.3 Gene regulation by polyamides in living systems

*In vitro* assays first established that polyamides could act as inhibitors of transcription factor binding. They have been shown to inhibit the DNA binding of both major and minor groove-binding transcription factors. The inhibition of major groove-binding transcription factors is presumably due to an allosteric mechanism, as polyamides bind only in the minor groove. A series of recent crystal structures of polyamide-bound DNA and the unliganded oligonucleotide at atomic resolution has elucidated the mechanism of allosteric inhibition



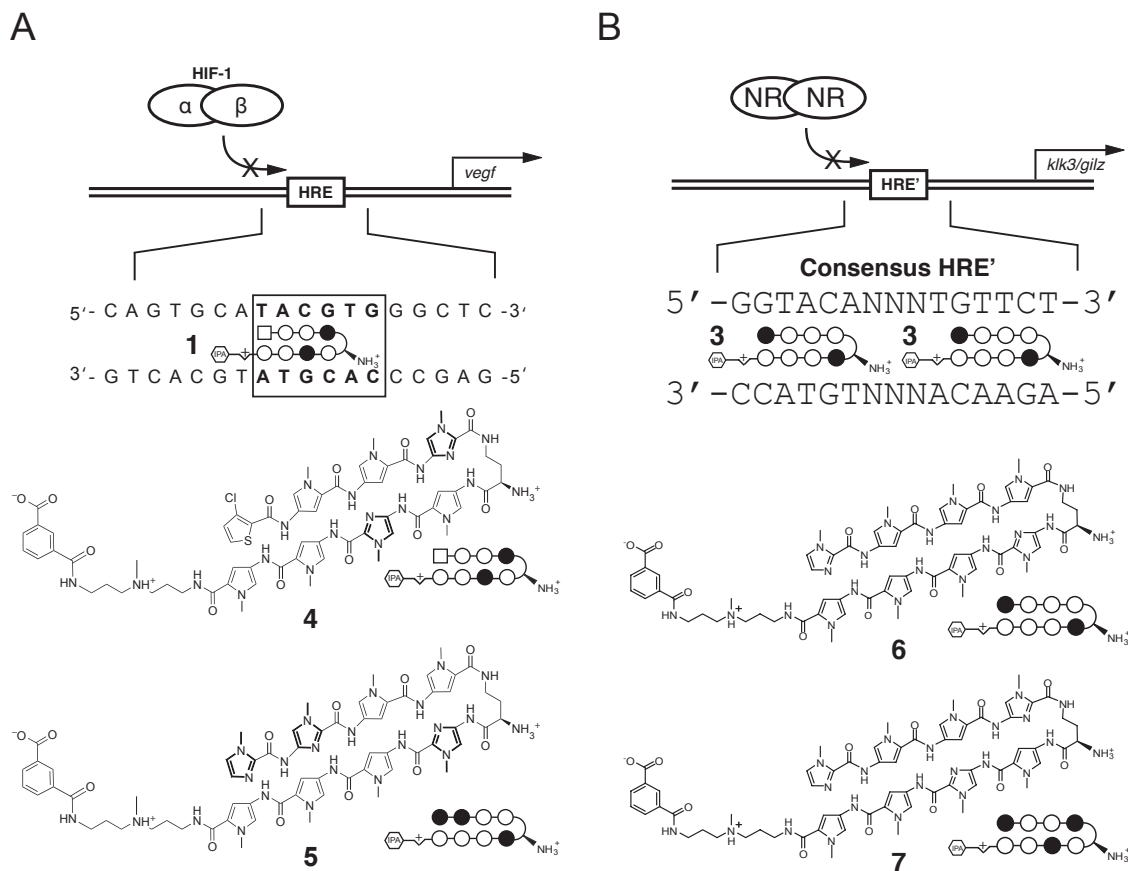
**Figure 1.13.** Structural basis for allosteric inhibition of major groove-binding transcription factors by minor groove-binding polyamides. (A) Chemical structure of ARE-targeted cyclic polyamide compound **3**. (B) Ball-and-stick representation of **3** bound to its target DNA sequence. (C) Atomic resolution crystal structure of the DNA duplex shown in (B) (PDB 1D8G). (D) Atomic resolution crystal structure of **3** bound to the DNA sequence shown in (B) (PDB 3OMJ). (E) Superimposition of structures (C) and (D) showing the 15° helical bend. (F) ~4 Å narrowing of the DNA minor groove and ~4 Å widening of the minor groove that occurs upon polyamide binding.



in great detail: polyamide binding widens the minor groove by 4 Å and narrows the major groove by 4 Å (Figure 1.13) (65,66). In addition, it introduces a 15° helical bend toward the major groove, further distorting it. Allosteric inhibition of major-groove binding transcription factors by polyamides has a sound structural and biochemical foundation.

Proof of principle for polyamides as transcriptional regulators was first established with a polyamide targeted to the DNA response elements of the major groove-binding transcription factor HIF1 $\alpha$  (67). Signaling through HIF1 $\alpha$  regulates the homeostatic response to low oxygen tension; this process is co-opted by cancer cells to drive neovascularization (Figure 1.2). The DNA binding consensus sequence for HIF1 $\alpha$  is 5'-TACGTG-3'; DNA elements that match this consensus and accommodate HIF1 $\alpha$  binding are known as Hypoxic Response Elements (HREs). We designed a polyamide (**X**) that binds the sequence 5'-WTWCGW-3' with nanomolar affinity (Figure 1.14). In HeLa and U251 cells, this compound inhibited the induction of some HIF1 $\alpha$  target genes when the cells were treated with desferioxamine (DFO), a small molecule mimic of hypoxia. Among these genes was VEGF, the vascular endothelial growth factor that promotes blood vessel formation required for extensive tumor growth (68). Microarray analysis of gene expression showed that **4** blocked a subset of all DFO-induced changes in gene expression. Chromatin immunoprecipitation experiments confirmed decreased HIF1 $\alpha$  occupancy at the VEGF HRE when cells were treated with **4** prior to DFO induction. Targeting the HRE for polyamide-mediated transcriptional modulation of an endogenous transcription factor in cell culture resulted in the inhibition of a medically important gene, VEGF.

We have also developed an 8-ring hairpin Py-Im polyamide (**6**) that antagonizes steroid-induced gene expression changes driven by the nuclear hormone receptors androgen receptor (AR) and glucocorticoid receptor (GR) (Figure 1.14) (69,70). In animals, these transcription factors form powerful nodes in the transduction of signals from circulating hormones: AR controls the development and maintenance of the male sexual phenotype in response to testosterone as described above, and GR regulates the expression of anti-



**Figure 1.14.** Py-Im polyamide inhibitors of HIF1 $\alpha$  and nuclear hormone receptor signaling. (A) Cartoon representation of inhibition of HIF1 $\alpha$  binding by compound **4**, which is designed to bind the hypoxia response element (HRE) in the VEGF promoter. Compound **4** inhibited the induction of VEGF mRNA expression and the expression of other HIF1 $\alpha$  target genes in response to desferrioxamine, a chemomimetic of hypoxic conditions. Mismatch compound **5** is targeted to a sequence unrelated to the HRE consensus and was much less effective. (B) Cartoon representation of nuclear receptor transcription factors binding to a hormone response element (HRE') as a homodimer in response to ligand binding. Compound **6** is designed to target the consensus DNA binding sequence shared by androgen receptor (AR) and glucocorticoid receptor (GR). Compound **6** inhibited the androgen-induced expression of some AR target genes, like KLK3 and the glucocorticoid-induced expression of some GR target genes (e.g. GILZ). Chromatin immunoprecipitation confirmed inhibition of transcription factor binding at the appropriate response elements for all three transcription factors.

inflammatory genes in response to cortisol (27,71). In the absence of ligand, these receptors are sequestered in the cytosol by heat shock proteins (HSPs). Steroid binding causes an allosteric conformational shift followed by homodimerization, nuclear localization, response element binding, and modulation of target gene transcription (72). AR and GR share the consensus binding sequence 5'-GGTACANNNTGTTCT-3'. Their genotropic actions can be partially inhibited in cancer tissue culture cells by polyamide **6**, which disrupts the protein-DNA interface by selectively binding the DNA sequence 5'-WGWWCW-3' (W=



A or T) (69,70). Chromatin immunoprecipitation experiments again confirmed decreased response element occupancy by the appropriate transcription factor when treated with **6**. This second example shows the generality of the polyamide approach; nuclear hormone receptors are structurally unrelated to HIF1 $\alpha$ , yet both can be successfully inhibited with the same class of compound.

The polyamide HIF1 $\alpha$  and AR/GR inhibitors are composed of 8 Py/Im ring pairs and bind 6 DNA base pairs. Larger hairpin compounds containing beta-alanine residues to increase flexibility and binding site size have been employed to inhibit the binding of a third class of transcription factor, AP-1. This compound inhibited AP-1 directed MMP9 expression in cell culture and mouse models of metastatic colorectal cancer with Py-Im polyamides (73). Our Japanese colleagues have begun preclinical work on these compounds in rats and achieved therapeutic dose levels following systemic administration without the use of delivery agents (74). These results are important first steps toward polyamide-based, transcription factor-targeted therapeutics.

## **1.8 Scope of this work**

This thesis builds on all of the optimization and proof-of-principle experiments in an effort to define the scope and the limitations of DNA-binding Py-Im polyamides as inhibitors of transcription factors. Chapter 2 describes an effort to define the polyamide tool kit: we design, synthesize, and measure the DNA-binding affinities and specificities of the set of polyamides targeted to all possible 5'-WGNNNW-3' (W = A or T, N = any nucleotide) sequences. Chapter 3 details the synthesis and biological activity of a cyclic polyamide designed to inhibit AR binding. This compound maintains its activity against AR-driven gene expression in hormone-sensitive prostate cancer cells and represents an alternative polyamide architecture for use in future animal experiments. In chapter 4, I present an attempt to extend the AR work in a hormone-resistant cell line. This attempt ultimately fails. We observe inhibition of AR-target gene mRNA expression, but this change

is not accompanied by decreased AR occupancy at those genes when cells are treated with polyamide. In addition, polyamide treatment is accompanied by significant cytotoxicity and activation of a stress response. These events are explored further in chapters 5 and 6. The former investigates a possible role for polyamides as DNA Topoisomerase II inhibitors, and the latter establishes **6** as an inhibitor of DNA synthesis in cell culture. I discuss the ramifications of these findings on the design and use of polyamides as transcription factor antagonists.

## References.

1. (2004) Finishing the euchromatic sequence of the human genome. *Nature*, **431**, 931-945.
2. Gilbert, D.G. (2002) euGenes: a eukaryote genome information system. *Nucleic Acids Res*, **30**, 145-148.
3. Reiter, L.T., Potocki, L., Chien, S., Gribskov, M. and Bier, E. (2001) A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res*, **11**, 1114-1125.
4. Orgel, L.E. and Crick, F.H. (1980) Selfish DNA: the ultimate parasite. *Nature*, **284**, 604-607.
5. Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E. *et al.* (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, **447**, 799-816.
6. Babu, M.M., Luscombe, N.M., Aravind, L., Gerstein, M. and Teichmann, S.A. (2004) Structure and evolution of transcriptional regulatory networks. *Curr Opin Struct Biol*, **14**, 283-291.
7. Latchman, D.S. (1997) Transcription factors: an overview. *Int J Biochem Cell Biol*, **29**, 1305-1312.
8. Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B., *et al.* (1989) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. *Cell*, **58**, 537-544.
9. Krishna, S.S., Majumdar, I. and Grishin, N.V. (2003) Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res*, **31**, 532-550.
10. Stros, M., Launholt, D., Grasser, K.D. (2007) The HMG-box: a versatile protein domain occurring in a wide variety of DNA-binding proteins. *Cell Mol Life Sci*, **64**, 19-20.
11. Love, J.J., Li, X., Case, D.A., Giese, K., Grosschedl, R. and Wright, P.E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature*, **376**, 791-795.
12. Remenyi, A., Lins, K., Nissen, L.J., Reinbold, R., Scholer, H.R. and Wilmanns, M. (2003) Crystal structure of a POU/HMG/DNA ternary complex suggests differential assembly of Oct4 and Sox2 on two enhancers. *Genes Dev*, **17**, 2048-2059.
13. Ptashne, M. and Gann, A. (1997) Transcriptional activation by recruitment. *Nature*, **386**, 569-577.
14. Roeder, R.G. (2005) Transcriptional regulation and the role of diverse coactivators in animal cells. *FEBS Lett*, **579**, 909-915.
15. Whiteside, S.T. and Goodbourn, S. (1993) Signal transduction and nuclear targeting: regulation of transcription factor activity by subcellular localisation. *J Cell Sci*, **104** ( Pt 4), 949-955.
16. Conaway, R.C., Brower, C.S. and Conaway, J.W. (2002) Emerging roles of ubiquitin in transcription regulation. *Science*, **296**, 1254-1258.
17. Bohmann, D. (1990) Transcription factor phosphorylation: a link between signal transduction and the regulation of gene expression. *Cancer Cells*, **2**, 337-344.

18. Brivanlou, A.H. and Darnell, J.E., Jr. (2002) Signal transduction and the control of gene expression. *Science*, **295**, 813-818.
19. Panne, D., Maniatis, T. and Harrison, S.C. (2007) An atomic model of the interferon-beta enhanceosome. *Cell*, **129**, 1111-1123.
20. Davison, S.L. and Bell, R. (2006) Androgen physiology. *Semin Reprod Med*, **24**, 71-77.
21. Jenster, G., van der Korput, H.A., Trapman, J. and Brinkmann, A.O. (1995) Identification of two transcription activation units in the N-terminal domain of the human androgen receptor. *J Biol Chem*, **270**, 7341-7346.
22. Shaffer, P.L., Jivan, A., Dollins, D.E., Claessens, F. and Gewirth, D.T. (2004) Structural basis of androgen receptor binding to selective androgen response elements. *Proc Natl Acad Sci U S A*, **101**, 4758-4763.
23. Kaku, N., Matsuda, K., Tsujimura, A. and Kawata, M. (2008) Characterization of nuclear import of the domain-specific androgen receptor in association with the importin alpha/beta and Ran-guanosine 5'-triphosphate systems. *Endocrinology*, **149**, 3960-3969.
24. He, B., Gampe, R.T., Jr., Kole, A.J., Hnat, A.T., Stanley, T.B., An, G., Stewart, E.L., Kalman, R.I., Minges, J.T. and Wilson, E.M. (2004) Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance. *Mol Cell*, **16**, 425-438.
25. Lanz, R.B. and Rusconi, S. (1994) A conserved carboxy-terminal subdomain is important for ligand interpretation and transactivation by nuclear receptors. *Endocrinology*, **135**, 2183-2195.
26. Veldscholte, J., Berrevoets, C.A., Brinkmann, A.O., Grootegoed, J.A. and Mulder, E. (1992) Anti-androgens and the mutated androgen receptor of LNCaP cells: differential effects on binding affinity, heat-shock protein interaction, and transcription activation. *Biochemistry*, **31**, 2393-2399.
27. Lamont, K.R. and Tindall, D.J. (2010) Androgen regulation of gene expression. *Adv Cancer Res*, **107**, 137-162.
28. Kyprianou, N. and Isaacs, J.T. (1988) Activation of programmed cell death in the rat ventral prostate after castration. *Endocrinology*, **122**, 552-562.
29. Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R. *et al.* (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*, **310**, 644-648.
30. Bracarda, S., Logothetis, C., Sternberg, C.N. and Oudard, S. (2011) Current and emerging treatment modalities for metastatic castration-resistant prostate cancer. *BJU Int*, **107 Suppl 2**, 13-20.
31. Chen, C.D., Welsbie, D.S., Tran, C., Baek, S.H., Chen, R., Vessella, R., Rosenfeld, M.G. and Sawyers, C.L. (2004) Molecular determinants of resistance to antiandrogen therapy. *Nat Med*, **10**, 33-39.
32. Tran, C., Ouk, S., Clegg, N.J., Chen, Y., Watson, P.A., Arora, V., Wongvipat, J., Smith-Jones, P.M., Yoo, D., Kwon, A. *et al.* (2009) Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science*, **324**, 787-790.

33. Scher, H.I., Beer, T.M., Higano, C.S., Anand, A., Taplin, M.E., Efstathiou, E., Rathkopf, D., Shelkey, J., Yu, E.Y., Alumkal, J. *et al.* (2010) Antitumour activity of MDV3100 in castration-resistant prostate cancer: a phase 1-2 study. *Lancet*, **375**, 1437-1446.
34. Darnell, J.E., Jr. (2002) Transcription factors as targets for cancer therapy. *Nat Rev Cancer*, **2**, 740-749.
35. Koehler, A.N. (2010) A complex task? Direct modulation of transcription factors with small molecules. *Curr Opin Chem Biol*, **14**, 331-340.
36. Block, K.M., Wang, H., Szabo, L.Z., Polaske, N.W., Henchey, L.K., Dubey, R., Kushal, S., Laszlo, C.F., Makhoul, J., Song, Z. *et al.* (2009) Direct inhibition of hypoxia-inducible transcription factor complex with designed dimeric epidithiodik etopiperazine. *J Am Chem Soc*, **131**, 18078-18088.
37. Kung, A.L., Zabudoff, S.D., France, D.S., Freedman, S.J., Tanner, E.A., Vieira, A., Cornell-Kennon, S., Lee, J., Wang, B., Wang, J. *et al.* (2004) Small molecule blockade of transcriptional coactivation of the hypoxia-inducible factor pathway. *Cancer Cell*, **6**, 33-43.
38. Kong, D., Park, E.J., Stephen, A.G., Calvani, M., Cardellina, J.H., Monks, A., Fisher, R.J., Shoemaker, R.H. and Melillo, G. (2005) Echinomycin, a small-molecule inhibitor of hypoxia-inducible factor-1 DNA-binding activity. *Cancer Res*, **65**, 9047-9055.
39. Pecot, C.V., Calin, G.A., Coleman, R.L., Lopez-Berestein, G. and Sood, A.K. (2011) RNA interference in the clinic: challenges and future directions. *Nat Rev Cancer*, **11**, 59-67.
40. Tan, S., Guschin, D., Davalos, A., Lee, Y.L., Snowden, A.W., Jouvenot, Y., Zhang, H.S., Howes, K., McNamara, A.R., Lai, A. *et al.* (2003) Zinc-finger protein-targeted gene regulation: genomewide single-gene specificity. *Proc Natl Acad Sci U S A*, **100**, 11997-12002.
41. Pavletich, N.P. and Pabo, C.O. (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*, **252**, 809-817.
42. Segal, D.J., Dreier, B., Beerli, R.R. and Barbas, C.F., 3rd. (1999) Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-GNN-3' DNA target sequences. *Proc Natl Acad Sci U S A*, **96**, 2758-2763.
43. Greisman, H.A. and Pabo, C.O. (1997) A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. *Science*, **275**, 657-661.
44. Liu, Q., Segal, D.J., Ghiara, J.B. and Barbas, C.F., 3rd. (1997) Design of polydactyl zinc-finger proteins for unique addressing within complex genomes. *Proc Natl Acad Sci U S A*, **94**, 5525-5530.
45. Carroll, D. (2008) Progress and prospects: zinc-finger nucleases as gene therapy agents. *Gene Ther*, **15**, 1463-1468.
46. Zhao, L. and Chmielewski, J. (2005) Inhibiting protein-protein interactions using designed molecules. *Curr Opin Struct Biol*, **15**, 31-34.
47. Moellering, R.E., Cornejo, M., Davis, T.N., Del Bianco, C., Aster, J.C., Blacklow, S.C., Kung, A.L., Gilliland, D.G., Verdine, G.L. and Bradner, J.E. (2009) Direct inhibition of the NOTCH transcription factor complex. *Nature*, **462**, 182-188.

48. Dickerson, R.E., Drew, H.R., Conner, B.N., Wing, R.M., Fratini, A.V. and Kopka, M.L. (1982) The anatomy of A-, B-, and Z-DNA. *Science*, **216**, 475-485.
49. Tse, W.C. and Boger, D.L. (2004) Sequence-selective DNA recognition: natural products and nature's lessons. *Chem Biol*, **11**, 1607-1617.
50. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol*, **13**, 284-299.
51. Coll, M., Frederick, C.A., Wang, A.H. and Rich, A. (1987) 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 U S A*, **84**, 8385-8389.
52. Pelton, J.G. and Wemmer, D.E. (1989) Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc Natl Acad Sci U S A*, **86**, 5723-5727.
53. Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B. (1992) Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc Natl Acad Sci U S A*, **89**, 7586-7590.
54. Wade, W.S., Mrksich, M. and Dervan, P.B. (1992) Design of Peptides That Bind in the Minor Groove of DNA at 5'-(a,T)G(a,T)C(a,T)-3' Sequences by a Dimeric Side-by-Side Motif. *J Am Chem Soc*, **114**, 8783-8794.
55. 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. and Dervan, P.B. (2007) Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron*, **63**, 6146-6151.
56. Trauger, J.W., Baird, E.E. and Dervan, P.B. (1998) Recognition of 16 base pairs in the minor groove of DNA by a pyrrole-imidazole polyamide dimer. *J Am Chem Soc*, **120**, 3534-3535.
57. Doss, R.M., Marques, M.A., Foister, S., Chenoweth, D.M. and Dervan, P.B. (2006) Programmable oligomers for minor groove DNA recognition. *J Am Chem Soc*, **128**, 9074-9079.
58. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
59. Trauger, J.W., Baird, E.E. and Dervan, P.B. (1996) Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature*, **382**, 559-561.
60. Herman, D.M., Turner, J.M., Baird, E.E. and Dervan, P.B. (1999) Cycle polyamide motif for recognition of the minor groove of DNA. *J Am Chem Soc*, **121**, 1121-1129.
61. Cohen, J.D., Sadowski, J.P. and Dervan, P.B. (2007) Addressing single molecules on DNA nanostructures. *Angew Chem Int Edit*, **46**, 7956-7959.
62. Edayathumangalam, R.S., Weyermann, P., Gottesfeld, J.M., Dervan, P.B. and Luger, K. (2004) Molecular recognition of the nucleosomal "supergroove". *Proc Natl Acad Sci U S A*, **101**, 6864-6869.
63. Edelson, B.S., Best, T.P., Olenyuk, B., Nickols, N.G., Doss, R.M., Foister, S., Heckel, A. and Dervan, P.B. (2004) Influence of structural variation on nuclear



- localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res*, **32**, 2802-2818.
64. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc Natl Acad Sci U S A*, **100**, 12063-12068.
  65. Chenoweth, D.M. and Dervan, P.B. (2010) Structural basis for cyclic Py-Im polyamide allosteric inhibition of nuclear receptor binding. *J Am Chem Soc*, **132**, 14521-14529.
  66. Chenoweth, D.M. and Dervan, P.B. (2009) Allosteric modulation of DNA by small molecules. *Proc Natl Acad Sci U S A*, **106**, 13175-13179.
  67. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G., Jr. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*, **101**, 16768-16773.
  68. Fong, G.H. (2009) Regulation of angiogenesis by oxygen sensing mechanisms. *J Mol Med*, **87**, 549-560.
  69. Muzikar, K.A., Nickols, N.G. and Dervan, P.B. (2009) Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc Natl Acad Sci U S A*, **106**, 16598-16603.
  70. Nickols, N.G. and Dervan, P.B. (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci U S A*, **104**, 10418-10423.
  71. Nicolaides, N.C., Galata, Z., Kino, T., Chrousos, G.P. and Charmandari, E. (2010) The human glucocorticoid receptor: molecular basis of biologic function. *Steroids*, **75**, 1-12.
  72. Aranda, A. and Pascual, A. (2001) Nuclear hormone receptors and gene expression. *Physiol Rev*, **81**, 1269-1304.
  73. Wang, X., Nagase, H., Watanabe, T., Nobusue, H., Suzuki, T., Asami, Y., Shinojima, Y., Kawashima, H., Takagi, K., Mishra, R. *et al.* (2010) Inhibition of MMP-9 transcription and suppression of tumor metastasis by pyrrole-imidazole polyamide. *Cancer Sci*, **101**, 759-766.
  74. Nagashima, T., Aoyama, T., Fukasawa, A., Watabe, S., Fukuda, N., Ueno, T., Sugiyama, H., Nagase, H. and Matsumoto, Y. (2009) Determination of pyrrole-imidazole polyamide in rat plasma by liquid chromatography-tandem mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci*, **877**, 1070-1076.

## Chapter 2

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

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

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



**Abstract**

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

## 2.1. Introduction

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

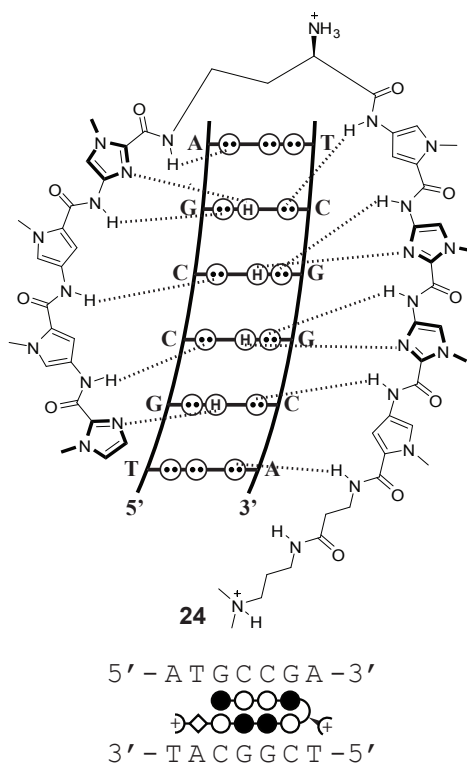
Polyamides constructed from *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methylhydroxypyrrole (Hp) amino acids comprise a class of synthetic ligands that bind within the minor groove of DNA in a sequence-specific manner.(18,19) Inspired by the natural products netropsin and distamycin A,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

Within the framework of the pairing rules, covalent linkages between two antiparallel

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

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

We have excluded Hp in this study in favor of more stable Py residue. The electron-rich Hp ring degrades in the presence of strong acid and free radical impurities. Based on the paradigm of unsymmetrical ring pairings, more robust rings such as the hydroxybenzimidazole/Py pair can replace Hp for T•A recognition.(33) In our work, hairpin polyamide design typically includes an Im/Py pairing at the hairpin terminus in order to impart G•C specificity at this position. As a result, eight-ring pyrrole-imidazole hairpin polyamides can specifically bind 7-bp sequences of the general form 5'-WWGNNNW-3' (where N = W, G, or C). There are 27 possible permutations that fall within these



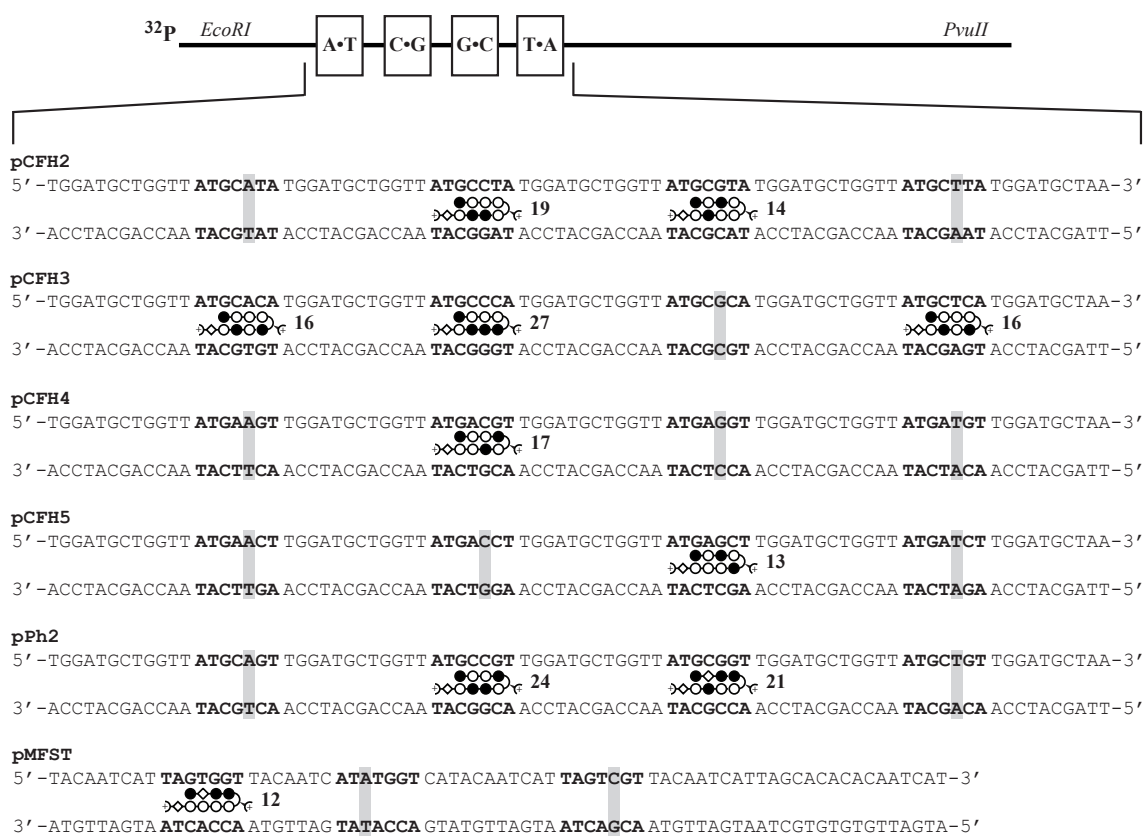
**Figure 2.1.** (Top) Model for the complex formed between hairpin polyamide **24** and its match DNA sequence. Circles with two dots represent the lone pairs of N(3) of purines and O(2) of pyrimidines. Circles containing an H represent the N(2) hydrogen of G. Hydrogen bonds are illustrated by dotted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

guidelines. Our group has published the energetics of 11 hairpin polyamides binding 11 distinct cognate sequences. Over half of the 27 sequences remain unreported. Searching our theses and notebooks confirmed 7 additional previously unpublished characterizations. Our studies during the past decade left only 9 out of 27 sequences unexamined; these are: 5'-WWGCGWW-3', 5'-WWGCCWW-3', 5'-WWGCWCW-3', 5'-WWGCCCW-3', 5'-WWGWCGW-3', 5'-WWGWGCW-3', 5'-WWGCGGW-3', 5'-WWGCCGW-3', and 5'-WWGWGGW-3'. To complete the library, we have synthesized nine hairpin polyamides designed to target these remaining sequences and assayed their binding affinity and sequence specificity by DNase I footprint titration experiments. In this study, we complete

the table of polyamide-DNA binding affinities, indicating each of the 27 general DNA sequences and an eight-ring hairpin polyamide that sequence specifically binds that 7-bp site. We hope that this centralized source of previously unpublished data proves helpful for other research groups currently modulating protein-DNA interfaces with DNA-binding small molecules.

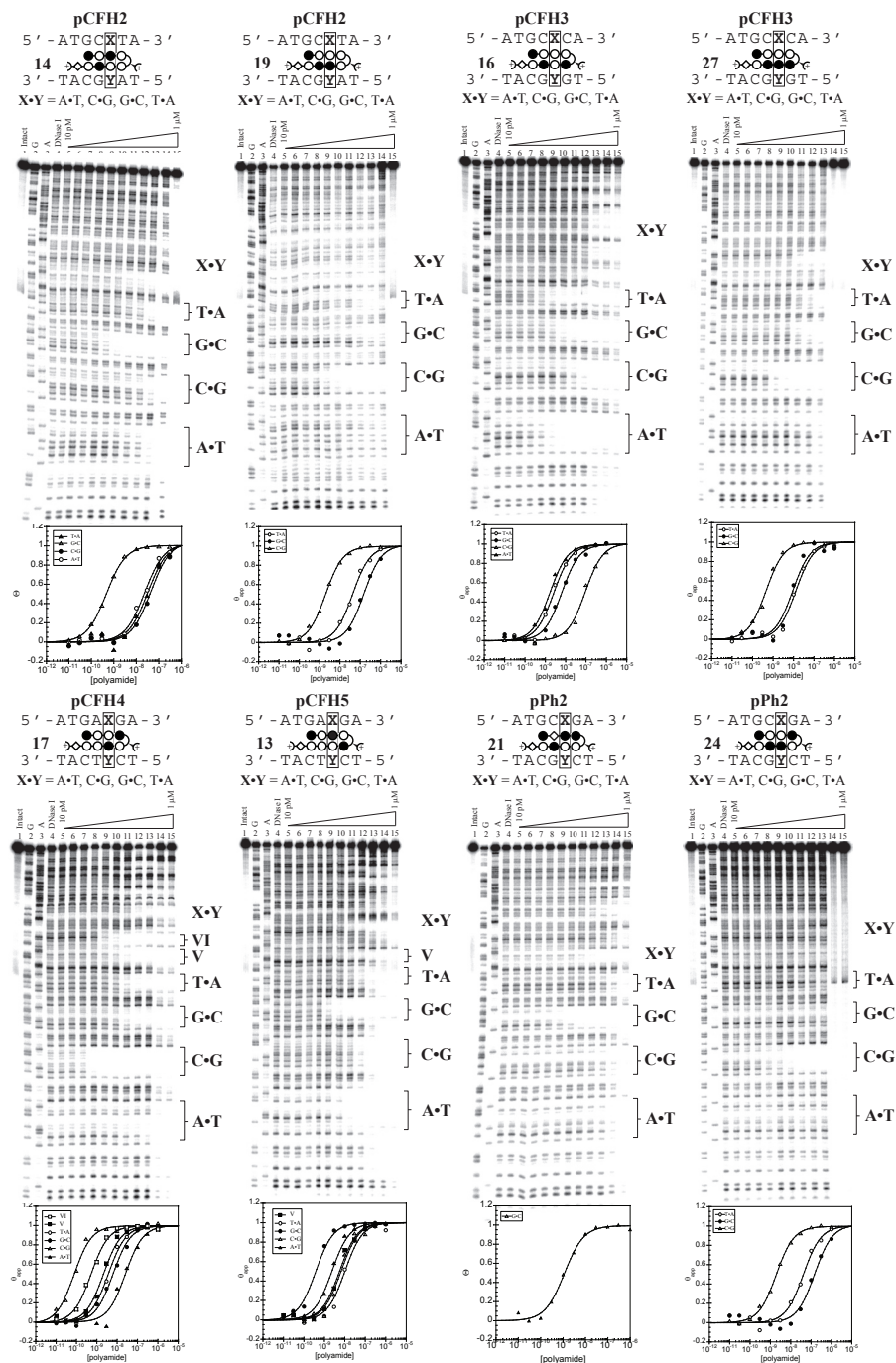
## 2.2. Results

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




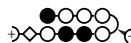
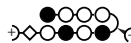






**Figure 2.2.** Plasmid design for pCFH2, pCFH3, pCFH4, pCFH5, pPh2, and pMFST, indicating the designed match and mismatch sites for hairpin polyamides **12**, **13**, **14**, **16**, **17**, **19**, **21**, **24**, and **27**. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;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

**2.2.2. DNA binding energetics.** Quantitative DNase I footprint titration experiments (10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub>, pH 7.0, 22°C) were performed on the 295 bp, 5'-<sup>32</sup>P-end-labeled PCR product of plasmids pCFH2,



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

pCFH3, pCFH4, pCFH5, pPh2, and pMFST.(35) Each of the nine polyamides was assayed on a plasmid containing its 7-bp match site according to the pairing rules, as well as three formal mismatch binding sites (Table 2.1 and Figure 2.2). The energetics of polyamide binding in the minor groove of DNA can be calculated from the Hill equation isotherms following DNase I cleavage and gel separation of the fragments. The equilibrium association constants ( $K_a$ ) determined in this way provide a quantitative measure of polyamide affinity at a given DNA binding site. Comparing these constants across the four potential binding sites allows a relative measure of specificity for each base pair at the targeted position.

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

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

The nine polyamides **12**, **13**, **14**, **16**, **17**, **19**, **21**, **24** and **27** bind their respective 7-bp match sites with equilibrium association constants that range from  $K_a = 7 \times 10^8 M^{-1}$  to  $4 \times 10^{10}$



$M^{-1}$  (Table 2.1 and Figure 2.3). The sequence specificities ( $K_{\text{match}}/K_{\text{single bp mismatch}}$ ) for these compounds vary from 4-fold to greater than 100-fold, further validating the pairing rules.

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

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

**Table 2.2.** Equilibrium association constants  $K_a$  ( $M^{-1}$ ). Values reported are the mean values from at least three DNase I footprint titration experiments. <sup>a</sup>Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; β-alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; the achiral γ-aminobutyric acid turn residue is shown as a semicircle connecting the two subunits; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits. <sup>b</sup> This paper, Table 2.1. <sup>c</sup> This paper, Figure 2.4.

### 2.3. Discussion

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

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

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

**Table 2.3.** Equilibrium association constants  $K_a$  (M<sup>-1</sup>). <sup>a</sup>  $K_a$  (M<sup>-1</sup>) values reported are the mean values from at least three DNase I footprint titration experiments. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0. Imidazole and pyrrole are shown as filled and non-filled circles, respectively;  $\beta$ -alanine is shown as a diamond; the dimethylaminopropylamide tail is shown as a half-circle with a plus; the achiral  $\gamma$ -aminobutyric acid turn residue is shown as a semicircle connecting the two subunits; and the chiral diaminobutyric acid turn residue is shown as a semicircle linked to a half-circle with a plus connecting the two subunits. <sup>b</sup> This paper, Table 1. <sup>c</sup> This paper, Figure 2.4. <sup>d</sup> Previously unpublished data.

## 2.4. Conclusion

We report here nine new eight-ring hairpin polyamides and assay their DNA-binding properties. These characterizations mark the completion of a 27-member polyamide library that covers all 7-bp DNA sequences 5'-WWGNNNW-3' (N = W, G, or C), where each compound binds its respective sequence with  $K_a \geq 1 \times 10^8 \text{ M}^{-1}$ . Typical fold change for match versus single base-pair mismatch sites is in the range of 4 to 100. Looking forward, one could imagine using a DNA microarray-based approach to interrogate the entire sequence-recognition profile of each member of the 27 hairpin polyamide library on every permutation of DNA sites seven base pairs in size (42). We hope that this compilation of polyamide-DNA binding affinities will serve as a resource for ongoing small molecule gene regulation projects.

## 2.5. Experimental

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



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

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

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

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

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

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

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

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

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

**2.5.2.7. Im-β-ImIm-(R)<sup>H<sub>2</sub>N</sup>γ-PyPyImPy-β-Dp (21).** Polyamide **21** was isolated upon lyophilization of the appropriate fractions as a white powder (0.3 mg, 0.2% recovery). UV (H<sub>2</sub>O) λ<sub>max</sub> = 309 nm; MALDI-TOF-MS (monoisotopic) *m/z* 1188.67 (1188.57 calcd



for  $[M + H]^+ C_{53}H_{70}N_{23}O_{10}^+$ ).

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

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

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

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

## References

1. Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E. and Dervan, P.B. (1997) Regulation of gene expression by small molecules. *Nature*, **387**, 202-205.
2. Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M. and Dervan, P.B. (1998) Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 12890-12895.
3. Mapp, A.K., Ansari, A.Z., Ptashne, M. and Dervan, P.B. (2000) Activation of gene expression by small molecule transcription factors. *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 3930-3935.
4. Janssen, S., Cuvier, O., Muller, M. and Laemmli, U.K. (2000) Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to *Drosophila melanogaster*. *Molecular Cell*, **6**, 1013-1024.
5. Maeshima, K., Janssen, S. and Laemmli, U.K. (2001) Specific targeting of insect and vertebrate telomeres with pyrrole and imidazole polyamides. *Embo Journal*, **20**, 3218-3228.
6. Gygi, M.P., Ferguson, M.D., Mefford, H.C., Lund, K.P., O'Day, C., Zhou, P., Friedman, C., van den Engh, G., Stolowitz, M.L. and Trask, B.J. (2002) Use of fluorescent sequence-specific polyamides to discriminate human chromosomes by microscopy and flow cytometry. *Nucleic Acids Research*, **30**, 2790-2799.
7. Crowley, K.S., Phillion, D.P., Woodard, S.S., Schweitzer, B.A., Singh, M., Shabany, H., Burnette, B., Hippenmeyer, P., Heitmeier, M. and Bashkin, J.K. (2003) Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorganic & Medicinal Chemistry Letters*, **13**, 1565-1570.
8. Kodadek, T., Reddy, M.M., Olivos, H.J., Bachhawat-Sikder, K. and Alluri, P.G. (2004) Synthetic molecules as antibody replacements. *Accounts of Chemical Research*, **37**, 711-718.
9. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 16768-16773.
10. Bando, T., Narita, A., Sasaki, S. and Sugiyama, H. (2005) Specific adenine alkylation by pyrrole-imidazole CBI conjugates. *Journal of the American Chemical Society*, **127**, 13890-13895.
11. Burnett, R., Melander, C., Puckett, J.W., Son, L.S., Wells, R.D., Dervan, P.B. and Gottesfeld, J.M. (2006) DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA·TTC repeats in Friedreich's ataxia. *Proceedings*

*of the National Academy of Sciences of the United States of America*, **103**, 11497-11502.

12. Belitsky, J.M., Leslie, S.J., Arora, P.S., Beerman, T.A. and Dervan, P.B. (2002) Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. *Bioorg Med Chem*, **10**, 3313-3318.
13. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc Natl Acad Sci U S A*, **100**, 12063-12068.
14. Edelson, B.S., Best, T.P., Olenyuk, B., Nickols, N.G., Doss, R.M., Foister, S., Heckel, A. and Dervan, P.B. (2004) Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res*, **32**, 2802-2818.
15. Pandolfi, P.P. (2001) Transcription therapy for cancer. *Oncogene*, **20**, 3116-3127.
16. Darnell, J.E. (2002) Transcription factors as targets for cancer therapy. *Nature Reviews Cancer*, **2**, 740-749.
17. Arkin, M.R. and Wells, J.A. (2004) Small-molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nature Reviews Drug Discovery*, **3**, 301-317.
18. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorganic & Medicinal Chemistry*, **9**, 2215-2235.
19. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol*, **13**, 284-299.
20. Pelton, J.G. and Wemmer, D.E. (1989) Structural Characterization of a 2-1 Distamycin A.D(Cgcaaattggc) Complex by Two-Dimensional Nmr. *Proceedings of the National Academy of Sciences of the United States of America*, **86**, 5723-5727.
21. Wade, W.S., Mrksich, M. and Dervan, P.B. (1992) Design of Peptides That Bind in the Minor Groove of DNA at 5'-(a,T)G(a,T)C(a,T)-3' Sequences by a Dimeric Side-by-Side Motif. *Journal of the American Chemical Society*, **114**, 8783-8794.
22. Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B. (1992) Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc Natl Acad Sci U S A*, **89**, 7586-7590.
23. Trauger, J.W., Baird, E.E. and Dervan, P.B. (1996) Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature*, **382**, 559-561.

24. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
25. Kielkopf, C.L., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) Structural basis for G·C recognition in the DNA minor groove. *Nature Structural Biology*, **5**, 104-109.
26. Kielkopf, C.L., White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) A structural basis for recognition of A·T and T·A base pairs in the minor groove of B-DNA. *Science*, **282**, 111-115.
27. Mrksich, M., Parks, M.E. and Dervan, P.B. (1994) Hairpin peptide motif - a new class of oligopeptides for sequence-specific recognition in the minor-groove of double-helical DNA. *Journal of the American Chemical Society*, **116**, 7983-7988.
28. Herman, D.M., Turner, J.M., Baird, E.E. and Dervan, P.B. (1999) Cycle polyamide motif for recognition of the minor groove of DNA. *Journal of the American Chemical Society*, **121**, 1121-1129.
29. Greenberg, W.A., Baird, E.E. and Dervan, P.B. (1998) A comparison of H-pin and hairpin polyamide motifs for the recognition of the minor groove of DNA. *Chemistry-a European Journal*, **4**, 796-805.
30. Heckel, A. and Dervan, P.B. (2003) U-pin polyamide motif for recognition of the DNA minor groove. *Chemistry-a European Journal*, **9**, 3353-3366.
31. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1999) Effects of gamma-turn and beta-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *J Am Chem Soc*, **121**, 1113-1120.
32. Herman, D.M., Baird, E.E. and Dervan, P.B. (1998) Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *Journal of the American Chemical Society*, **120**, 1382-1391.
33. Renneberg, D. and Dervan, P.B. (2003) Imidazopyridine/pyrrole and hydroxybenzimidazole/pyrrole pairs for DNA minor groove recognition. *Journal of the American Chemical Society*, **125**, 5707-5716.
34. Baird, E.E. and Dervan, P.B. (1996) Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *Journal of the American Chemical Society*, **118**, 6141-6146.
35. Trauger, J.W. and Dervan, P.B. (2001) Footprinting methods for analysis of pyrrole-imidazole/DNA complexes. *Methods in Enzymology*, **340**, 450-466.
36. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1997) A pyrrole-imidazole polyamide

motif for recognition of eleven base pair sequences in the minor groove of DNA. *Chem-Eur J*, **3**, 1600-1607.

37. Turner, J.M., Swalley, S.E., Baird, E.E. and Dervan, P.B. (1998) Aliphatic/aromatic amino acid pairings for polyamide recognition in the minor groove of DNA. *J Am Chem Soc*, **120**, 6219-6226.
38. Wang, C.C.C., Ellervik, U. and Dervan, P.B. (2001) Expanding the recognition of the minor groove of DNA by incorporation of beta-alanine in hairpin polyamides. *Bioorganic & Medicinal Chemistry*, **9**, 653-657.
39. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. 2nd ed. Cold Spring Harbor Laboratory, Plainview, NY.
40. Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods in Enzymology*, **65**, 499-560.
41. Iverson, B.L. and Dervan, P.B. (1987) Adenine specific DNA chemical sequencing reaction. *Nucleic Acids Research*, **15**, 7823-7830.
42. Warren, C.L., Kartochovil, N.C.S., Hauschild, K.E., Foister, S., Brezinski, M.L., Dervan, P.B., Phillips, G.N., Jr., Ansari, A.Z. (2006) Defining the sequence-recognition profile of DNA-binding molecules. *Proc Natl Acad Sci USA*, **103**, 867-872.

## Chapter 3

### Cyclic Pyrrole-Imidazole Polyamides Targeted to the Androgen Response Element

The text of this chapter was taken in part from a manuscript co-authored with David M. Chenoweth, Daniel A. Harki, Christian Dose, and Peter B. Dervan (Caltech).

(Chenoweth, D.M., Harki, D.A., Phillips, J.W., Dose, C., and Dervan, P.B. (2009) *J. Am. Chem. Soc.* **131**, 7182-7188).

**Abstract**

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

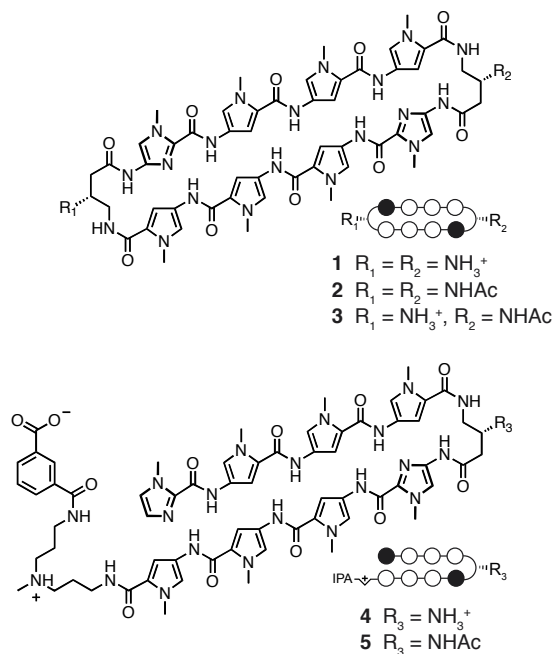
### 3.1. Introduction

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

In parallel with our gene regulation studies, a significant effort has been devoted to maximizing the biological potency of hairpin Py-Im polyamides through structural modifications. In particular, we have recently demonstrated that hairpin polyamides bearing the (*R*)- $\beta$ -amino- $\gamma$ -turn, such as polyamide **4**, possess favorable binding affinities to DNA and are useful in gene regulation studies (Figure 3.1).<sup>5g</sup> A significant effort exists in our laboratory to regulate aberrant AR-activated gene expression in prostate cancer.<sup>8</sup> To further optimize lead oligomer **4** it would seem reasonable that closing the hairpin with an identical linker, yielding a cyclic structure **1**, may further enhance DNA affinity (Figure 3.1). Previous syntheses of cyclic polyamides using solid-phase protocols are characterized by low reaction yields due to inefficient macrocyclization.<sup>10</sup> We report here the solution-phase synthesis of cyclic polyamides **1-3** with an improved high yield cyclization step. In addition we examined the DNA binding properties of these compounds by thermal duplex DNA melting as well as preliminary studies of their *in vitro* ADMET properties. Cyclic



Py-Im polyamides **1-3** were shown to regulate endogenous gene expression in cell culture experiments.



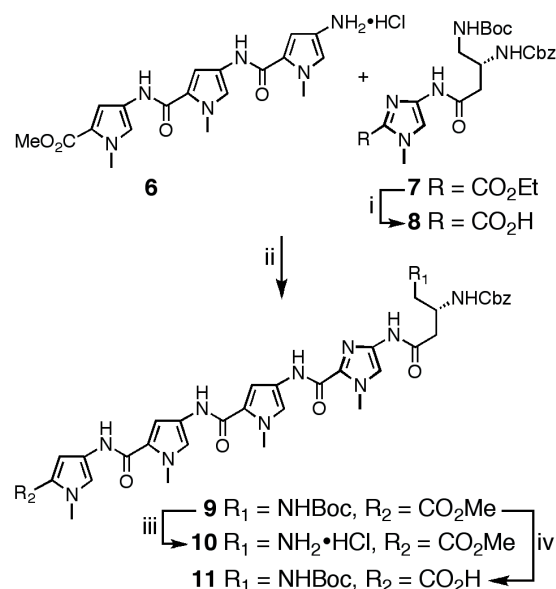
**Figure 3.1.** Chemical structures for cyclic and hairpin polyamides 1-5 targeted to the DNA-sequence 5'-WGWWCW-3' are shown with ball-and-stick models. Ball-and-stick representation legend: black and white circles represent N-methylimidazole and N-methylpyrrole units, respectively, IPA denotes the terminal isophthalic acid substituent, and white half-diamond with + sign represents the triamine linker unit.

## 3.2 Results and Discussion

### 3.2.1 Solution-Phase Synthesis of Cyclic Polyamides

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

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



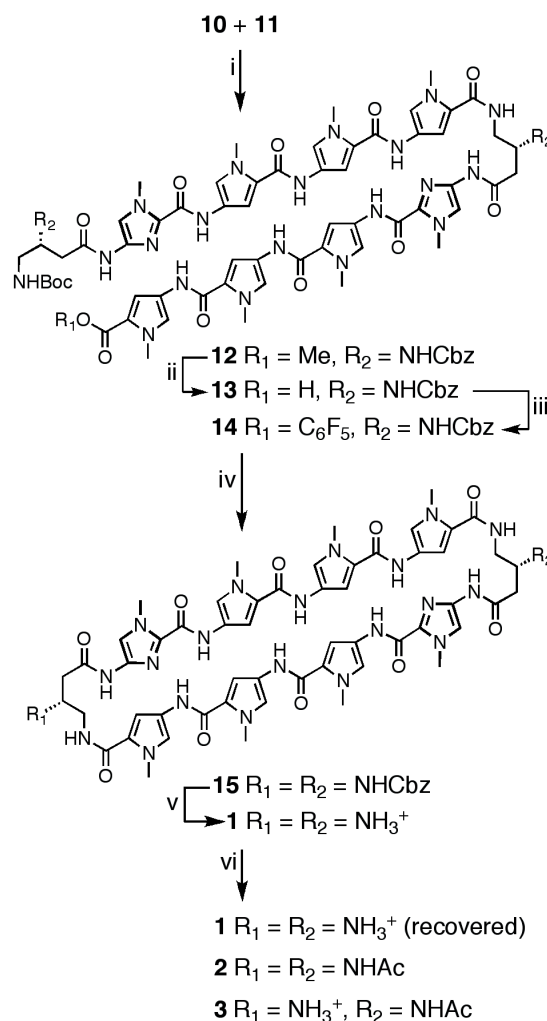
**Scheme 3.1.** Preparation of **10** and **11**.

Reagents: (i) 1.0 M KOH (aq), MeOH, 37°C, 2 h, 95%; (ii) **8**, PyBOP, DMF, DIEA, **6**, 23°C, 4 h, 96%; (iii) 4.0 M HCl in 1,4-dioxane, 23°C, 2 h, 99% (iv) 1.0 M NaOH (aq), 1,4-dioxane, 42°C, 3 h, 95%.

polyamide **1**, compound **9** was elaborated to amine salt **10** (99% yield) by reaction with HCl in 1,4-dioxane. Carboxylic acid **11** was generated by saponification of **10** with NaOH

in 1,4-dioxane in 95 % yield.

Assembly of the acyclic advanced intermediate **12** was achieved by PyBOP-mediated coupling of intermediates **10** and **11** in 94 % yield (Scheme 3.2). A small excess of amine salt **10** was utilized to drive the reaction to completion. Saponification of ester **12** proceeded smoothly with aqueous NaOH in 1,4-dioxane, yielding **13** in 93% yield. Activation of acid **13** as the pentafluorophenol ester **14** provided the necessary functionality to afford macrocyclization following removal of the terminal t-butyl carbamate (Boc) protecting group. In our hands, we found the pentafluorophenol ester sufficiently activated the terminal acid for amide coupling while avoiding undesired oligomerization and/or decomposition processes that are conceivable with more reactive functionalities, such as acid chlorides. Premature initiation of the macrocyclization reaction was tempered by keeping the terminal amine protonated



**Scheme 3.2.** Preparation of **1**, **2**, and **3**. Reagents: (i) PyBOP, DMF, DIEA, 23°C, 2 h, 94%; (ii) 1.0 M NaOH (aq), 1,4-dioxane, 40°C, 4 h, 93%; (iii) DCM, DCC, pentafluorophenol, DMAP, 23°C, 12 h, 80%; (iv) a) TFA, DCM, 23°C, concentrate; b) DMF, acetonitrile, DIEA, 0-23°C, 3 days; (v)  $\text{CF}_3\text{SO}_3\text{H}$ ,  $\text{CF}_3\text{CO}_2\text{H}$ , 23°C, 5 min, 68% over 3 steps; (vi) NMP,  $\text{Ac}_2\text{O}$ , 23°C, 18% of **1** (recovered), 40% of **2**, 22% of **3**.




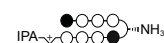
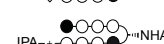
until transferred into a dilute solution of acetonitrile. Addition of an amine base (DIEA) then generated the free terminal amine, which could then undergo macrocyclization in dilute solvent conditions to deliver **15**, which was directly deprotected. The benzyl carbamate protecting groups were cleaved by treatment with superacid conditions (trifluoromethylsulfonic acid-trifluoroacetic acid) to provide **1** in 68% yield over 3 steps.

Controlled acetylation of **1** by reaction with substoichiometric quantities of Ac<sub>2</sub>O in NMP/DIEA provided a statistical population of **1** (18%), **2** (40%), and **3** (22%) that were easily separable by preparative HPLC. Acetylated hairpin **5** was prepared using excess Ac<sub>2</sub>O/pyridine in 95% yield from previously reported amine hairpin **4**.<sup>11</sup>

### 3.2.2. Thermal Stabilization of DNA-duplexes by Polyamides

Quantitative DNase I footprint titrations have historically been utilized to measure polyamide-DNA binding affinities and specificities.<sup>13</sup> However, this method is limited to measuring  $K_a$  values  $\leq 2 \times 10^{10} \text{ M}^{-1}$ , which invalidates this technique for quantifying the exceptionally high DNA-binding affinities of cycles **1-3**.<sup>14</sup> The magnitude of DNA thermal stabilization ( $\Delta T_m$ ) of DNA-polyamide complexes has been utilized to rank order polyamides with high DNA binding affinities.<sup>5g,15</sup> Accordingly, we have employed melting temperature analysis ( $\Delta T_m$ ) for dissecting differences in DNA-binding affinities of hairpin versus cyclic polyamides. Spectroscopic analyses were performed on a 14-mer duplex DNA mimicking the androgen response element (ARE) DNA sequence, 5'-TTGCTGTTCTGCAA-3' DNA duplex, which contains one polyamide binding site. As shown in Table 3.1 polyamides

**1-5** provided an increase in the duplex DNA melting temperature relative to the individual DNA duplex, thereby confirming polyamide-DNA binding. Chiral hairpin **4** led to an increased melting temperature  $\Delta T_m = 18.4 \text{ }^\circ\text{C}$  whereas cyclic polyamide **1** yielded a higher  $\Delta T_m$ -value of  $23.6 \text{ }^\circ\text{C}$ . Cyclic polyamides **1-3** reveal stronger stabilizations than parent hairpin analogs **4** and **5**. Acylation of the  $\beta$ -amino turns was shown to decrease the thermal stabilization

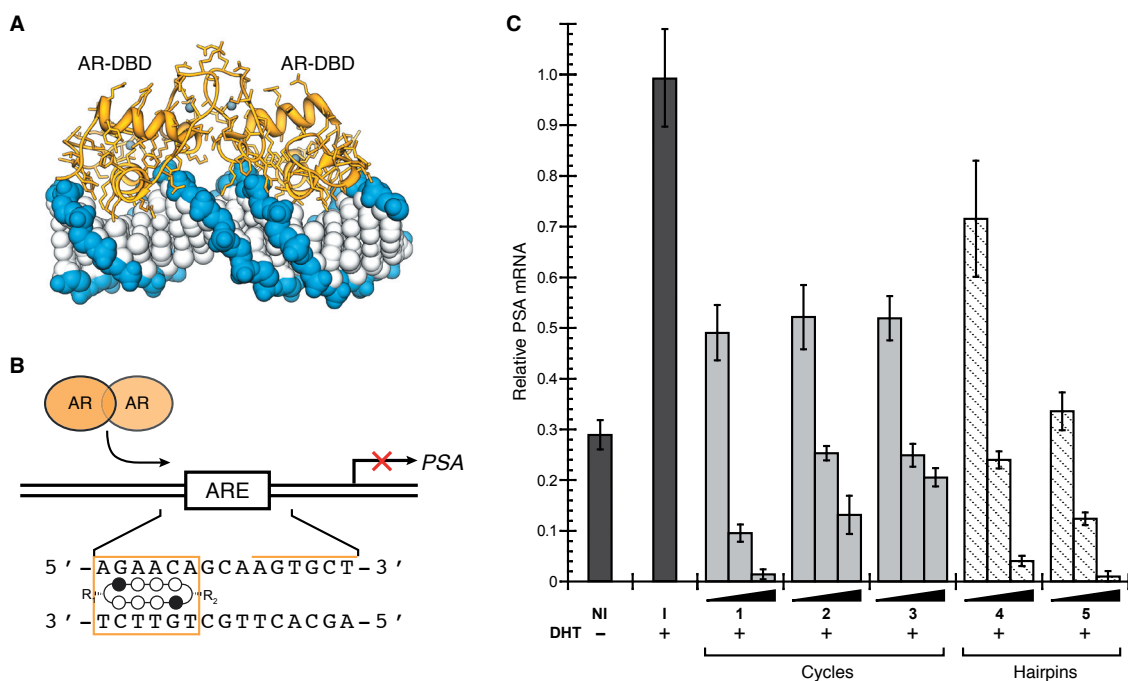
ARE dsDNA sequence =		5' -TTGC	<b>TGTTCT</b>	GCAA-3'
		3' -AACG	<b>ACAAGA</b>	CGTT-5'
Polyamides		$T_m / ^\circ\text{C}$		$\Delta T_m / ^\circ\text{C}$
—		60.0 ( $\pm 0.3$ )		—
 (1)		83.5 ( $\pm 0.5$ )		23.6 ( $\pm 0.6$ )
 (2)		81.2 ( $\pm 0.2$ )		21.3 ( $\pm 0.4$ )
 (3)		82.0 ( $\pm 0.0$ )		22.1 ( $\pm 0.3$ )
 (4)		78.4 ( $\pm 0.5$ )		18.4 ( $\pm 0.6$ )
 (5)		76.0 ( $\pm 0.5$ )		16.1 ( $\pm 0.6$ )

**Table 3.1.**  $T_m$  values for polyamides for **1 - 5**. All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses.  $\Delta T_m$  values are given as  $T_m(\text{DNA/polyamide}) - T_m(\text{DNA})$ . The propagated error in  $\Delta T_m$  measurements is the square root of the sum of the square of the standard deviations for the  $T_m$  values.

values in both hairpin and cyclic motifs, presumably due to the loss of the electrostatic benefit of the cation protonated amine on the turn unit.

### 3.2.3. Biological Assay for Cell Permeability

Hairpin polyamides have been shown to modulate endogenous gene expression in living cells by disrupting transcription factor-DNA binding in gene promoters.<sup>2,7-9</sup> Recently, hairpin polyamide **4** was shown to inhibit androgen receptor-mediated expression of prostate specific antigen (PSA) in LNCaP cells by targeting the DNA sequence 5'-AGAACA-3' found in the ARE.<sup>8</sup> We utilized this cell culture transcription assay to investigate the biological activity of cyclic polyamides **1-3** in comparison to hairpin polyamides **4** and **5** (Figure 3.2). Since small structural changes to polyamides have been shown to



**Figure 3.2.** Targeting the ARE with DNA-binding polyamides. (a) X-ray crystal structure of androgen receptor homodimer DNA-binding domain bound to the sequence 5'-CTGTTCTTGATGTTCTGG-3' (PDB 1r4i).<sup>16</sup> (b) Map of the PSA-ARE site (top) and schematic representation of a cyclic polyamide targeting the PSA-ARE site 5'-AGAACA-3'. (c) Inhibition of induced PSA mRNA expression in LNCaP cells by cyclic Py-Im polyamides **1-3** and hairpin polyamides **4** and **5** (dosed at 0.3, 3, and 30  $\mu$ M) by real-time quantitative PCR. The results were normalized to a DHT-induced, untreated control (control = 1), and the error bars represent the standard error of the mean of a single experiment performed in biological triplicate. The entire experiment was reproduced four times, with similar results. NI = non-induced, I = induced, DHT = dihydrotestosterone.

correlate with differences in cellular uptake properties,<sup>5</sup> it was not obvious whether cyclic polyamides **1-3** would permeate cell membranes and exhibit biological activity comparable to hairpin polyamides **4** and **5**. Quantitative real-time RT-PCR analysis of DHT-induced PSA expression revealed cyclic polyamides **1-3** all decreased PSA mRNA levels in LNCaP cells, with cycle **1** exhibiting comparable activity to acetylated hairpin polyamide **5**. Based on these results we can infer that this class of cyclic Py-Im polyamides is cell permeable and can regulate endogenous gene expression in cell culture.

#### **3.2.4. ADMET Studies of Polyamides 1 and 5**

Due to the promising cell culture results obtained with cyclic polyamide **1** and hairpin polyamide **5** against PSA gene expression, we contracted preclinical *in vitro* absorption, distribution, metabolism, excretion and toxicity (ADMET)<sup>17</sup> studies for both compounds.<sup>18</sup> Results from this study are summarized as follows and additional detail can be found in the Supporting Information. Polyamides **1** and **5** were both found to exhibit low Caco-2 permeability, suggesting neither compound may be orally available. Both **1** and **5** were found to be almost exclusively protein bound in plasma with half-lives greater than 2 h. A recent positron emission tomography (PET)-based biodistribution study of a related hairpin polyamide in mice revealed high levels of liver occupancy following tail vein dosage.<sup>19</sup> Based on this result we investigated the liver stability of candidate polyamides **1** and **5**. Microsomal intrinsic clearance studies found half-lives greater than 3 hr for **1** and **5** in both human and rat liver microsomes, and no significant inhibition was measured against any cytochrome P450 isoform examined (Cyp1A2/CEC, Cyp2C8/DBF, Cyp2C9/DBF, Cyp2C19/DBF, Cyp2D6/AMMC, Cyp3A4/BFC, Cyp3A4/DBF). Furthermore, no obvious toxicity ( $IC_{50} > 100 \mu\text{M}$ ) was observed in the human hepatocellular carcinoma cell line HepG2. In addition, standard hERG FastPatch assays of cardiac toxicity found both polyamides (**1** and **5**) were devoid of unwanted inhibition ( $IC_{50} > 100 \mu\text{M}$ ).

### 3.3. Conclusion

We describe a solution-phase synthesis methodology for preparing cyclic Py-Im polyamides, highlighted by an efficient macrocyclization between the alkyl linker amine and a pentafluorophenyl ester-activated amino acid. The three cyclic Py-Im polyamides possessed high DNA-binding affinities and were capable of accessing the nucleus in cell culture as judged by their ability to downregulate AR-activated PSA expression in cell culture. Preclinical ADMET analysis of cyclic polyamide **1** and hairpin polyamide **5** revealed favorable druglike properties such as high liver stability and low toxicity. Ongoing work is focused on characterizing the precise molecular interactions between cyclic polyamides and their cognate DNA sequences by high-resolution crystallographic studies.

### 3.4. Materials and methods

**3.4.1. General.** Chemicals and solvents were purchased from Sigma-Aldrich and were used without further purification. (*R*)-3,4-Cbz-Dbu(Boc)-OH was purchased from Senn Chemicals AG (code number 44159). All DNA oligomers were purchased HPLC purified from Integrated DNA Technologies. Water (18 M $\Omega$ ) was purified using a Millipore MilliQ purification system. Centrifugation was performed in a Beckman Coulter bench-top centrifuge (Allegra 21R) equipped with a Beckman swing-out rotor (model S4180). Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250  $\times$  4.6 mm, 5  $\mu$ m), a diode array detector, and the mobile phase consisted of a gradient of acetonitrile (MeCN) in 0.1% (v/v) aqueous CF<sub>3</sub>CO<sub>2</sub>H. Preparative HPLC was performed on an Agilent 1200 system equipped with a solvent degasser, diode array detector, and a Phenomenex Gemini column (250  $\times$  21.2 mm, 5  $\mu$ m). A gradient of MeCN in 0.1% (v/v) aqueous CF<sub>3</sub>CO<sub>2</sub>H was utilized as the mobile phase. UV-Vis measurements were made on a Hewlett-Packard diode array spectrophotometer (Model 8452 A) and polyamide concentrations were measured in 0.1% (v/v) aqueous CF<sub>3</sub>CO<sub>2</sub>H using an extinction coefficient of 69200 M<sup>-1</sup>·cm<sup>-1</sup> at  $\lambda_{\text{max}}$  near 310 nm. NMR spectroscopy was performed on a Varian instrument operating at 499.8 MHz (for <sup>1</sup>H) or 125.7 MHz (for <sup>13</sup>C) at ambient temperature. All NMR analyses were performed in DMSO-*d*<sub>6</sub>, and chemical shifts are reported in parts per million relative to the internal solvent peak referenced to 2.49 (for <sup>1</sup>H) or 39.5 (for <sup>13</sup>C). High-resolution mass spectrometry (HRMS) was recorded in positive-ion mode by fast-atom bombardment (FAB<sup>+</sup>) on a JOEL JMS-600H instrument or by electrospray ionization (ESI<sup>+</sup>) on a Waters Acquity UPLC-LCT Premiere XE TOF-MS system.

**3.4.2. UV Absorption Spectrophotometry.** Melting temperature analysis was performed on a Varian Cary 100 spectrophotometer equipped with a thermo-controlled cell holder



possessing a cell path length of 1 cm. A degassed aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM MgCl<sub>2</sub>, and 5 mM CaCl<sub>2</sub> at pH 7.0 was used as analysis buffer. DNA duplexes and hairpin polyamides were mixed in 1:1 stoichiometry to a final concentration of 2 μM for each experiment. Prior to analysis, samples were heated to 90 °C and cooled to a starting temperature of 23 °C with a heating rate of 5 °C/min for each ramp. Denaturation profiles were recorded at λ = 260 nm from 23 °C to 90 °C with a heating rate of 0.5 °C/min. The reported melting temperatures were defined as the maximum of the first derivative of the denaturation profile.

**3.4.3. Measurement of Androgen-Induced PSA mRNA.** Experiments were performed as described previously<sup>8</sup> with the following modifications: (1) all compounds and controls were prepared in neat DMSO then diluted with media to a final concentration of 0.1% DMSO, and (2) mRNA was isolated with the RNEasy 96 kit (Qiagen, Valencia, CA).

**3.4.4. BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-Im-CO<sub>2</sub>H (8).** A solution of BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-Im-CO<sub>2</sub>Et **7** (450 mg, 0.894 mmol) dissolved in MeOH (1.0 mL) and aqueous KOH (1N, 2.0 mL, 2.0 mmol) was stirred at 37 °C for 2 h. The reaction mixture was added to a cooled (ice bath) solution of distilled H<sub>2</sub>O (10 mL) pre-acidified with aqueous HCl (1N, 2.0 mL, 2.0 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H<sub>2</sub>O (10 mL) and collected by centrifugation. The resultant solid, which contained a small amount of residual H<sub>2</sub>O, was frozen and lyophilized to dryness, and then suspended in excess anhydrous Et<sub>2</sub>O, filtered, and the filter cake washed with copious amounts of anhydrous Et<sub>2</sub>O. Drying of the brown solid *in vacuo* yielded saponified dimer **8** (404 mg, 95%). <sup>1</sup>H NMR: δ 10.46 (s, 1H), 7.47 (s, 1H), 7.31-7.26 (m, 5H), 7.02 (d, *J* = 8.3 Hz, 1H), 6.79 (t, *J* = 5.4 Hz, 1H), 4.97 (s, 2H), 3.92 (m, 1H), 3.88 (s, 3H), 3.60 (s, 1H), 3.01 (m, 2H), 2.41 (m, 2H), 1.35 (s, 9H); <sup>13</sup>C NMR: δ 167.6, 160.0, 155.8, 155.4, 137.11, 137.09, 131.6, 128.2, 127.66, 127.55, 114.6, 77.7, 65.1, 48.7, 43.5, 38.0, 35.4, 28.2; HRMS

(FAB<sup>+</sup>) calc'd for C<sub>22</sub>H<sub>30</sub>N<sub>5</sub>O<sub>7</sub> [M+H]<sup>+</sup> 476.2145, found 476.2130.

**3.4.5. BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me (9).** A solution of BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImCO<sub>2</sub>H **8** (300 mg, 0.631 mmol) and PyBOP (345 mg, 0.663 mmol) in DMF (3.2 mL) and DIEA (330 mL, 1.9 mmol) was stirred at 23°C for 10 min. The solution was then treated with solid (powdered) HCl•H<sub>2</sub>N-PyPyPy-CO<sub>2</sub>Me **6** (288 mg, 0.663 mmol) and stirred at 23°C for 4 h. The solution was then added to distilled H<sub>2</sub>O (10 mL) pre-acidified with aqueous HCl (1N, 2 mL, 2 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H<sub>2</sub>O (10 mL) and collected by centrifugation (repeated 3X). The resultant solid, which contained a small amount of residual H<sub>2</sub>O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et<sub>2</sub>O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et<sub>2</sub>O and dried *in vacuo* to yield BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me **9** as a tan solid (518 mg, 96%). <sup>1</sup>H NMR: δ 10.17 (s, 1H), 10.00 (s, 1H), 9.95 (s, 1H), 9.93 (s, 1H), 7.46 (d, *J* = 1.7 Hz, 1H), 7.44 (s, 1H), 7.31-7.29 (m, 5H), 7.27 (d, *J* = 1.7 Hz, 1H), 7.23 (d, *J* = 1.7 Hz, 1H), 7.14 (d, *J* = 1.7 Hz, 1H), 7.07 (d, *J* = 1.7 Hz, 1H), 7.04 (d, *J* = 8.3 Hz, 1H), 6.90 (d, *J* = 2.0 Hz, 1H), 6.81 (t, *J* = 5.9 Hz, 1H), 4.98 (s, 2H), 3.96 (m, 1H), 3.95 (s, 3H), 3.85 (s, 3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.73 (s, 3H), 3.03 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H); <sup>13</sup>C NMR: δ 167.8, 160.8, 158.5, 158.4, 155.84, 155.81, 155.5, 137.1, 136.0, 134.0, 128.3, 127.7, 127.6, 123.06, 123.00, 122.5, 122.2, 121.2, 120.7, 118.7, 118.6, 118.5, 114.0, 108.4, 104.9, 77.8, 65.2, 50.9, 48.8, 43.6, 38.2, 36.20, 36.18, 36.09, 34.9, 28.2; HRMS (FAB<sup>+</sup>) calc'd for C<sub>41</sub>H<sub>49</sub>N<sub>11</sub>O<sub>10</sub> [M]<sup>+</sup> 855.3663, found 855.3688.

**3.4.6. HCl•H<sub>2</sub>N-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me (10).** A solution of BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me **9** (125 g, 0.146 mmol) in anhydrous HCl in 1,4-dioxane (4.0 M, 10 mL) was stirred at 23°C for 2 h. The mixture was then diluted with 100 mL of anhydrous

Et<sub>2</sub>O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et<sub>2</sub>O and dried *in vacuo* to yield HCl•H<sub>2</sub>N-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me **10** as a brown solid (114 mg, 99%). <sup>1</sup>H NMR: δ 10.38 (s, 1H), 9.98 (s, 1H), 9.96 (s, 1H), 9.94 (s, 1H), 8.10 (m, 3H), 7.46 (d, *J* = 1.7 Hz, 1H), 7.45 (s, 1H), 7.42 (d, *J* = 8.3 Hz, 1H), 7.34-7.28 (m, 5H), 7.28 (d, *J* = 1.7 Hz, 1H), 7.24 (d, *J* = 1.7 Hz, 1H), 7.16 (d, *J* = 1.7 Hz, 1H), 7.08 (d, *J* = 1.7 Hz, 1H), 6.90 (d, *J* = 1.7 Hz, 1H), 5.02 (m, 2H), 4.14 (m, 1H), 3.95 (s, 3H), 3.85 (s, 3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.73 (s, 3H), 3.02 (m, 2H), 2.63 (m, 2H); <sup>13</sup>C NMR: δ 167.0, 160.8, 158.5, 158.4, 155.7, 136.8, 135.7, 134.0, 128.3, 127.8, 127.7, 123.05, 122.97, 122.5, 122.2, 121.1, 120.7, 118.64, 118.60, 118.5, 108.4, 104.9, 65.6, 50.9, 46.7, 42.0, 38.2, 36.2, 36.1, 36.0, 34.9; HRMS (FAB<sup>+</sup>) calc'd for C<sub>36</sub>H<sub>42</sub>N<sub>11</sub>O<sub>8</sub> [M+H]<sup>+</sup> 756.3218, found 756.3211.

**3.4.7. BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>H (11).** A solution of BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me **10** (200 mg, 0.234 mmol) dissolved in 1,4-dioxane (2.3 mL) and aqueous NaOH (1N, 2.3 mL, 2.3 mmol) was stirred at 42 °C for 3 h. The solution was then added to distilled H<sub>2</sub>O (5 mL) pre-acidified with aqueous HCl (1N, 2.3 mL, 2.3 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H<sub>2</sub>O (10 mL) and collected by centrifugation (repeated 2X). The resultant solid, which contained a small amount of residual H<sub>2</sub>O, was frozen and lyophilized to dryness, and then suspended in excess anhydrous Et<sub>2</sub>O, filtered, and the filter cake washed with copious amounts of anhydrous Et<sub>2</sub>O. Drying of the tan solid *in vacuo* yielded BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>H **11** (187 mg, 95%). <sup>1</sup>H NMR: δ 12.15 (s, 1H), 10.21 (s, 1H), 10.00 (s, 1H), 9.96 (s, 1H), 9.92 (s, 1H), 7.44 (s, 1H), 7.42 (d, *J* = 1.7 Hz, 1H), 7.31-7.29 (m, 5H), 7.28 (d, *J* = 1.5 Hz, 1H), 7.24 (d, *J* = 1.5 Hz, 1H), 7.16 (d, *J* = 1.5 Hz, 1H), 7.08 (m, 2H), 6.85 (d, *J* = 1.7 Hz, 1H), 6.82 (t, *J* = 5.7 Hz, 1H), 4.98 (s, 2H), 3.95 (m, 4H), 3.85 (s, 3H), 3.84 (s, 3H), 3.81 (s, 3H), 3.03 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H); <sup>13</sup>C NMR: δ 167.8, 162.0, 158.44, 158.38, 155.80, 155.77, 155.4, 137.1, 136.0, 133.9,

128.3, 127.7, 127.6, 123.0, 122.7, 122.6, 122.2, 121.2, 120.2, 119.5, 118.6, 118.5, 114.0, 108.4, 104.87, 104.83, 77.7, 65.1, 48.8, 43.5, 38.2, 36.2, 36.13, 36.06, 34.9, 28.2; HRMS (FAB<sup>+</sup>) calc'd for C<sub>40</sub>H<sub>47</sub>N<sub>11</sub>O<sub>10</sub> [M<sup>+</sup>] 841.3507, found 841.3498.

**3.4.8. BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me (12).** A solution of BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>H **11** (25 mg, 0.029 mmol) and PyBOP (17 mg, 0.031 mmol) in DMF (150 mL) and DIEA (16 mL, 0.089 mmol) was stirred at 23°C for 20 min. The solution was then treated with solid (powdered) HCl•H<sub>2</sub>N-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me **10** (25 mg, 0.031 mmol) and stirred at 23°C for 2 h. The solution was then added to distilled H<sub>2</sub>O (10 mL) pre-acidified with aqueous HCl (1N, 1 mL, 1 mmol), yielding a precipitate that was isolated by centrifugation (~ 4500 rpm). The residual solid was again suspended in distilled H<sub>2</sub>O (10 mL) and collected by centrifugation (repeated 3X). The resultant solid, which contained a small amount of residual H<sub>2</sub>O, was frozen and lyophilized to dryness. The solid was triturated with anhydrous Et<sub>2</sub>O and filtered over a sintered glass funnel. The resultant solid was washed with copious amounts of anhydrous Et<sub>2</sub>O and dried *in vacuo* to yield BocHN-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-(R)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me **12** as a tan solid (44 mg, 94%). <sup>1</sup>H NMR: δ 10.20 (s, 1H), 10.16 (s, 1H), 9.98 (s, 2H), 9.94-9.91 (m, 4H), 7.99 (m, 1H), 7.46 (d, *J* = 1.7 Hz, 1H), 7.45 (s, 1H), 7.44 (s, 1H), 7.32-7.14 (m, 18H), 7.07 (m, 2H), 7.03 (d, *J* = 8.3 Hz, 1H), 6.92 (s, 1H), 6.90 (d, *J* = 2.0 Hz, 1H), 6.80 (t, *J* = 5.6 Hz, 1H), 4.99 (m, 4H), 4.10 (m, 1H), 3.95 (m, 7H), 3.85-3.83 (m, 15H), 3.79 (s, 3H), 3.73 (s, 3H), ~3.30 (m, 2H, obstructed by H<sub>2</sub>O peak), 3.04 (m, 2H), 2.53-2.44 (m, 4H, partially obstructed by NMR solvent), 1.36 (s, 9H); <sup>13</sup>C NMR: δ 167.9, 167.8, 161.6, 160.8, 158.5, 158.44, 158.42, 155.8, 155.6, 155.5, 137.1, 136.0, 134.00, 133.98, 128.3, 127.7, 127.63, 127.60, 123.11, 123.07, 123.00, 122.80, 122.77, 122.5, 122.3, 122.2, 122.1, 121.3, 120.8, 118.69, 118.66, 118.62, 118.52, 118.0, 114.1, 108.4, 104.9, 104.8, 104.5, 77.8, 65.21, 65.16, 50.9, 48.83, 48.78, 43.6, 42.2, 38.4, 38.2, 36.2, 36.10, 36.07, 36.0, 34.9, 28.2; HRMS (TOF-ESI<sup>+</sup>) calc'd for C<sub>76</sub>H<sub>87</sub>N<sub>22</sub>O<sub>17</sub> [M+H]<sup>+</sup> 1579.6620, found 1579.6580.

**3.4.9. BocHN-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>H (13).** A solution of BocHN-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Me **12** (25 mg, 0.0158 mmol) dissolved in 1,4-dioxane (376 mL) and aqueous NaOH (1N, 253 mL, 0.253 mmol) was stirred at 40 °C for 4 h. The solution was then added to distilled H<sub>2</sub>O (5 mL) pre-acidified with aqueous HCl (1N, 253 mL, 0.253 mmol), yielding a precipitate which was diluted with another 15 mL H<sub>2</sub>O and was then isolated by centrifugation (~ 4500 rpm). The resultant solid, which contained a small amount of residual H<sub>2</sub>O, was frozen and lyophilized to dryness, and then suspended in excess anhydrous Et<sub>2</sub>O, triturated, filtered, and the filter cake washed with copious amounts of anhydrous Et<sub>2</sub>O. Drying of the tan solid *in vacuo* yielded BocHN-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>H **13** (23 mg, 93%). <sup>1</sup>H NMR: δ 12.13 (br s, 1H), 10.23 (s, 1H), 10.20 (s, 1H), 9.98-9.90 (m, 6H), 8.01 (m, 1H), 7.443 (s, 1H), 7.439 (s, 1H), 7.42 (d, *J* = 1.7 Hz, 1H), 7.30-7.15 (m, 18H), 7.07 (m, 3H), 6.92 (m, 1H), 6.84 (d, *J* = 2.0 Hz, 1H), 6.81 (t, *J* = 5.6 Hz, 1H), 5.00 (m, 2H), 4.98 (s, 2H), 4.10 (m, 1H), 3.95 (m, 7H), 3.85-3.83 (m, 12H), 3.81 (s, 3H), 3.78 (s, 3H), ~3.30 (m, 2H, obstructed by H<sub>2</sub>O peak), 3.03 (m, 2H), 2.53 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H); <sup>13</sup>C NMR: δ 167.94, 167.87, 162.4, 162.0, 161.6, 158.54, 158.50, 158.43, 155.8, 155.75, 155.74, 155.6, 155.5, 137.1, 136.0, 133.91, 133.90, 128.3, 127.7, 127.60, 127.57, 123.10, 123.07, 122.8, 122.7, 122.6, 122.3, 122.24, 122.17, 121.16, 121.15, 120.3, 119.5, 118.64, 118.61, 118.5, 118.0, 114.10, 114.06, 108.5, 105.0, 104.5, 77.8, 65.2, 65.1, 48.8, 43.6, 42.2, 38.4, 38.2, 36.2, 36.14, 36.10, 36.08, 36.0, 35.8, 34.9, 28.2; HRMS (TOF-ESI<sup>+</sup>) calc'd for C<sub>75</sub>H<sub>86</sub>N<sub>22</sub>O<sub>17</sub> [M+2H]<sup>2+</sup>/2 783.3271, found 783.3237.

**3.4.10. BocHN-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>Pfp (14).** A solution of BocHN-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-(*R*)<sup>β</sup>-CbzHN<sub>γ</sub>-ImPyPyPy-CO<sub>2</sub>H **13** (250 mg, 0.160 mmol) and DCC (66 mg, 0.320 mmol) in CH<sub>2</sub>Cl<sub>2</sub> (8.8 mL) was stirred at 23°C for 45 min. The solution was then treated with DMAP (2 mg, 0.016 mmol) followed by pentafluorophenol (175

mg, 0.950 mmol) and stirred at 23°C for 12 h. The reaction mixture was then loaded onto a silica gel column with CH<sub>2</sub>Cl<sub>2</sub> and eluted with step gradients of 100% CH<sub>2</sub>Cl<sub>2</sub> to 100% acetone with incremental steps of 5% acetone. The product was concentrated *in vacuo* to yield BocHN-(R)<sup>β-CbzHN</sup>γ-ImPyPyPy-(R)<sup>β-CbzHN</sup>γ-ImPyPyPy-CO<sub>2</sub>Pfp **14** as a brown solid (221 mg, 80%). <sup>1</sup>H NMR: δ 10.20 (s, 1H), 10.16 (s, 1H), 10.08 (s, 1H), 9.99-9.91 (m, 5H), 7.99 (m, 1H), 7.73 (d, *J* = 1.7 Hz, 1H), 7.444 (s, 1H), 7.440 (s, 1H), 7.30-7.12 (m, 20 H), 7.06 (d, *J* = 1.2 Hz, 1H), 7.03 (d, *J* = 8.5 Hz, 1H), 6.92 (s, 1H), 6.80 (t, *J* = 5.4 Hz, 1H), 5.00 (m, 2H), 4.98 (m, 2H), 4.11 (m, 1H), 3.95 (m, 7H), 3.88 (s, 3H), 3.86-3.84 (m, 12H), 3.78 (s, 3H), ~3.30 (m, 2H, obstructed by H<sub>2</sub>O peak), 3.04 (m, 2H), 2.52 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H); HRMS (TOF-ESI<sup>+</sup>) calc'd for C<sub>81</sub>H<sub>85</sub>F<sub>5</sub>N<sub>22</sub>O<sub>17</sub> [M+2H]<sup>2+</sup>/2 866.3192, found 866.3236.

**3.4.11. cyclo-(-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-) (1).** A solution of BocHN-(R)<sup>β-CbzHN</sup>γ-ImPyPyPy-(R)<sup>β-CbzHN</sup>γ-ImPyPyPy-CO<sub>2</sub>Pfp **14** (84 mg, 0.049 mmol) in anhydrous CF<sub>3</sub>CO<sub>2</sub>H:CH<sub>2</sub>Cl<sub>2</sub> (1:1, 4 mL) was stirred at 23 °C for 10 min prior to being concentrated to dryness in a 500 mL round bottom flask. The residue was then dissolved in cold (0 °C) DMF (10 mL), followed by immediate dilution with MeCN (300 mL) and DIEA (1.6 mL). The reaction mixture was left at 23 °C for 3 days without stirring. (Note: The solution turns cloudy as the macrocyclization proceeds.) The reaction mixture was concentrated to a volume of 11 mL and added to a solution of H<sub>2</sub>O (30 mL) and aqueous HCl (1N, 9.2 mL) at 0 °C. The protected intermediate cyclo-(-ImPyPyPy-(R)<sup>β-CbzHN</sup>γ-ImPyPyPy-(R)<sup>β-CbzHN</sup>γ-) **15** was isolated by centrifugation (~ 4500 rpm), lyophilized to dryness, and then suspended in excess anhydrous Et<sub>2</sub>O, triturated, filtered, and the filter cake washed with copious amounts of anhydrous Et<sub>2</sub>O. Drying of the tan solid *in vacuo* yielded the protected intermediate cyclo-(-ImPyPyPy-(R)<sup>β-CbzHN</sup>γ-ImPyPyPy-(R)<sup>β-CbzHN</sup>γ-) **15**, HRMS (TOF-ESI<sup>+</sup>) calc'd for C<sub>70</sub>H<sub>76</sub>N<sub>22</sub>O<sub>14</sub> [M+2H]<sup>2+</sup>/2 724.2956, found 724.2925. This material was immediately deprotected by dissolving in CF<sub>3</sub>CO<sub>2</sub>H (2 mL) followed by addition of

CF<sub>3</sub>SO<sub>3</sub>H (100 μL) at 23 °C for 5 min. The solution was then frozen and DMF (2 mL) was layered over the frozen solution. The thawed solution was diluted with H<sub>2</sub>O (6 mL) and purified by reverse-phase HPLC to give a white solid after lyophilization. The solid was suspended in excess anhydrous Et<sub>2</sub>O, triturated, filtered, and the filter cake washed with copious amounts of anhydrous Et<sub>2</sub>O. Drying of the white solid *in vacuo* yielded cyclo(-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-) **1** (46 mg, 68%). <sup>1</sup>H NMR: δ 10.56 (s, 2H), 9.91 (s, 4H), 9.88 (s, 2H), 8.17 (t, *J* = 5.6 Hz, 2H), 7.96 (m, 6H), 7.40 (s, 2H), 7.31 (d, *J* = 1.6 Hz, 2H), 7.27 (d, *J* = 1.6 Hz, 2H), 7.19 (d, *J* = 1.6 Hz, 2H), 7.00 (d, *J* = 1.7 Hz, 2H), 6.96 (d, *J* = 1.6 Hz, 2H), 6.94 (d, *J* = 1.7 Hz, 2H), 3.94 (s, 6H), 3.83 (s, 12H), 3.80 (s, 6H), 3.71 – 3.66 (m, 2H), 3.49 – 3.27 (m, 4H, partially obstructed by H<sub>2</sub>O peak), 2.79 (dd, *J* = 16.1 Hz, 6.0 Hz, 2H), 2.60 (dd, *J* = 15.2 Hz, 5.2 Hz, 2H). HRMS (TOF-ESI<sup>+</sup>) calc'd for C<sub>54</sub>H<sub>63</sub>N<sub>22</sub>O<sub>10</sub> [M+H]<sup>+</sup> 1179.5098, found 1179.5087.

**3.4.12. cyclo(-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-) (2) and cyclo(-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-) (3).** A solution of cyclo(-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-) **1** (2.81 mg, 2.0 mmol) in anhydrous NMP (200 mL) and DIEA (20 mL) at 23 °C was treated with a solution of Ac<sub>2</sub>O in NMP (0.122 M, 6.8 mL). After 10 min the reaction mixture was treated with another 6.8 mL of Ac<sub>2</sub>O in NMP (0.122 M) and allowed to stand for 5 hr. The reaction mixture was then diluted to a volume of 10 mL by addition of a 4:1 solution of aqueous CF<sub>3</sub>CO<sub>2</sub>H (0.1% v/v):MeCN (5 mL), followed by additional aqueous CF<sub>3</sub>CO<sub>2</sub>H (0.1% v/v, 5 mL), and then purified by reverse-phase HPLC to yield cyclo(-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-) **1** (363 nmol, 18%), cyclo(-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-) **2** (800 nmol, 40%), and cyclo(-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-) **3** (432 nmol, 22%). Cyclo(-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-) **2** HRMS (TOF-ESI<sup>+</sup>) calc'd for C<sub>56</sub>H<sub>65</sub>N<sub>22</sub>O<sub>11</sub> [M+H]<sup>+</sup> 1221.5203, found 1221.5204. Cyclo(-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-ImPyPyPy-(R)<sup>β-AcHN</sup>γ-) **3** HRMS (TOF-ESI<sup>+</sup>) calc'd for C<sub>58</sub>H<sub>68</sub>N<sub>22</sub>O<sub>12</sub> [M+2H]<sup>2+</sup>/2 633.2646, found 633.2631.

**3.4.12. ImPyPyPy-(R)<sup>β-H<sub>2</sub>N</sup>γ-ImPyPyPy-(+)-IPA (4).** Prepared as described in the preceding manuscript.<sup>11</sup>

**3.4.13. ImPyPyPy-(R)<sup>β-AcHN</sup>γ-ImPyPyPy-(+)-IPA (5).** A solution of polyamide **4**<sup>11</sup> (7.4 mg, 5.03 μmoles, assumes **4** as the mono-CF<sub>3</sub>CO<sub>2</sub>H salt) in DMF (1.76 mL) was treated with a solution of Ac<sub>2</sub>O in pyridine (10% v/v, 240 μL, 0.254 mmoles Ac<sub>2</sub>O). The solution was allowed to stand at 23 °C for 30 min, and then acidified with aqueous CF<sub>3</sub>CO<sub>2</sub>H (15% v/v, 2 mL). After 5 min the solution was further diluted with distilled H<sub>2</sub>O (5 mL), purified by preparative RP-HPLC, and lyophilized to dryness. Suspension of the residual solid in anhydrous Et<sub>2</sub>O, following by filtration and drying under high vacuum yielded **5** (6.7 mg, 95%). HRMS (FAB<sup>+</sup>) calc'd for C<sub>67</sub>H<sub>79</sub>N<sub>22</sub>O<sub>13</sub> [M+H]<sup>+</sup> 1399.6191, found 1399.6181.

**3.5. Supporting Information Available:** A summary of ADMET results, <sup>1</sup>H and <sup>13</sup>C NMR spectra, and analytical HPLC purity analyses of selected compounds. This material is available free of charge via the Internet at <http://pubs.acs.org>. The full ADMET report including experimental conditions can be accessed via the Dervan laboratory homepage at <http://dervan.caltech.edu>.



### 3.6. References

- (1) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215-2235.
- (2) Dervan, P. B., and Edelson, J. A. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299.
- (3) (a) Trauger, J. W., Baird, E. E., and Dervan, P. B. *Nature* **1996**, *382*, 559-561. (b) White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. *Nature* **1998**, *391*, 468-470. (c) Kielkopf, C. L., Baird, E. E., Dervan, P. B., and Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104-109. (d) Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B., and Rees, D. C. *Science* **1998**, *282*, 111-115.
- (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., and Dervan, P. B. *Tetrahedron* **2007**, *63*, 6146-6151.
- (5) (a) Belitsky, J. M., Leslie, S. J., Arora, P. S., Beerman, T. A., and Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 3313-3318. (b) Crowley, K. S., Phillion, D. P., Woodard, S. S., Scheitzer, B. A., Singh, M., Shabany, H., Burnette, B., Hippenmeyer, P., Heitmeier, M., and Bashkin, J. K. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1565-1570. (c) Best, T. P., Edelson, B. S., Nickols, N. G., and Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 12063-12068. (d) Edelson, B. S., Best, T. P., Olenyuk, B., Nickols, N. G., Doss, R. M., Foister, S., Heckel, A., and Dervan, P. B. *Nucleic Acids. Res.* **2004**, *32*, 2802-2818. (e) Xiao, X., Yu, P., Lim, H. S., Sikder, D., and Kodadek, T. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 2865-2868. (f) Nickols, N. G., Jacobs, C. S., Farkas, M. E., and Dervan, P. B. *Nucleic Acids. Res.* **2007**, *35*, 363-370. (g) Dose, C., Farkas, M. E., Chenoweth, D. M., and Dervan, P. B. *J. Am. Chem. Soc.* **2008**, *130*, 6859-6866. (h) Hsu, C. F., and Dervan, P. B. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5851-5855.
- (6) (a) Gottesfeld, J. M., Melander, C., Suto, R. K., Raviol, H., Luger, K., and Dervan, P. B.

- Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides. *J. Mol. Biol.* **2001**, *309*, 615-629. (b) Suto, R. K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., and Luger, K. Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.* **2003**, *326*, 371-380. (c) Edayathumangalam, R. S., Weyermann, P., Gottesfeld, J. M., Dervan, P. B., and Luger, K. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 6864-6869. (d) Dudouet, B., Burnett, R., Dickinson, L. A., Wood, M. R., Melander, C., Belitsky, J. M., Edelson, B., Wurtz, N., Briehn, C., Dervan, P. B., and Gottesfeld, J. M. *Chem. Biol.* **2003**, *10*, 859-867.
- (7) (a) Olenyuk, B. Z., Zhang, G. J., Klco, J. M., Nickols, N. G., Kaelin, W. G., and Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 16768-16773. (b) Kageyama, Y., Sugiyama, H., Ayame, H., Iwai, A., Fujii, Y., Huang, L. E., Kizaka-Kondoh, S., Hiraoka, M., and Kihara, K. *Acta. Oncol.* **2006**, *45*, 317-324. (c) Nickols, N. G., Jacobs, C. S., Farkas, M. E., and Dervan, P. B. *ACS Chem. Biol.* **2007**, *2*, 561-571.
- (8) Nickols, N. G., and Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 10418-10423.
- (9) (a) Matsuda, H., Fukuda, N., Ueno, T., Tahira, Y., Ayame, H., Zhang, W., Bando, T., Sugiyama, H., Saito, S., Matsumoto, K., and others, O. *J. Am. Soc. Neph.* **2006**, *17*, 422-432. (b) Yao, E. H., Fukuda, N., Ueno, T., Matsuda, H., Matsumoto, K., Nagase, H., Matsumoto, Y., Takasaka, A., Serie, K., Sugiyama, H., and Sawamura, T. *Hypertension* **2008**, *52*, 86-92.
- (10) (a) Cho, J.; Parks, M. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10389-10392. (b) Zhang, Q., Dwyer, T. J., Tsui, V., Case, D. A., Cho, J., Dervan, P. B., Wemmer, D. E. *J. Am. Chem. Soc.* **2004**, *126*, 7958-7966. (c) Herman, D. M., Turner, J. M., Baird, E. E., Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 1121-1129. (d) Melander, C., Herman, D.

M., Dervan, P. B. *Chem. Eur. J.* **2000**, *6*, 4487-4497.

(11) Chenoweth D. M., Harki, D. A., and Dervan, P.B. *J. Am. Chem. Soc.* **2009**, *131*, 7175–7181.

(12) (a) Baird, E. E., Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146. (b) Wurtz, N. R., Turner, J. M., Baird, E. E., Dervan, P. B. *Org. Lett.* **2001**, *3*, 1201-1203. (c) Jaramillo, D., Liu, Q., Aldrich-Wright, J., Tor, Y. *J. Org. Chem.* **2004**, *69*, 8151-8153.

(13) Trauger, J. W., Dervan P. B. *Methods Enzymol.* **2001**, *340*, 450-466.

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

(15) (a) Pilch, D. S., Poklar, N., Gelfand, C. A., Law, S. M., Breslauer, K. J., Baird, E. E., Dervan, P. B. *Proc. Natl. Acad. Sci. USA.* **1996**, *93*, 8306-8311. (b) Pilch, D. S., Poklar, N., Baird, E. E., Dervan, P. B., Breslauer, K. J. *Biochemistry.* **1999**, *38*, 2143-2151.

(16) Shaffer, P. L., Jivan, A., Dollins, D. E., Claessens, F., Gewirth, D. T. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 4758-4763.

(17) For a review of pharmacokinetics in drug discovery see: Ruiz-Garcia, A., Bermejo, M., Moss, A., Casabo, V.G. *J. Pharm. Sci.* **2008**, *97*, 654-690.

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

(19) Harki, D. A., Satyamurthy, N., Stout, D. B., Phelps, M. E., Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 13039-13044.

## **Chapter 4**

### **Characterization of Py-Im Polyamide Androgen Receptor Antagonists in Hormone-Refractory Prostate Cancer Cells**

John W. Phillips and Peter B. Dervan

**Abstract**

We have previously developed DNA-binding polyamides targeted to the consensus androgen response element (ARE). These compounds bind to DNA in the nucleus of live, hormone-sensitive cultured prostate cancer cells and reduce the occupancy of androgen receptor at some AREs, including the one in the intronic enhancer region of FKBP5. Expression of FKBP5, KLK3 and other androgen receptor (AR) target genes is correspondingly reduced. In this study, we investigate the activity of these compounds in the LNCaP-AR cell line, which is a tissue culture model of hormone-refractory disease. We determine that the polyamides maintain their activity against PSA mRNA expression but fail to decrease AR occupancy at the FKBP5 ARE in LNCaP-AR cells, which suggests an alternative mechanism for inhibition of AR-target gene expression exists in this cell line. Cytotoxic stress is also known to disrupt AR signaling; we show that p53 is stabilized and programmed cell death pathways are activated in response to polyamide treatment. Further investigation into the cytotoxic effects of polyamide treatment is recommended.

#### 4.1. Introduction

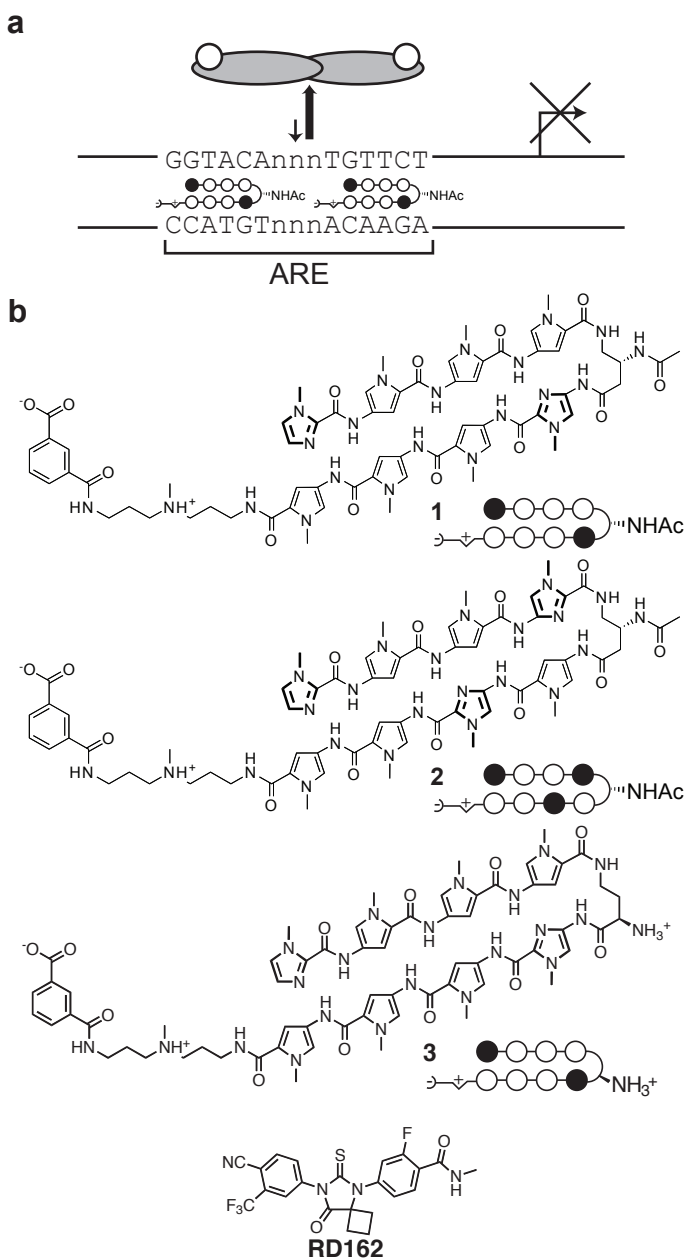
The dramatic clinical response of primary prostate cancer to androgen deprivation therapy highlights the importance of the androgen receptor (AR) in its pathogenesis (1). In fact, chemical or surgical castration can establish clinical remission in many patients. Pharmacologic agents used for this purpose include bicalutamide, a synthetic anti-androgen that binds the ligand-binding pocket of the AR and prevents assembly of the coactivators required for AR-driven gene expression (2,3). For many patients, bicalutamide and other hormonally active agents maintain efficacy at least six months (4). When the disease recurs, it is invariably resistant to conventional hormonal therapy, metastatic, and lethal.

Recent studies have shown that the progression to hormone refractory prostate cancer (HRPC) is critically dependent on androgen receptor signaling. Recurrent tumors often show evidence of AR function despite low levels of circulating androgens (5). Mechanisms proposed to explain the reactivation of AR signaling include mutation of the ligand binding domain or upregulation of the receptor, ligand-independent AR stimulation through other signaling pathways, and changes in AR-coactivators (4). Of all of these, gene amplification is likely the most important mechanism of resistance, as it is the expression change most often found in patient samples of recurrent disease and the only change common to several mouse xenograft models of HPRC. AR remains a crucial therapeutic target even after first-line anti-hormone therapies have failed.

We have previously reported the characterization of a Py-Im polyamide antagonist that can disrupt AR binding to androgen response elements (ARE) in the hormone-sensitive prostate cancer cell line LNCaP (6). Polyamides are programmable, high-affinity, DNA sequence-specific molecules that can permeate cells, traffic to the nucleus, and bind DNA in chromatin (7,8). The polyamide androgen antagonist was designed to bind 5'-WGWWCW-3' (W = A or T) sequences in DNA, which blocked AR binding at a subset of the 5'-GGTACAnnnTGTTCT-3' sequences in regulatory regions of AR-target genes

(Figure 4.1a). This small molecule inhibited the dihydrotestosterone (DHT)-induced expression of many AR-target genes, including *TMPRSS2*, *FKBP5*, and *PSA*. The activity of the polyamide was comparable to bicalutamide, a synthetic anti-androgen currently in clinical use.

Given the dependence of recurrent prostate cancer progression on AR and the ability of a simple upregulation of AR expression to defeat anti-androgens targeted to the ligand-binding pocket, the continued ability to abrogate AR signaling in recurrent prostate cancer disease models by disrupting the protein-DNA interface would represent an important proof of concept. In this study, we explore the activity of a polyamide-ARE antagonist in a cell culture model of recurrent, hormone-sensitive prostate cancer: the LN-AR cell line (5). This cell line is a derivative of LNCaP cells containing a stably integrated vector that overexpresses



**Figure 4.1.** Disrupting the AR/ARE interface in HRPC. (a) A DNA-binding polyamide targeting the ARE consensus sequence disrupts the AR/ARE interface and offers an alternative anti-androgen strategy. (b) Chemical structures and ball-and-stick representations of the polyamides and the synthetic anti-androgen RD162.



AR. In this model, traditional anti-androgens like bicalutamide lose their efficacy. In fact, AR overexpression *in vitro* transforms bicalutamide into an agonist of AR activity. It is possible that compounds targeted to the protein-DNA interface would not be defeated by this resistance mechanism.

In this study we investigate the viability of targeting the protein-DNA interface as an alternative anti-androgen strategy for recurrent prostate cancer. We examine the activity of an improved polyamide ARE antagonist (compound **1**) in LNCaP and LN-AR cells (Figure 4.1b). We compare its activity to that of the original polyamide ARE antagonist (compound **3**). These studies also include mismatch compound **2**, which is targeted to the unrelated DNA sequence 5'-WGWCGW-3'. Additionally, we include a comparison of the polyamides with the second- and third-generation anti-androgens, bicalutamide and RD162 (9). We show that **1** maintains its potency against AR-driven mRNA expression in LN-AR cells, but that it no longer blocks AR binding to the ARE of a known AR-responsive gene, FKBP5. We show that the inhibition of PSA mRNA expression produced by **1** is accompanied by activation of cell stress and programmed cell death pathways. These results have important implications for the proposed mechanism of AR inhibition by polyamides.

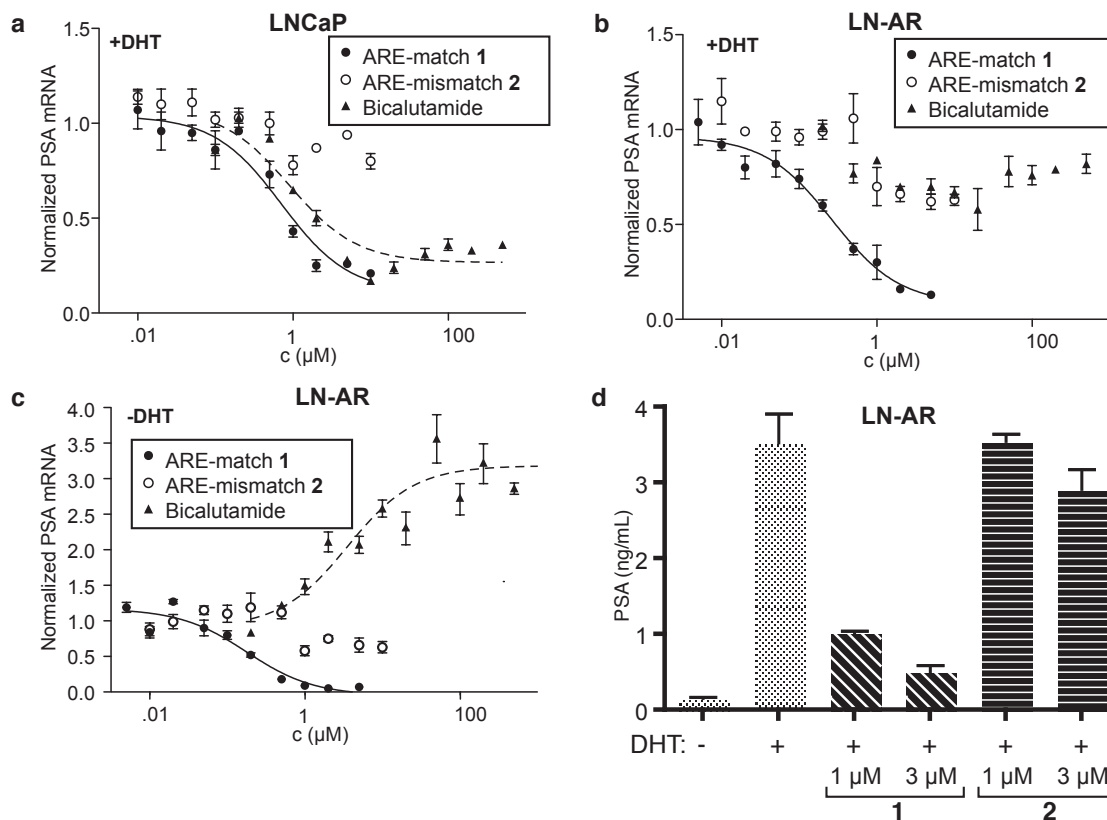
## 4.2. Results

**4.2.1. Compound synthesis.** Using a combination of solution- and solid-phase techniques, we prepared an improved ARE-antagonist polyamide (compound **1**) that incorporated an acetylated beta-amino hairpin turn (Figure 4.1). This design change was recently demonstrated to improve the affinity and biological activity of the alpha-amino compound used in our hormone-sensitive studies. We also synthesized an appropriate mismatch compound designed to bind an unrelated sequence (5'-WGWCGW-3') (compound **2**). Compound **3** is the original, unacetylated alpha-turn compound from the hormone sensitive cell lines studies and is included in some assays as a control. We also synthesized the third-generation anti-androgen, RD162, using previously-published protocols for use as an additional control (9).

**4.2.2. Inhibition of PSA expression in LNCaP and LN-AR cells.** We tested the ability of the improved ARE-match polyamide **1** to inhibit AR-driven gene expression in LNCaP and LN-AR cells and compared its activity to the mismatch compound **2** and to bicalutamide. Using qPCR analysis of DHT-induced PSA mRNA expression, we found that polyamide **1** inhibited AR-driven gene expression in the hormone-sensitive and the hormone-refractory setting (Figure 4.2a, b). The polyamide also inhibited basal PSA expression in hormone-stripped media without DHT; bicalutamide acted as an agonist in this setting (Figure 4.2c). The inhibition was dose dependent and efficacious, as PSA levels could be reduced to their non-induced levels. The  $IC_{50}$  for the inhibition of PSA mRNA by **1** was approximately  $0.5 \mu\text{M}$  in LNCaP cells and  $0.3 \mu\text{M}$  in LN-AR cells (Table 4.1). The downregulation was also sequence specific, for no

Compound	Cell line		
	LNCaP +DHT	LN-AR +DHT	LN-AR -DHT
<b>1</b>	500±100 nM	300±30 nM	120±30 nM
<b>2</b>	>10 $\mu\text{M}$	>10 $\mu\text{M}$	>10 $\mu\text{M}$
Bicalutamide	900 nM	>1 mM	(3.5 $\mu\text{M}$ )

**Table 4.1.**  $IC_{50}$  values for inhibition of PSA mRNA expression. Values represent the average of  $IC_{50}$  determinations for three biological replicates, mean  $\pm$  s.d.



**Figure 4.2.** A Py-Im polyamide antagonist of AR-ARE binding inhibits expression of AR target gene PSA. (a) Dose-dependent inhibition of DHT-induced PSA mRNA expression by polyamide 1 and bicalutamide in hormone-sensitive LNCaP cells as measured by qRT-PCR analysis of cDNA (untreated, DHT-induced = 1). Inhibition by 2 is much less potent effective. (b) Inhibition of PSA mRNA expression in DHT-induced LN-AR cells (untreated, DHT-induced = 1). Bicalutamide and 2 are now similarly ineffective but compound 1 maintains its potency. (c) Effects of polyamides and bicalutamide on basal PSA mRNA expression in LN-AR cells (untreated = 1). Bicalutamide induces PSA expression under these conditions. Polyamide 2 is ineffective. (d) Effects of polyamides 1 and 2 on PSA protein expression in LN-AR cells as measured by ELISA. Compound 1 inhibits PSA production; compound 2 is less effective.

dose-dependent inhibition was observed with the mismatch compound. As expected, bicalutamide failed to inhibit DHT-induced PSA mRNA expression in LN-AR cells across a wide range of concentrations.

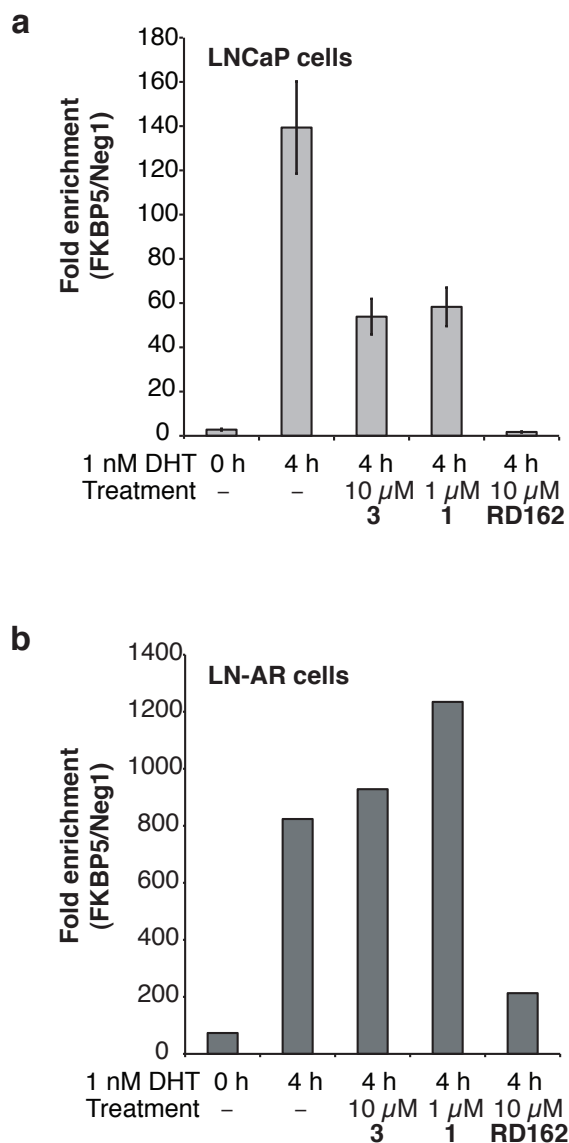
We also examined the polyamide's activity under non-induced conditions in charcoal-stripped media (no DHT). In this setting, bicalutamide induced PSA mRNA expression in accordance with the previously published result. The match polyamide again inhibited PSA expression, and the mismatch compound had little effect.

We confirmed the results of the mRNA expression analysis with an enzyme-linked immunosorbent assay (ELISA) for PSA in LN-AR cell culture supernatants (Figure 4.2d).

The protein expression levels of PSA were also decreased for the match polyamide, and relatively unaffected for cells treated with the mismatch compound or with bicalutamide.

**Polyamide treatment does not reduce AR occupancy at AREs in LN-AR cells.**

Previously, we have shown that compound **3** downregulates the expression of the AR-responsive gene FKBP5 in DHT-stimulated LNCaP cells at 10  $\mu$ M. Additionally, this compound decreased AR occupancy at the FKBP5 intronic ARE to 50% of DHT-induced levels. This locus was chosen as a test case for polyamide disruption of AR binding in the nuclei of live cells because of its particularly strong signal and large dynamic range between the non-induced and induced states. The original data obtained with 10  $\mu$ M compound **3** in DHT-stimulated LNCaP cells have been reproduced in LNCaP cells alongside the results obtained with 3  $\mu$ M compound **1** (Figure 4.3). However, in LN-AR cells, neither compound was able to block AR binding to this locus. LN-AR cells treated with 10  $\mu$ M of the third-generation, synthetic anti-androgen RD162 did behave as expected; AR bound to this compound is excluded from the nucleus and thus ARE



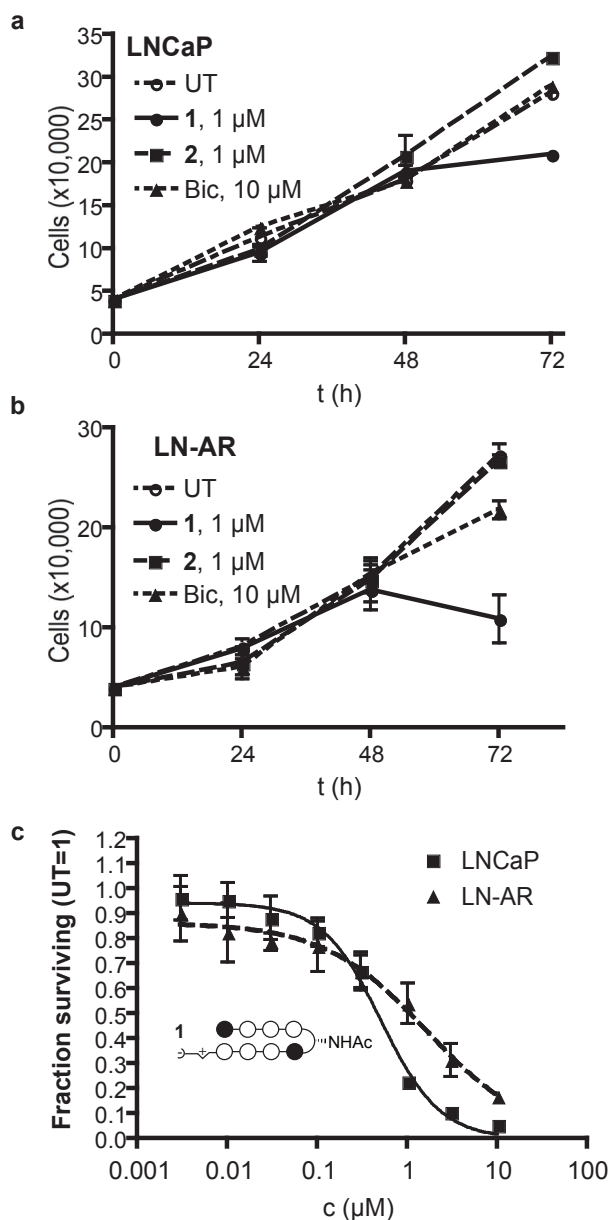
**Figure 4.3.** Inhibition of AR occupancy at the FKBP5 intronic enhancer. (a) Chromatin immunoprecipitation (ChIP) experiment in LNCaP cells showing inhibition of ARE occupancy by ARE-targeted polyamides **3** and **1**, as well as RD162. The data represent the mean and s.d. of three independent pulldowns. (b) AR ChIP in LN-AR fails to show any inhibition of AR occupancy by **3** or **1**. RD162 retains its activity. The data represent the mean of two independent pulldowns and are representative of 3 biological replicates.

occupancy is reduced to the non-induced state.

#### 4.2.3. Inhibition of AR-target gene expression is accompanied by cytotoxicity.

In addition to targeted compounds that interrupt AR signaling, cytotoxic DNA-binding compounds can also disrupt the AR axis (10). To look for polyamide-mediated cytotoxicity, we first performed a cell growth assay in both LNCaP and LN-AR cells with compounds **1**, **2**, at 1  $\mu\text{M}$  and bicalutamide at 10  $\mu\text{M}$  concentration (Figure 4.4A). Growth inhibition caused by **1** became apparent after 72 hours continuous treatment, as measured by manual hemocytometry.

The cell growth assay indicated that compound **1** was cytostatic at 1  $\mu\text{M}$  after 72 hours. To look for a cytotoxic effect and to define the cytotoxic concentration, we treated LNCaP and LN-AR cells with increasing concentrations of compound **1** for 72 and 96 hours (Figure 4.4B). The surviving cells were stained with sulfarhodamine B and quantitated by spectrophotometry. The compound showed dose-dependent cytotoxicity in



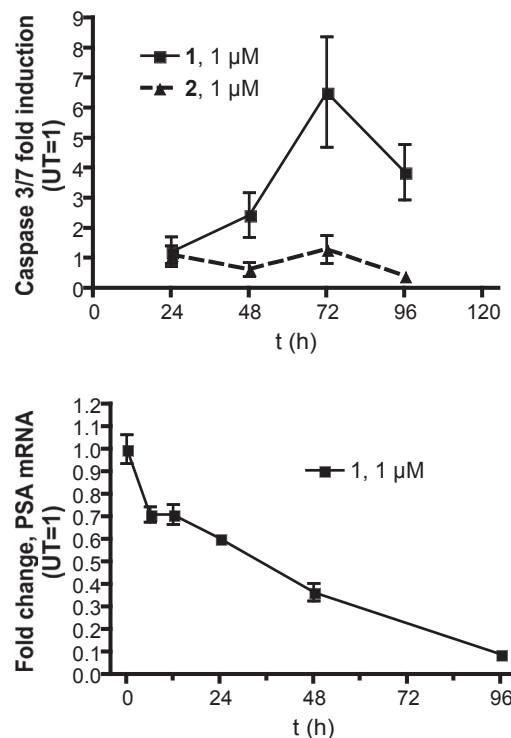
**Figure 4.4.** Inhibition of prostate cancer cell growth and cytotoxic response following treatment with ARE-targeted polyamide **1**. (a) Growth curve of LNCaP cells treated with either **1** (1  $\mu\text{M}$ ), **2** (1  $\mu\text{M}$ ), or bicalutamide (10  $\mu\text{M}$ ). Compound **1** displays a cytostatic effect at 72 h. (b) Growth curve of LN-AR cells treated with either **1** (1  $\mu\text{M}$ ), **2** (1  $\mu\text{M}$ ), or bicalutamide (10  $\mu\text{M}$ ). Compound **1** displays a slight cytotoxic effect at 72 h. (c) Cytotoxicity assay displaying dose-responsive cell death in response to **1** in LNCaP and LN-AR cells at 96 h. as measured by sulfarhodamine B staining.

both cell lines. The data were fit with non-linear least squares analysis to give the  $IC_{50}$  values shown in Table 4.2.

Cell line	AR	Cytotoxicity $IC_{50}$ values ( $\mu M$ )	
		72h	96h
LNCaP	+	$1.8 \pm 0.9$	$0.6 \pm 0.2$
LNAR	+++	$3 \pm 1$	$1.5 \pm 0.2$

**Table 4.2.** Cytotoxicity  $IC_{50}$  values in LNCaP and LN-AR cells in response to treatment with **1**.

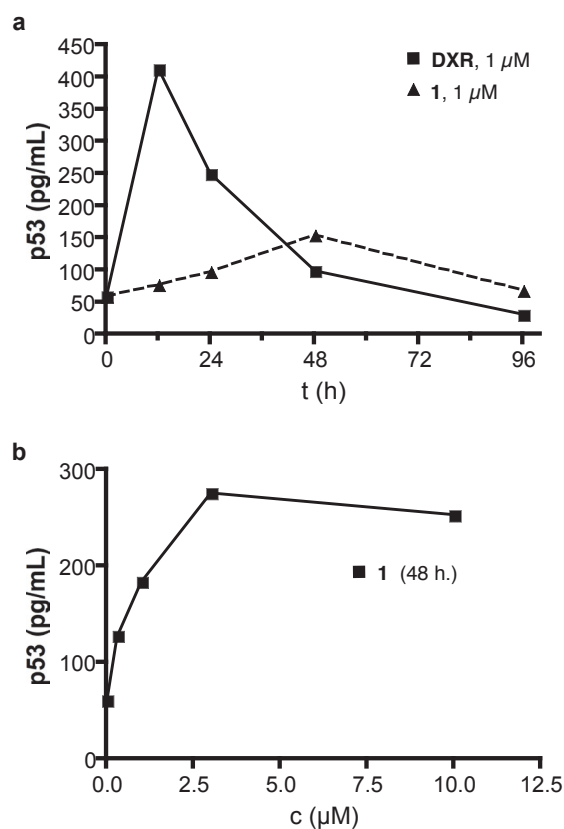
**4.2.4. Polyamide treatment of LN-AR cells activates cell death and cell stress response pathways on the same timescale as PSA down-regulation.** The cell death response to cytotoxic compounds generally follows either a necrotic or apoptotic pathway. To look for induction of apoptosis (11), we treated LN-AR cells with  $1 \mu M$  **1** or  $1 \mu M$  **2** for up to 96 hours using a luciferogenic substrate for Caspases 3 and 7 (Figure 4.5A). Cells treated with **1** showed a time-dependent increase in Caspase 3/7 activity starting at 48 hours. Untreated cells and cells treated with **2** showed no such induction.



**Figure 4.5.** Caspase 3/7 activation accompanies PSA downregulation in unstimulated LN-AR cells. (a) Time course of Caspase 3/7 fold induction (UT=1) following treatment with  $1 \mu M$  match polyamide **1** or  $1 \mu M$  mismatch polyamide **2** as measured by cleavage of a luciferogenic substrate for Caspases 3 and 7. Activity is first detectable at 48 h. and peaks at 72 h. (b) Time course of fold changes (UT=1) in PSA mRNA levels as measured by RT-PCR. Inhibition of basal PSA expression is detectable after 6 h of treatment and continues to decline for the entire time course.

To examine the effects of polyamide inhibition of PSA mRNA expression on the same time course as the caspase induction, we treated unstimulated LN-AR cells in rich media with 1  $\mu\text{M}$  **1** and measured mRNA levels at intervals up to 96 hours (Figure 4.5B). This assay measures inhibition of PSA from the steady state; cell culture media contains dilute concentrations of steroid hormones, so there is some basal expression of AR-target genes. In this assay, we observed time-dependent inhibition of PSA expression. Inhibition was detectable even after 6 hours of polyamide treatment and expression continued to decline until the 96-hour time point. Measurement of PSA inhibition in DHT-stimulated cells was performed after 64 hours incubation with the polyamide, a timepoint at which unstimulated, polyamide-treated cells showed significant Caspase 3/7 activity.

The tumor suppressor protein p53 can be activated by a wide array of stressors, including DNA damage and replicative or oxidative stress (12). In response to stress, p53 accumulates as it is rescued from ubiquitination and degradation. To look for accumulation of p53 in response to polyamide treatment, we treated LN-AR cells with 1  $\mu\text{M}$  **1** for up to 96 hours and assayed total cellular lysates for accumulation p53 by sandwich ELISA (Figure 4.6). A standard curve using purified protein was created for this and the subsequent



**Figure 4.6.** Stabilization of p53 in response to polyamide treatment. (a) Time course of p53 stabilization in LN-AR cells treated with 1  $\mu\text{M}$  doxorubicin (DXR) or 1  $\mu\text{M}$  **1** as measured by sandwich ELISA. Levels of p53 peak at 48 h. in cells treated with polyamide **1**. (b) Dose-response of p53 stabilization in LN-AR cells treated with increasing concentrations of **1** as measured by sandwich ELISA. Levels of p53 peaked at ~6-fold induction at 3  $\mu\text{M}$  **1**.

experiment. Cells treated with  $1\ \mu\text{M}$  doxorubicin, a DNA-damaging agent known to induce p53 stabilization, were included as a positive control. The polyamide-treated cells showed a time-dependent induction of p53 that peaked at 48 hours. The induction was weak when compared to that produced by doxorubicin. To look for a dose-dependent effect, LN-AR cells were treated with increasing concentrations of **1** for 48 hours and then subjected to the same ELISA. The induction of p53 was dose dependent and appeared to plateau at  $3\ \mu\text{M}$ . Maximal induction was about 4-fold over the untreated condition ( $t = 0$ ).



### 4.3. Discussion

In the hormone-resistant LN-AR cells, the ARE-targeted polyamide **1** maintains potency against PSA mRNA expression (Figure 4.1) but fails to decrease the binding of AR to the ARE in the FKBP5 intronic enhancer (Figure 4.2). The original polyamide antagonist, compound **1** also fails to inhibit AR binding to this locus in LN-AR cells. The LN-AR cell line is an LNCaP daughter cell line that contains a stably integrated construct that overexpresses AR. It is unclear how this single change would prevent polyamide-mediated inhibition of AR binding without disrupting the compound's ability to inhibit AR-driven gene expression.

It is possible that the polyamide and AR can co-occupy the FKBP5 intronic ARE when AR expression levels are raised to several times wild-type levels. Since AR binds in the major groove and the polyamide in the minor groove, AR inhibition by this compound is presumed to be allosteric. The transactivating function of AR is known to be allosterically sensitive, and there is precedent for a DNA-bound, transcriptionally inactive state of AR (13). Exploration of polyamide and receptor co-occupancy would be greatly facilitated by positive confirmation of polyamide binding to specific loci; the development of polyamide ChIP-seq methods is currently underway.

This study also revealed that treatment with **1** is accompanied by cytotoxicity, activation of the programmed cell death effector Caspases 3 and 7 (Figures 4.3 and 4.4). Further investigation of these findings should include an examination of their AR dependence, as hormone withdrawal is also known to induce an apoptotic cell death response in LNCaP cells (14). PC-3 and DU145 are prostate cancer cell lines that express very low levels of AR and display androgen-independent growth (15); these would make excellent model systems for further experimentation.

An additional consideration is that the observed effects on AR-driven gene expression could be a result of the cytotoxic insult. In particular, the data suggest a role for

p53 in the observed downregulation of AR-target gene expression. When p53 is stabilized in LNCaP cells, it acts as a repressor of AR homodimerization and suppresses expression of the AR gene(16,17). Additionally, it can bind a response element in the PSA promoter and repress its transcription (18). Experiments with polyamides **1** and **3** with p53 knockdown constructs in LNCaP cells could define the dependence of the observed transcriptional inhibition and cytotoxicity on p53 signaling.

#### 4.4. Materials and methods

**4.4.1. Compound synthesis.** Polyamides **1**, **2**, and **3** were synthesized by solid-phase methods on Kaiser oxime resin (Nova Biochem, Darmstadt, Germany) according to established protocols (19). Polyamides were cleaved from resin with 3,3'-diamino-*N*-methyl-dipropylamine and purified by reverse-phase HPLC. Isophthalic acid was activated with PyBOP (Nova Biochem) and conjugated to the polyamides as described. Purities and identities of the polyamides were assessed by HPLC, UV-visible spectroscopy, and MALDI-TOF MS. RD162 was synthesized by microwave-assisted solution-phase synthesis according to previously published protocols (9). Purity and identity of RD162 were assessed with <sup>1</sup>H NMR and ESI-MS. Doxorubicin was purchased from Sigma-Aldrich and used without further purification, as were all other compounds and supplies unless otherwise specified.

**4.4.2. Cell culture.** LNCaP and LN-AR cells were maintained in RPMI 1640 with 10% FBS at 37°C under 5% CO<sub>2</sub>. LNCaP cells were purchased from ATCC (Manassas, VA). LN-AR cells were a gift from C.L. Sawyers at Memorial Sloan-Kettering Cancer Center (NY, NY).

**4.4.3. Measurement of PSA mRNA and Protein.** 40,000 LNCaP or LN-AR cells were plated in RPMI1640 (Invitrogen) supplemented with 10% FBS (Irvine Scientific, Santa Ana, CA). For DHT-induced experiments, the medium was replaced with RPMI medium 1640 containing 10% charcoal-stripped FBS after 72 h with or without polyamides at the designated concentrations. Cells were grown for an additional 48 h and then treated with 1 nM DHT for 16 h. When appropriate, bicalutamide was added 2 h before DHT stimulation. For assays of basal PSA expression in normal media, 40,000-80,000 cells were plated in 24-well plates and compounds were added in fresh, complete media after 24 h. Isolation of

RNA and cDNA synthesis was performed as described (20). Quantitative real-time RT-PCR was performed with SYBR Green PCR Master Mix (Applied Biosystems, Foster City, CA) on an ABI 7300 instrument. PSA mRNA was measured relative to  $\beta$ -glucuronidase as an endogenous control. Primer sequences are available upon request. Cell-culture supernatants were collected for an ELISA (R & D Systems, Minneapolis, MN) to measure PSA protein according to the manufacturer's protocol.

**4.4.4. Chromatin immunoprecipitation.** ChIP analysis of AR occupancy at the FKBP5 intronic ARE was performed as described using enrichment at a negative locus for normalization (21). Briefly,  $80 \times 10^6$  LNCaP or LN-AR cells were plated in RPMI1640 containing 10% charcoal-treated FBS. Polyamides were added after 24 h and allowed to incubate for an additional 48 h. RD162 was added 2 h prior to induction with DHT. DHT was added and the cells were harvested after 4 h. Workup and analysis of fold enrichment were performed as described, using a negative locus for quantitation of immunoprecipitated DNA using qPCR.

**4.4.5. Growth inhibition assay.** 40,000 LNCaP or LN-AR cells were plated in triplicate in 24-well plates. Compounds and controls were added after 24 hours pre-incubation. At each time point, cells were trypsinized, pelleted by centrifugation and resuspended in 1 mL complete media. Cells were hand counted with a hemocytometer. The data represent the mean and standard deviation of these technical replicates.

**4.4.6. Cytotoxicity assays.**  $IC_{50}$  values for cytotoxicity were determined using a previously described, sulfarhodamine-based, colorimetric assay for cellular protein content in 96-well microplates (22). LNCaP or LNAR cells were plated at 3,000 or 4,000 cells per well for the 72 h and 96 h timepoints, respectively. Compounds were added in 100  $\mu$ L RPMI1640 supplemented with 10% FBS 24 h after plating. Quadruplicate wells were used for each

concentration. At the appropriate time, the cells were fixed with 100  $\mu$ L 10% trichloroacetic acid solution, washed, stained, and dried as described elsewhere.(22) After solubilization of the bound dye in 10 mM Tris, the absorbance was measured at 490 nm.

The data are charted as a percentage of untreated controls, corrected for background absorbance.  $IC_{50}$  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  $H = \text{Hill Slope}$ . Three independent trials were averaged; stated  $IC_{50}$  values represent the mean and standard deviation. These calculations were performed using Prism 4 (GraphPad) software.

**4.4.7. Apoptosis assays.** LN-AR cells were plated in 96-well microplates at 4,000 cells per well. As above, compounds and controls were added 24h after plating. Each time point was assayed in triplicate. At harvest, Caspase 3/7 activity was assessed using 100  $\mu$ L of Caspase-Glo reagent (Promega), which contains the proluminescent caspase substrate DEVD-aminoluciferin. Luminescence was measured after 30 minutes incubation at room temperature. Luminescence data are expressed as the mean and standard deviation of technical replicates in arbitrary units.

**4.4.8. P53 stabilization assay.** 400,000–800,000 LN-AR cells were plated in 10 cm diameter dishes. Compounds were added after 24h and were allowed to incubate for various times up to 72 hours. At harvest, cells were washed once with PBS then treated with 1 mL ice-cold lysis buffer and collected with a rubber policeman. After 15 min incubation at 5°C, lysates were centrifuged at 4,000 x g, reserving the supernatants. Protein concentrations were determined by Bradford assay (Promega). P53 protein was assayed by sandwich ELISA (R&D Biosystems) and performed according to the manufacturer's recommendations. 10  $\mu$ g total protein was used for each experimental sample. A standard curve using the provided purified p53 protein was created for each assay run, and p53 levels

for each experimental sample were determined by interpolation. The data are expressed as the mean of two technical replicates.

#### 4.5. References.

1. Huggins, C., Stevens, R.E. and Hodges, C.V. (1941) Studies on prostatic cancer: II. The effects of castration on advanced carcinoma of the prostate gland. *Archives of Surgery*, **43**, 169-185.
2. Furr, B.J.A., Valcaccia, B., Curry, B., Woodburn, J.R., Chesterson, G. and Tucker, H. (1987) Ici-176,334 - a Novel Nonsteroidal, Peripherally Selective Antiandrogen. *Journal of Endocrinology*, **113**, R7-R9.
3. Hodgson, M.C., Astapova, I., Hollenberg, A.N. and Balk, S.P. (2007) Activity of androgen receptor antagonist bicalutamide in prostate cancer cells is independent of NCoR and SMRT corepressors. *Cancer Res*, **67**, 8388-8395.
4. Chen, Y., Sawyers, C.L. and Scher, H.I. (2008) Targeting the androgen receptor pathway in prostate cancer. *Curr Opin Pharmacol*, **8**, 440-448.
5. Chen, C.D., Welsbie, D.S., Tran, C., Baek, S.H., Chen, R., Vessella, R., Rosenfeld, M.G. and Sawyers, C.L. (2004) Molecular determinants of resistance to antiandrogen therapy. *Nat Med*, **10**, 33-39.
6. Nickols, N.G. and Dervan, P.B. (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci U S A*, **104**, 10418-10423.
7. Hsu, C.F., Phillips, J.W., Trauger, J.W., Farkas, M.E., Belitsky, J.M., Heckel, A., Olenyuk, B.Z., Puckett, J.W., Wang, C.C. and Dervan, P.B. (2007) Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron*, **63**, 6146-6151.
8. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol*, **13**, 284-299.
9. Tran, C., Ouk, S., Clegg, N.J., Chen, Y., Watson, P.A., Arora, V., Wongvipat, J., Smith-Jones, P.M., Yoo, D., Kwon, A. *et al.* (2009) Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science*, **324**, 787-790.
10. Mantoni, T.S., Reid, G. and Garrett, M.D. (2006) Androgen receptor activity is inhibited in response to genotoxic agents in a p53-independent manner. *Oncogene*, **25**, 3139-3149.
11. Tewari, M., Quan, L.T., O'Rourke, K., Desnoyers, S., Zeng, Z., Beidler, D.R., Poirier, G.G., Salvesen, G.S. and Dixit, V.M. (1995) Yama/CPP32 beta, a mammalian homolog of CED-3, is a CrmA-inhibitable protease that cleaves the death substrate poly(ADP-ribose) polymerase. *Cell*, **81**, 801-809.
12. Lavin, M.F. and Gueven, N. (2006) The complexity of p53 stabilization and activation. *Cell Death Differ*, **13**, 941-950.
13. Masiello, D., Cheng, S., Buble, G.J., Lu, M.L. and Balk, S.P. (2002) Bicalutamide functions as an androgen receptor antagonist by assembly of a transcriptionally inactive receptor. *J Biol Chem*, **277**, 26321-26326.
14. Lee, E.C., Zhan, P., Schallhom, R., Packman, K. and Tenniswood, M. (2003) Antiandrogen-induced cell death in LNCaP human prostate cancer cells. *Cell Death Differ*, **10**, 761-771.
15. Alimirah, F., Chen, J., Basrawala, Z., Xin, H. and Choubey, D. (2006) DU-145 and

- PC-3 human prostate cancer cell lines express androgen receptor: implications for the androgen receptor functions and regulation. *FEBS Lett*, **580**, 2294-2300.
16. Shenk, J.L., Fisher, C.J., Chen, S.Y., Zhou, X.F., Tillman, K. and Shemshedini, L. (2001) p53 represses androgen-induced transactivation of prostate-specific antigen by disrupting hAR amino- to carboxyl-terminal interaction. *J Biol Chem*, **276**, 38472-38479.
  17. Alimirah, F., Panchanathan, R., Chen, J., Zhang, X., Ho, S.M. and Choubey, D. (2007) Expression of androgen receptor is negatively regulated by p53. *Neoplasia*, **9**, 1152-1159.
  18. Gurova, K.V., Roklin, O.W., Krivokrysenko, V.I., Chumakov, P.M., Cohen, M.B., Feinstein, E. and Gudkov, A.V. (2002) Expression of prostate specific antigen (PSA) is negatively regulated by p53. *Oncogene*, **21**, 153-157.
  19. Belitsky, J.M., Nguyen, D.H., Wurtz, N.R. and Dervan, P.B. (2002) Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg Med Chem*, **10**, 2767-2774.
  20. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G., Jr. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*, **101**, 16768-16773.
  21. Puckett, J.W. (2009) Dissertation (Ph.D.), California Institute of Technology, Pasadena.
  22. Vichai, V. and Kirtikara, K. (2006) Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat Protoc*, **1**, 1112-1116.



## **Chapter 5**

### **Py-Im polyamides inhibit DNA Topoisomerase II activity in vitro by disrupting enzyme-DNA binding**

John W. Phillips and Peter B. Dervan

**Abstract**

We have developed 8-ring, hairpin, Py-Im polyamides that downregulate the expression of Androgen Receptor (AR) target genes in LNCaP cells. These compounds are cell-permeable, sequence-specific, high-affinity DNA minor groove binders that are designed to prefer the sequence 5'-WGWWCW-3' (W=A or T). In this study, we investigate the interaction of these two compounds with topoisomerase II (Top2) isozymes using *in vitro* and cell culture techniques. We determine the polyamides to inhibit Top2 catalytic activity by preventing enzyme-DNA binding. However, the DU145 cell culture results in Top2 isozyme knockdown constructs are more consistent with Top2 poisons. Further investigation using flow cytometric analysis of cell cycle distribution indicates that polyamide treatment results in S-phase arrest, a finding that is not associated with Top2 inhibitors or Top2 poisons. The possibility that polyamides are acting as DNA-damage inducing agents or as DNA synthesis inhibitors is discussed.

## 5.1. Introduction

Pyrrole-imidazole polyamides are programmable, sequence-specific, high-affinity DNA minor groove binders (1). Using pairwise configurations of N-methylpyrrole (Im) and N-methylimidazole (Im) amino acid monomers arranged in antiparallel fashion in a hairpin configuration, they can be designed to bind specifically to a wide array of DNA sequences using a simple set of pairing rules (2). An Im-Py pair recognizes G-C, a Py-Im pair recognizes C-G, and a Py-Py pair is degenerate for A-T or T-A (W) (3). Due to steric constraints, hairpin polyamides have additional binding preferences for an A-T or T-A base pair at either side of the core set of heterocycles; therefore an 8-ring hairpin polyamide specifies a 6 base pair binding site (4). In addition to their DNA binding properties, polyamides permeate cells, localize to the nucleus, and bind chromatin without the need for delivery agents (5,6).

When combined with affinity and sequence specificity approaching that of endogenous transcription factors, these favorable biocompatibility properties have made it possible to use polyamides as tools for establishing chemical control of gene expression. Polyamides designed to target the consensus DNA-binding sequences of endogenous transcription factors have been used to disrupt the genotropic actions of those proteins. By blocking transcription factor-response element binding, polyamides can abrogate activation of target gene expression. This technology has been used successfully in cancer cell tissue culture to target the transcription factors HIF1 $\alpha$ , androgen receptor (AR), glucocorticoid receptor (GR), and AP-1 (7-10). In each case, genomic occupancy of the target transcription factor at relevant loci was decreased approximately 2-fold as measured by chromatin immunoprecipitation (ChIP) when cells were treated with an appropriately targeted polyamide.

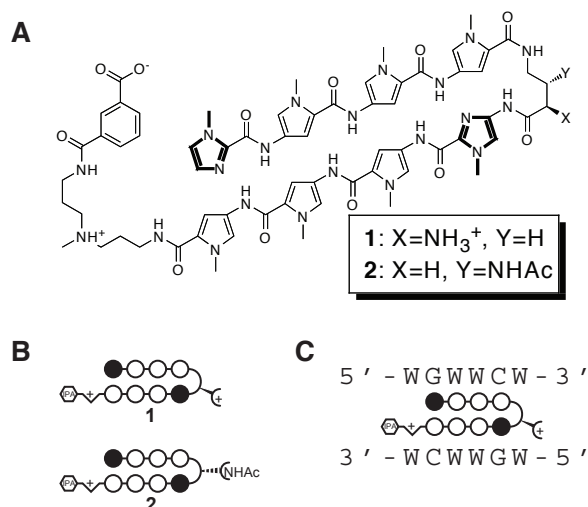
As DNA minor groove binders, polyamides would be predicted to have a wide range of effects on DNA-dependent cellular processes (11). In addition to their effects on

transcription factor binding and inhibition of target gene induction, interaction with other biological processes could be contributing to the observed changes in gene expression (12). Minor groove binders are known to inhibit the DNA topoisomerases Top1 and Top2, as well as DNA helicases (13-18). Inhibition of any one of these enzymes could contribute to non-specific transcriptional inhibition. Yet data from cDNA microarray analyses of polyamide-treated cells suggest that the resultant mRNA inhibition is not a simple global downregulation (7,8,19). In fact, the magnitude and number of transcriptional changes produced by the polyamide used to antagonize AR and GR activity are similar to those produced by other small molecule, non-DNA binding antagonists of those pathways.

Top1 and the two isozymes of Top2, Top2 $\alpha$  and Top2 $\beta$ , are known to be intimately involved in gene transcription in mammalian cells (20). The strand separation required to accommodate the RNA Polymerase II holoenzyme produces superhelical strain in the form of positive supercoiling in advance of the transcription bubble (21). Top1 and Top2 $\alpha$  are associated with this complex and are required to enable it to advance (22); inhibition of either of these two enzymes would be expected to produce a global decrease in transcription. The isozymes of Top2 are highly homologous and differ chiefly in their regulation; Top2 $\alpha$  expression is correlated with the cell cycle, Top2 $\beta$  is not (23). Yet inhibition of Top2 $\beta$  activity would produce a more limited response, as activity of this isozyme is only required for expression of a subset of regulated genes (24). This subset corresponds to the target genes of several nuclear hormone receptors, including AR and GR (24-26). The interaction of polyamides with the two Top2 isoenzymes is the chief subject of this study.

To define the interaction of polyamides with Top2, we chose to study two AR-targeted polyamides, compounds **1** and **2** (Figure 5.1). Using *in vitro* relaxation assays with a purified Top2 $\alpha$  fragment, we demonstrate that both compounds displayed dose-dependent inhibition of topoisomerase activity without forming a cleavable complex. Further analysis with a titration of increasing amounts of Top2 $\alpha$  suggested that the inhibition was due to blocking Top2 binding to DNA. Cell culture assays using a Top2 $\alpha$  shRNA-

knockdown cell line conferred significant resistance to polyamide-induced cytotoxicity. We then investigated the effects of polyamide treatment on the cell cycle to look for the characteristic G<sub>2</sub>/M arrest produced by Top2 inhibitors and Top2 poisons. Unexpectedly, we observed S-phase arrest in both the wild-type and Top2 $\alpha$  knockdown cell line, which suggests that the primary effects of polyamides are unrelated to Top2.



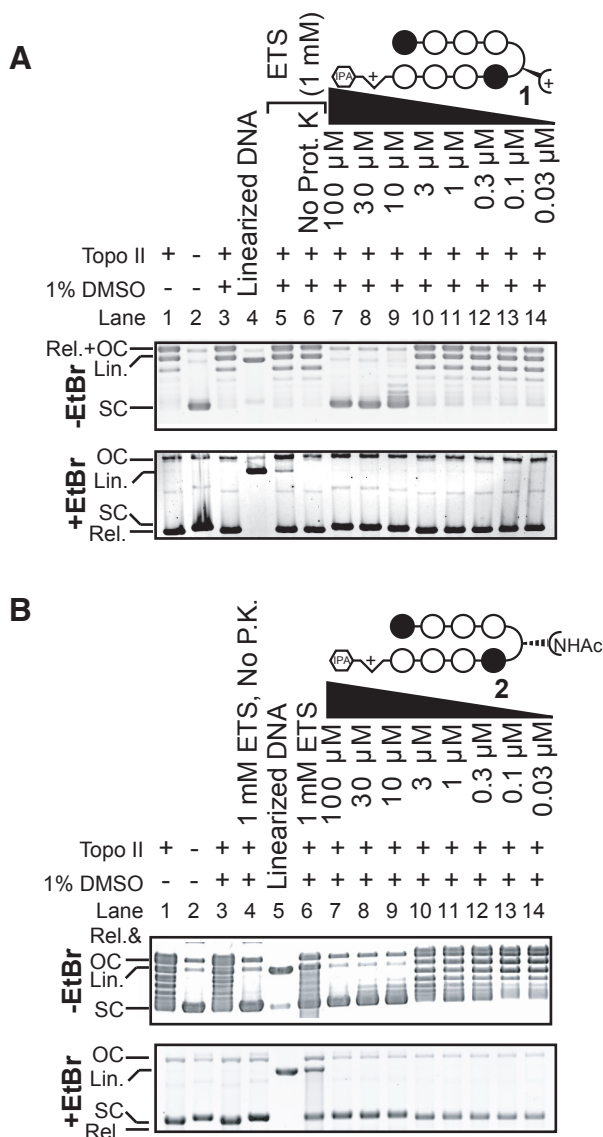
**Figure 5.1.** Chemical structure and binding preferences of the Py-Im polyamides used in this study. (A) Chemical structure showing the two different turn functionalities. (B) Ball-and-stick representation of the polyamides. Open circles represent N-methylpyrrole residues, filled circles represent N-methylimidazoles. The hexagon represents the isophthalic acid moiety. (C) Both 1 and 2 bind the same 5'-WGWWCW-3' DNA sequence, where W = A or T.

## 5.2. Results

### 5.2.1. Polyamides inhibit Top2 catalysis by preventing the enzyme from binding DNA.

Top2-targeted small molecules generally fall into two categories: inhibitors of enzymatic activity and topoisomerase poisons (27). Inhibitors prevent the enzyme from catalyzing the DNA strand breakage, either by inhibiting enzyme binding or by inhibiting the ATP hydrolysis Top2 requires to perform its catalytic cycle. They do not produce DNA damage. Topoisomerase poisons trap the enzyme in a covalent complex with the DNA, stabilizing the otherwise transient double-strand DNA break. Further interactions of this complex with cellular proteins quickly result in DNA damage. Minor groove DNA-binding compounds are predicted to inhibit topoisomerases by preventing enzyme binding.

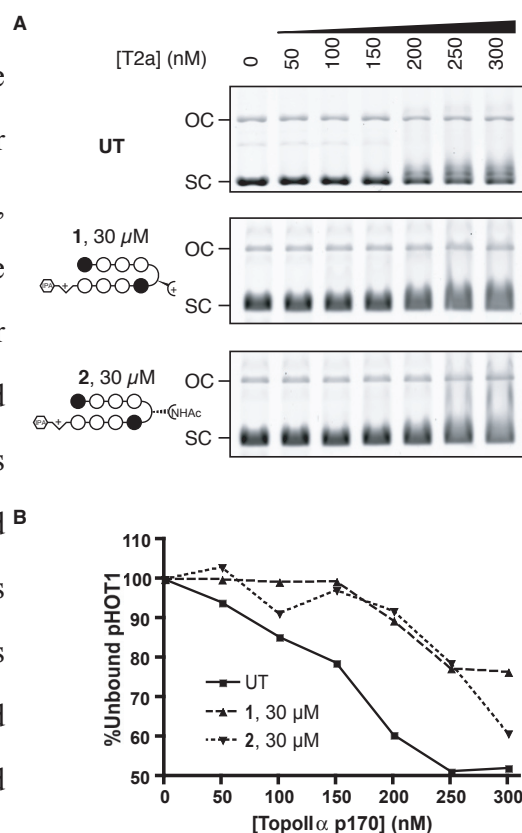
To look for polyamide-mediated inhibition of Top2, we conducted *in vitro* relaxation assays using supercoiled pHOT1 plasmid DNA and a purified Top2 $\alpha$  fragment (Figure 5.2). The



**Figure 5.2.** *In vitro* DNA relaxation assay demonstrating polyamide-mediated, dose-dependent inhibition of Top2 $\alpha$ -p170 catalytic activity without cleavage complex formation. Results of agarose gel electrophoretic analysis of supercoiled pHOT1 plasmid DNA relaxation by Top2 $\alpha$ -p170 run in the absence (top) or presence (bottom) of 0.5  $\mu$ g/mL ethidium bromide (EtBr). (A) Results obtained by adding increasing concentrations of compound 1 (0.03-100  $\mu$ M) to relaxation reactions. (B) Results obtained by adding increasing concentrations of compound 2 (0.03-100  $\mu$ M) to relaxation reactions.

reaction products were analyzed by agarose gel electrophoresis after quenching, proteinase digestion, and phenol-chloroform extraction. To look for inhibition of relaxation, the gels were run in the absence of ethidium bromide (EtBr) (Figure 5.2a and 5.2b, upper panel) and post-stained with SYBR-Gold; supercoiled substrates have a smaller relative size and migrate faster on the gel. The linearized, relaxed, and open-circular forms of the plasmid have similar apparent size and tend to migrate together (compare Lanes 3 and 4, Figure 5.2a, upper panel). Etoposide was included as a positive control for cleavage complex formation (Figure 5.2a, Lane 5; Figure 5.2b, Lane 6, upper panel). In the presence of ATP, both polyamides inhibited plasmid relaxation in a dose-dependent manner (Figure 5.2a and 5.2b, upper panel, lanes 7-14).

Samples from these reactions were also run in gels containing EtBr to look for formation of a cleavage product. In these gels, the unwinding effect of the intercalative dye dominates the topology; all closed circular forms of DNA will be positively supercoiled and will migrate with similar  $R_f$ . This allows unambiguous identification of the linearized DNA that forms when a small molecule causes Top2 cleavage complex formation (compares Lanes 3 and 4, Figure 5.2a). Etoposide-treated samples and linearized DNA were included as positive controls for cleavage complex formation (lane 5, Figure 5.2a, lower panel and lane 6, Figure 5.2b, lower panel). An etoposide-treated, proteinase K-free sample was also included to demonstrate detection of



**Figure 5.3.** Polyamides 1 & 2 inhibit Top2 $\alpha$ -p170 binding *in vitro*. (A) Gel electrophoretic mobility shift assay for Top2 $\alpha$ -p170 DNA-binding to pHOT1 in the presence or absence of 30  $\mu$ M 1 or 30  $\mu$ M 2. The amount of free DNA decreases as the protein is titrated up to 300 nM. No ATP is present in this assay. (B) Top2 $\alpha$ -p170 DNA-binding curves based on relative quantitation of (A).

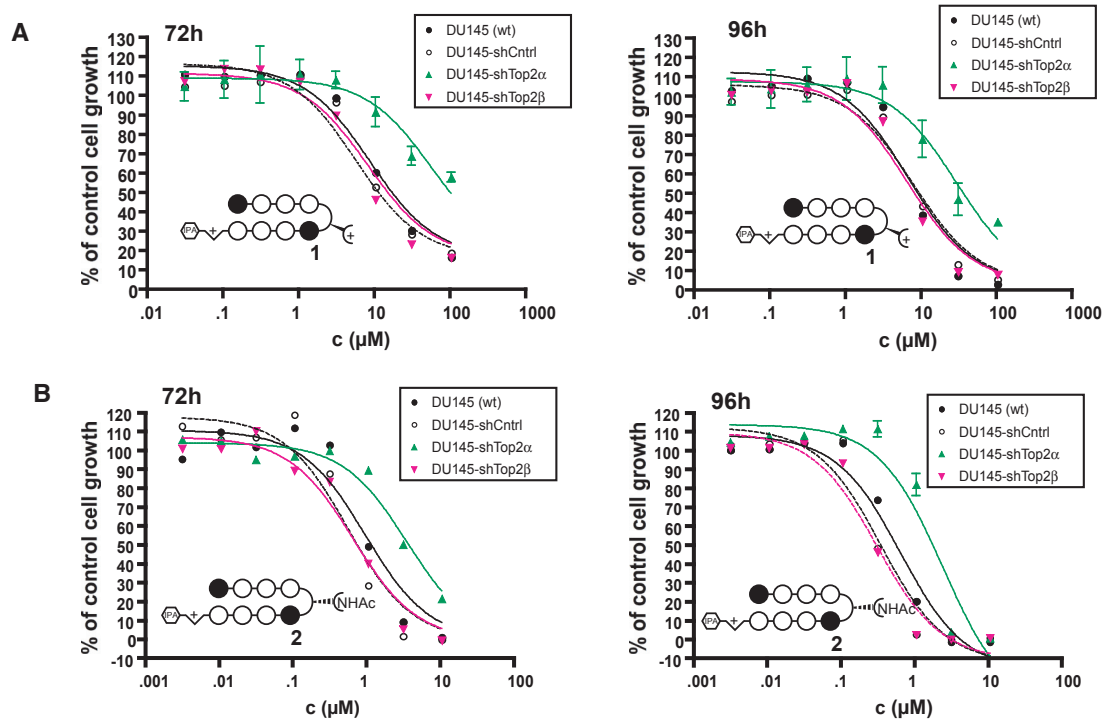
the covalent complex. Failure to digest the protein resulted in extraction of the covalent complex into the phenol-chloroform layer (compare lanes 5 and 6, Figure 5.2a, lower panel). Neither polyamide caused cleavage complex formation over a wide range of concentrations (Lanes 7-14, Figures 5.2a and Figure 5.2b, lower panel).

Enzymatic inhibition in the absence of cleavage complex formation is consistent with a mechanism based on disruption of Top2 binding. To address this mechanism more directly, we treated samples of pHOT1 plasmid DNA with 30  $\mu$ M concentrations of **1** or **2** for 30 minutes and then added increasing amounts of Top2 $\alpha$  in the absence of ATP (Figure 5.3). Enzyme-DNA binding was measured by gel-shift assay. Treatment with 30  $\mu$ M of either compound increased the amount of free DNA per microgram of Top2 $\alpha$  added. Compound **1** was more potent at inhibiting Top2-DNA binding than compound **2**.

**5.2.2. Topo2a knockdown confers apparent resistance to polyamide-mediated cytotoxicity.** The expression level of Top2 isozymes is known to modulate the cytotoxicity of agents targeted to these enzymes (28,29). As polyamides inhibit Top2 activity *in vitro* by inhibiting enzyme binding, decreasing Top2 expression would be expected to sensitize cells to their cytotoxic effects. On the other hand, if polyamides were acting as Top2 poisons despite the *in vitro* data, Top2 knockdown would be expected to confer resistance to their cytotoxicity, as decreased Top2 expression would produce fewer of the lethal DNA double-strand breaks that result from cleavage complexes.

To look for a correlation between Top2 isozyme expression and polyamide cytotoxicity, we determined cytotoxicity IC<sub>50</sub> values for compounds **1** and **2** in DU145 prostate cancer cells and three derivative cell lines containing stably incorporated shRNA constructs: DU145-shT2 $\alpha$ , DU145-shT2 $\beta$ , and DU145-shCntrl. The first two contain Top2 $\alpha$ - and Top2 $\beta$ -targeted shRNAs, and the third is included as a scrambled vector control (30). The polyamides demonstrated dose-dependent inhibition of cell growth in each cell line (Figure 5.4a-b). The data also show that Top2 $\alpha$  knockdown confers approximately





**Figure 5.4.** Dose-dependent induction of cytotoxicity by polyamides **1** and **2** in DU145 (wt) and Top2 knockdown cell lines. Dose-response curves generated by sulfarhodamine B staining of DU145 (filled circles), DU145-shCntrl (open circles), DU145-shTop2 $\alpha$  (green triangles), or DU145-shTop2 $\beta$  cells (pink triangles) treated with **1** (A) or **2** (B) for 72 or 96 h.

5-fold resistance to polyamide treatment for cells treated for 72 or 96 h (Table 5.1), which is similar to the resistance conferred on known cleavage-complex producing agents like etoposide (30). These results would appear to contradict the *in vitro* data that suggest that polyamides inhibit enzyme-DNA binding and do not cause cleavage complex formation.

### 5.2.3. Polyamide treatment causes S-phase arrest. Top2 inhibitors and Top2 poisons are

Compound	Time (h)	DU145 (wt)	DU145-shCntrl	DU145-shTop2 $\alpha$	DU145-shTop2 $\beta$
<b>1</b>	72	14 $\pm$ 4	13 $\pm$ 5	90 $\pm$ 30 ( <b>6</b> )	9 $\pm$ 1
	96	8 $\pm$ 4	10 $\pm$ 5	40 $\pm$ 10 ( <b>5</b> )	7 $\pm$ 2
<b>2</b>	72	1.5 $\pm$ 0.2	1.1 $\pm$ 0.1	8 $\pm$ 1 ( <b>5</b> )	1.1 $\pm$ 0.1
	96	0.76 $\pm$ 0.06	0.7 $\pm$ 0.1	2.5 $\pm$ 0.6 ( <b>3</b> )	0.6 $\pm$ 0.1

**Table 5.1.** Cytotoxicity IC<sub>50</sub> values ( $\mu$ M) of compounds **1** and **2** in DU145 and DU145-shTop2 cell lines. The data are determined from non-linear least squares analysis and are shown as the mean  $\pm$  s.d. of three independent biological replicates. The fold resistance conferred to polyamide-mediated cytotoxicity by Top2 $\alpha$  knockdown is shown in bold.

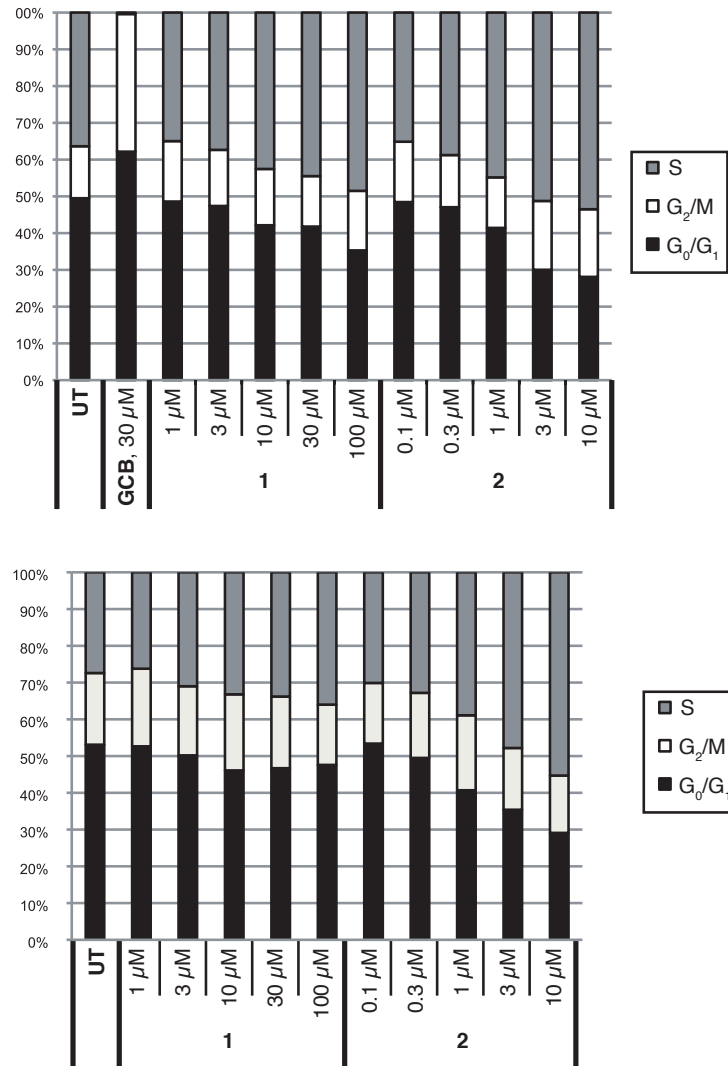
known to produce cell cycle arrest in G<sub>2</sub>/M. The former inhibit decatenation of replicated chromosomes during mitosis, and the latter activate DNA damage repair pathways at the G<sub>2</sub>/M checkpoint (31). To look for phenotypic evidence consistent with polyamide-mediated inhibition of Top2 suggested by the *in vitro* studies and the cytotoxicity results, we

performed a flow cytometric cell cycle assay. We subjected DU145 cells to treatment with various concentrations of **1** and **2** for 24 hours before pulse-labeling them with EdU for absolute S-phase analysis by flow cytometry (Figure 5.5a).

Fixed cells were labeled for EdU incorporation using an AlexaFluor dye and then co-stained for total DNA content with 7-aminoactinomycin D to display the entire cell cycle.

Contrary to our expectations, both compounds produced a dose-dependent increase in the number of cells in S-phase, not G<sub>2</sub>/M. Gemcitabine, a DNA synthesis inhibitor, was included as a positive control for S-phase disturbance.

We then performed a



**Figure 4.5.** Polyamide treatment causes S-phase arrest in DU145 and DU145-shTop2a cells. (A) Cell cycle distribution of DU145 cells treated with compounds or controls for 24h as measured by two-color flow cytometric evaluation of A488-EdU pulse-labeled cells stained for DNA content with 7AAD. Gemcitabine is included as a control for a known DNA synthesis inhibitor. Both compounds increase the percentage of cells in S-phase in a dose-dependent manner. (B) Analogous experiment conducted in DU145-shTop2a cells showing a lower number of S-phase cells in the untreated condition but a similar dose-dependent increase in the S-phase population for both compounds. 20,000 cells were counted for each sample. The experiment was conducted in duplicate.

similar experiment in the Top2 $\alpha$ -knockdown cell line that showed resistance to polyamide-induced cytotoxicity. The polyamides displayed similar dose-dependent induction of S-phase arrest to that observed in the parental cell line (Figure 5.5b). The disturbance was detected over the same concentration range. When compared to the results in the parental cell line, the percentage of DU145-shT2 $\alpha$  cells in S-phase was lower, but this was also true in the untreated controls. The percentage of cells in G<sub>2</sub>/M was higher than in the parental DU145 cells. This pattern is consistent with slowed transit through G<sub>2</sub>/M due to decreased Top2 $\alpha$  and the absolute requirement for this enzyme to perform decatenation prior to mitosis.

### 5.3. Discussion

The *in vitro* relaxation assay data appear to conflict with the results obtained in the Top2-isozyme knockdown cell lines. The former demonstrate that polyamides **1** and **2** inhibit Top2 activity by preventing enzyme-DNA binding, which agrees with the results observed for the related compound, distamycin A (32). The polyamides inhibit Top2 binding and enzymatic activity at micromolar concentrations in a rank order according to their relative affinities for DNA (Figure 5.3). Neither compound shows any evidence of creating the cleavable complex that leads to DNA double-strand breaks. On the other hand, the cytotoxicity  $IC_{50}$  values are consistent with a compound that does create Top2 $\alpha$ -mediated DNA damage (Figure 5.4), although this time the relative potencies are reversed (Table 5.1). This apparent contradiction prompted further investigation of the cellular response to polyamide treatment in an attempt to provide phenotypic evidence to support one of the two competing hypotheses.

Surprisingly, the results of the cell cycle investigation did not produce the G<sub>2</sub>/M-phase arrest that usually accompanies Top2 $\alpha$  inhibition or poisoning. Instead, they unambiguously demonstrate that compounds **1** and **2** produce S-phase arrest. The arrest is both dose dependent and relatively rapid compared to the 72 or 96 h timescale of the cytotoxicity results. The two compounds showed the same pattern of potency as in the cytotoxicity assays, with **2** being more potent and producing greater S-phase arrest than **1**. The observed cell cycle disruption due to polyamide treatment is not consistent with a Top2 $\alpha$  inhibitor or a Top2 $\alpha$  poison; S-phase arrest is characteristic of compounds that produce S-phase specific DNA damage or inhibition of DNA replication (33).

In addition, the cell cycle pattern observed in the untreated condition in the DU145-shT2 $\alpha$  cell line suggests an explanation for its relative resistance to polyamide-mediated cytotoxicity. This cell line showed a higher percentage of cells in G<sub>2</sub>/M, and a lower percentage in S-phase, which is likely due to decreased Top2 $\alpha$  availability for decatenation. If polyamides were indeed acting as S-phase-specific toxins, then having fewer cells in S-

phase at any given time would produce fewer cytotoxic interactions.

Taken together, these data suggest that Top2 inhibition is not a primary effect of polyamide treatment in cell culture. The *in vitro* data would appear to conflict with the cytotoxicity results in the knockdown models, and the S-phase interaction is not phenotypically consistent with Top2 inhibition. S-phase arrest is usually due to either S-phase specific DNA damage, or inhibition of DNA synthesis (33). It is possible that the polyamides are inhibiting Topoisomerase I, which is highly active during S-phase (20). Treatment with the Top1 poison camptothecin does cause DNA damage as the replication holoenzyme encounters the cleavable complex (27). But as non-reactive DNA minor groove binding ligands, polyamides would not be expected to form adducts with DNA. However, it is possible that cellular metabolism could transform the compound into a reactive species. The polyamides could also be directly impeding fork progression, perhaps by inhibiting a replicative helicase; distamycin is known to inhibit the DNA damage repair helicases WRN and BLM *in vitro* (34). Further investigation of the S-phase arrest produced by polyamide treatment should look for polyamide-induced DNA damage. Elucidation of the cell signaling responsible for producing the S-phase delay could also be informative; activation of the ATM/ATR kinases would signal checkpoint activation in response to DNA damage or replicative stress (33,35). If no DNA damage is detected, *in vitro* replication assays could provide further support to a helicase-based mechanism of DNA synthesis inhibition.

## 5.4. Materials and methods

**5.4.1. Chemicals and reagents.** Compounds **1** and **2** were synthesized on solid phase Kaiser oxime resin using previously published protocols (36). Cell culture media was purchased from Invitrogen, and fetal bovine serum from Irvine Scientific. Gemcitabine was purchased from AvaScientific. Etoposide was purchased from Sigma-Aldrich, as were all other reagents unless otherwise noted.

**5.4.2. *In vitro* DNA relaxation and Top2 $\alpha$ -DNA binding assays.** Materials for these assays were supplied by Topogen (Port Orange, FL). For relaxation assays, 540 ng Top2 $\alpha$ -p170 fragment (16 units) was added to 250 ng supercoiled pHOT1 DNA in assay buffer (0.05 M Tris-HCl (pH 8), 0.15 M NaCl, 10 mM MgCl<sub>2</sub>, 0.5 mM dithiothreitol) plus 2 mM ATP with or without test compounds in a total volume of 20  $\mu$ L. The dimethylsulfoxide (DMSO) concentration was standardized to 1% for all samples except the no-DMSO solvent controls. Reactions were incubated at 37 °C for 30 min. and then quenched with 2  $\mu$ L 10% sodium dodecyl sulfate solution. Samples were then extracted with chloroform: isoamyl alcohol 24:1, mixed with 2  $\mu$ L 10x glycerol loading buffer and loaded onto 1% agarose gels in tris-acetic acid-EDTA (TAE) buffer with or without 0.5  $\mu$ g/mL ethidium bromide (EtBr). Gels run without EtBr were post-stained with SYBR-Gold (Invitrogen). For Top2 $\alpha$ -DNA binding assays, increasing amounts of Top2 $\alpha$ -p170 fragment were added to aliquots of 0.325  $\mu$ g supercoiled pHOT1 plasmid DNA in assay buffer plus 2.5% glycerol with or without test compounds. Samples were incubated at 37 °C without ATP for 30 minutes and then loaded directly onto 1% agarose gels, run in TAE buffer, and then post-stained with 1x SYBR-Gold.

**5.4.3. Cell culture.** DU145 cells were purchased from ATCC (Manassas, VA) and maintained in RPMI 1640 with 10% FBS at 37 °C under 5% CO<sub>2</sub>. DU145-shCntrl, DU145-shT2 $\alpha$ ,

and DU145-shT2 $\beta$  cells were a gift from R. Dorr at Arizona Cancer Center (Tucson, AZ) and were maintained in RPMI1640/10% FBS supplemented with 0.5  $\mu$ g/ml puromycin to maintain selection for the knockdown constructs (30).

**5.4.4. Cytotoxicity assays.** IC<sub>50</sub> values for cytotoxicity were determined using a previously described, sulfarhodamine-based, colorimetric assay for cellular protein content in 96-well microplates (37). Cells were plated at 2,000 or 2,500 cells per well. Compounds were added in 100  $\mu$ L RPMI1640 supplemented with 10% FBS 24 h after plating. Quadruplicate wells were used for each concentration. At the appropriate time, the cells were fixed with 100  $\mu$ L 10% trichloroacetic acid solution, washed, stained, and dried as described. 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. IC<sub>50</sub> 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=max., B=min., and H=Hill Slope. Three independent trials were averaged; stated IC<sub>50</sub> values represent the mean and standard deviation. These calculations were performed using Prism 4 (GraphPad) software.

**5.4.5. Cell cycle analysis.** 800,000 cells were plated in 10 cm diameter dishes for 24 h before treatment with test compounds for an additional 24 h. 10  $\mu$ M EdU was added 30 min before harvest. The cells were harvested by trypsinization and combined with the cell culture supernatant before pelleting at 300 x g. Following overnight fixation in 70% ethanol, the cells were rehydrated in 1% BSA/PBS and processed with the Click-it EdU Alexa Fluor 488 Flow Cytometry assay kit (Invitrogen) using half the recommended A488 reagent. After overnight treatment with 0.2 mg/mL RNase A in 1% BSA/PBS, the cells were stained for DNA content with the provided 7-aminoactinomycin D and analyzed on

a FACSCalibur (Becton-Dickinson) instrument. The data were analyzed using FlowJo v8.8.2 (TreeStar) and are representative of two trials.



## 5.5. References.

1. Dickinson, L.A., Burnett, R., Melander, C., Edelson, B.S., Arora, P.S., Dervan, P.B. and Gottesfeld, J.M. (2006) Arresting cancer proliferation by small-molecule gene regulation. *Chem Biol*, **13**, 339-339.
2. 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. and Dervan, P.B. (2007) Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron*, **63**, 6146-6151.
3. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
4. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1996) Recognition of a 5'-(A,T)GGG(A,T)(2)-3' sequence in the minor groove of DNA by an eight-ring hairpin polyamide. *J Am Chem Soc*, **118**, 8198-8206.
5. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc Natl Acad Sci U S A*, **100**, 12063-12068.
6. Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B. and Luger, K. (2003) Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J Mol Biol*, **326**, 371-380.
7. Nickols, N.G. and Dervan, P.B. (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci U S A*, **104**, 10418-10423.
8. Muzikar, K.A., Nickols, N.G. and Dervan, P.B. (2009) Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc Natl Acad Sci U S A*, **106**, 16598-16603.
9. Wang, X., Nagase, H., Watanabe, T., Nobusue, H., Suzuki, T., Asami, Y., Shinojima, Y., Kawashima, H., Takagi, K., Mishra, R. *et al.* (2010) Inhibition of MMP-9 transcription and suppression of tumor metastasis by pyrrole-imidazole polyamide. *Cancer Sci*, **101**, 759-766.
10. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G., Jr. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*, **101**, 16768-16773.
11. Nelson, S.M., Ferguson, L.R. and Denny, W.A. (2007) Non-covalent ligand/DNA interactions: minor groove binding agents. *Mutat Res*, **623**, 24-40.
12. Baron, R.M., Lopez-Guzman, S., Riascos, D.F., Macias, A.A., Layne, M.D., Cheng, G., Harris, C., Chung, S.W., Reeves, R., von Andrian, U.H. *et al.* (2010) Distamycin A inhibits HMGA1-binding to the P-selectin promoter and attenuates lung and liver inflammation during murine endotoxemia. *PLoS One*, **5**, e10656.
13. Beerman, T.A., McHugh, M.M., Sigmund, R., Lown, J.W., Rao, K.E. and Bathini, Y. (1992) Effects of analogs of the DNA minor groove binder Hoechst-33258 on Topoisomerase-II and Topoisomerase-I mediated activities. *Biochim Biophys Acta*,

- 1131**, 53-61.
14. Beerman, T.A., Woynarowski, J.M., Sigmund, R.D., Gawron, L.S., Rao, K.E. and Lown, J.W. (1991) Netropsin and bis-netropsin analogs as inhibitors of the catalytic activity of mammalian DNA topoisomerase II and topoisomerase cleavable complexes. *Biochim Biophys Acta*, **1090**, 52-60.
  15. McHugh, M.M., Sigmund, R.D. and Beerman, T.A. (1990) Effects of minor groove binding drugs on camptothecin-induced DNA lesions in L1210 nuclei. *Biochem Pharmacol*, **39**, 707-714.
  16. McHugh, M.M., Woynarowski, J.M., Sigmund, R.D. and Beerman, T.A. (1989) Effect of minor groove binding drugs on mammalian topoisomerase I activity. *Biochem Pharmacol*, **38**, 2323-2328.
  17. Woynarowski, J.M., Sigmund, R.D. and Beerman, T.A. (1989) DNA minor groove binding agents interfere with topoisomerase II mediated lesions induced by epipodophyllotoxin derivative VM-26 and acridine derivative m-AMSA in nuclei from L1210 cells. *Biochemistry*, **28**, 3850-3855.
  18. Aggarwal, M., Sommers, J.A., Shoemaker, R.H. and Brosh, R.M., Jr. (2011) Inhibition of helicase activity by a small molecule impairs Werner syndrome helicase (WRN) function in the cellular response to DNA damage or replication stress. *Proc Natl Acad Sci U S A*, **108**, 1525-1530.
  19. Nickols, N.G., Jacobs, C.S., Farkas, M.E. and Dervan, P.B. (2007) Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem Biol*, **2**, 561-571.
  20. Wang, J.C. (2002) Cellular roles of DNA topoisomerases: A molecular perspective. *Nat Rev Mol Cell Bio*, **3**, 430-440.
  21. LIU, L.F. and Wang, J.C. (1987) Supercoiling of the DNA-Template during Transcription. *Proc Natl Acad Sci U S A*, **84**, 7024-7027.
  22. Mondal, N., Zhang, Y., Jonsson, Z., Dhar, S.K., Kannapiran, M. and Parvin, J.D. (2003) Elongation by RNA polymerase II on chromatin templates requires topoisomerase activity. *Nucleic Acids Res*, **31**, 5016-5024.
  23. Austin, C.A. and Marsh, K.L. (1998) Eukaryotic DNA topoisomerase II beta. *Bioessays*, **20**, 215-226.
  24. Ju, B.G., Lunyak, V.V., Perissi, V., Garcia-Bassets, I., Rose, D.W., Glass, C.K. and Rosenfeld, M.G. (2006) A topoisomerase II beta-mediated dsDNA break required for regulated transcription. *Science*, **312**, 1798-1802.
  25. Lin, C.R., Yang, L.Q., Tanasa, B., Hutt, K., Ju, B.G., Ohgi, K., Zhang, J., Rose, D.W., Fu, X.D., Glass, C.K. *et al.* (2009) Nuclear Receptor-Induced Chromosomal Proximity and DNA Breaks Underlie Specific Translocations in Cancer. *Cell*, **139**, 1069-1083.
  26. Haffner, M.C., Aryee, M.J., Toubaji, A., Esopi, D.M., Albadine, R., Gurel, B., Isaacs, W.B., Bova, G.S., Liu, W.N., Xu, J.F. *et al.* (2010) Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nat Genet*, **42**, 668-U645.
  27. Pommier, Y., Leo, E., Zhang, H.L. and Marchand, C. (2010) DNA Topoisomerases and Their Poisoning by Anticancer and Antibacterial Drugs. *Chem Biol*, **17**, 421-433.

28. Errington, F., Willmore, E., Tilby, M.J., Li, L., Li, G., Li, W., Baguley, B.C. and Austin, C.A. (1999) Murine transgenic cells lacking DNA topoisomerase II beta are resistant to acridines and mitoxantrone: Analysis of cytotoxicity and cleavable complex formation. *Mol Pharmacol*, **56**, 1309-1316.
29. Snapka, R.M., Gao, H.L., Grabowski, D.R., Brill, D., Chan, K.K., Li, L.G., Li, G.C. and Ganapathi, R. (2001) Cytotoxic mechanism of XK469: Resistance of topoisomerase II beta knockout cells and inhibition of topoisomerase I. *Biochem Biophys Res Commun*, **280**, 1155-1160.
30. Pourpak, A., Landowski, T.H. and Dorr, R.T. (2007) Ethonafide-induced cytotoxicity is mediated by topoisomerase II inhibition in prostate cancer cells. *Journal of Pharmacology and Experimental Therapeutics*, **321**, 1109-1117.
31. Nitiss, J.L. (2009) Targeting DNA topoisomerase II in cancer chemotherapy. *Nature Reviews Cancer*, **9**, 338-350.
32. Woynarowski, J.M., McHugh, M., Sigmund, R.D. and Beerman, T.A. (1989) Modulation of topoisomerase II catalytic activity by DNA minor groove binding agents distamycin, Hoechst 33258, and 4',6-diamidine-2-phenylindole. *Mol Pharmacol*, **35**, 177-182.
33. Branzei, D. and Foiani, M. (2010) Maintaining genome stability at the replication fork. *Nat Rev Mol Cell Bio*, **11**, 208-219.
34. Brosh, R.M., Jr., Karow, J.K., White, E.J., Shaw, N.D., Hickson, I.D. and Bohr, V.A. (2000) Potent inhibition of werner and bloom helicases by DNA minor groove binding drugs. *Nucleic Acids Res*, **28**, 2420-2430.
35. Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y. *et al.* (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, **316**, 1160-1166.
36. Belitsky, J.M., Nguyen, D.H., Wurtz, N.R. and Dervan, P.B. (2002) Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg Med Chem*, **10**, 2767-2774.
37. Vichai, V. and Kirtikara, K. (2006) Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat Protoc*, **1**, 1112-1116.

## Chapter 6

### **Mechanism of Py-Im Polyamide-Induced Cytotoxicity in Prostate Cancer Cells**

John W. Phillips, Benjamin C. Li, Kenneth K. Karanja, Judith L. Campbell, and  
Peter B. Dervan.

*Manuscript in preparation.*

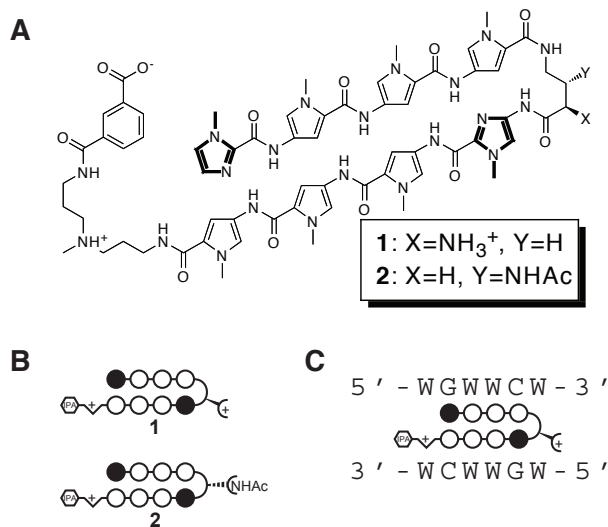
**Abstract**

We have developed 8-ring, hairpin, Py-Im polyamides that downregulate the expression of Androgen Receptor (AR) target genes in LNCaP cells. These compounds are cell-permeable, sequence-specific, high-affinity DNA minor groove binders that are designed to prefer the sequence 5'-WGWWCW-3' (W=A or T). This study investigates the cytotoxicity of polyamides targeted to this sequence in prostate cancer cells. We determine that these compounds are cytotoxic at micromolar concentrations and that this property is not dependent on intact AR signaling. Polyamide-treated cells undergo apoptotic cell death as marked by Caspase 3/7 activation and PARP cleavage after 24 hours. Prior to inducing programmed cell death, polyamide-treated cells arrest in S-phase. S-phase arrest occurs despite the apparent absence of DNA damage and without activating the DNA damage-induced S-phase checkpoint. The cytotoxic effects and cell cycle disruption produced by the AR-targeted polyamides in prostate cancer cells suggest that inhibition of AR-driven gene expression with these compounds is accompanied by AR-independent cell death.

## 6.1. Introduction

Py-Im polyamides are programmable small molecules capable of binding to DNA with affinities and specificities comparable to DNA-binding proteins (1). The sequence specificity arises from pairs of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) amino acids that interact with the hydrogen bonding pattern produced by DNA base pairs in the minor groove: Im/Py specifies a G-C base pair, Py/Im a C-G base pair, and Py/Py specifies T-A or A-T (2). They are cell-permeable compounds that localize to the nucleus in live cells and bind genomic DNA in its chromatin form (3-7). When designed to bind DNA sequences that match the response element of an endogenous transcription factor, they can disrupt the transcription factor-DNA interface and selectively inhibit inducible changes in downstream target gene expression (8-11). The Py-Im polyamide represents a framework for the development of chemical tools to control gene expression.

We have developed a polyamide (1) that antagonizes steroid-induced gene expression changes driven by the nuclear hormone receptors androgen receptor (AR) and glucocorticoid receptor (GR) (Figure 6.1). In animals, these transcription factors form powerful nodes in the transduction of signals from circulating hormones: AR controls the development and maintenance of the male sexual phenotype in response to testosterone, and GR regulates the expression of anti-inflammatory genes in response to



**Figure 6.1.** Chemical structure and binding preferences of the Py-Im polyamides used in this study. (A) Chemical structure showing the two different turn functionalities (B) Ball-and-stick representation of the polyamides. Open circles represent *N*-methylpyrrole residues, filled circles represent *N*-methylimidazoles. The hexagon represents the isophthalic acid moiety. (C) Both 1 and 2 bind the same 5'-WGWWCW-3' DNA sequence, where W = A or T.

cortisol (12,13). In the absence of ligand, nuclear hormone receptors are sequestered in the cytosol by heat shock proteins (HSPs). Steroid binding causes an allosteric conformational shift followed by homodimerization, nuclear localization, response element binding, and modulation of target gene transcription (14). AR and GR share the consensus binding sequence 5'-GGTACANNNTGTTCT-3' (15). Their genotropic actions can be partially inhibited in cancer tissue culture cells by polyamide **1**, which disrupts the protein-DNA interface by selectively binding the DNA sequence 5'-WGWWCW-3' (W= A or T) (8,10). In prostate cancer cells, AR signaling is mitogenic; whether or not this part of the androgen axis is inhibited by ARE-targeted polyamides is one of the subjects of this study.

Polyamides may be specific inhibitors of transcription factor-DNA binding in cell culture, but they are also part of a larger class of non-covalent, DNA-minor groove binders. This class of compounds includes the natural products distamycin A and netropsin as well as synthetic compounds like Hoechst 33342 and pentamidine (16). Of all of these, distamycin is perhaps the best studied. This compound binds to DNA with 2:1 stoichiometry and high affinity ( $K_a \approx 3 \times 10^8$ ), and is selective for A-T tracts (17-19). In cell culture, distamycin is known to inhibit the activity of topoisomerases I and II and to interfere with gene regulation by disrupting the binding of high mobility group (HMG) domain-containing proteins to DNA (20,21). *In vitro*, it is known to inhibit the activity of RNA and DNA polymerases as well as WRN and BLM helicases (22,23). Its cytotoxicity is in the 400  $\mu$ M range, and it is unknown which of these processes induces cell death. Derivatives appended with alkylating or other adduct-forming functionalities are two to three orders of magnitude more potent (24).

This study addresses the growth-inhibitory and cytotoxic properties of two structurally related Py-Im polyamides (compounds **1** and **2**, Figure 6.1). They are selective for the same 5'-WGWWCW-3' DNA sequence, and both downregulate the expression of AR target genes in LNCaP cell culture, albeit with different potencies (25). To explore the AR dependence of any potential growth inhibitory effects, we establish the cytotoxicity

of these compounds in cell culture models of hormone-sensitive and hormone-refractory prostate cancer. Using assays for Caspase 3/7 activity and PARP cleavage, we show the cell death to be apoptotic in nature. Flow cytometric cell cycle analysis reveals a dose-dependent S-phase arrest. S-phase stress is often the result of DNA damage, but single cell alkaline gel electrophoresis and Western blot analysis of checkpoint proteins show no evidence of single- or double-strand DNA breaks. The results have important implications for the design and use of Py-Im polyamides as chemical probes of protein-DNA interactions in biological systems.



## 6.2. Results.

**6.2.1. Cytotoxicity of a Py-Im polyamide in prostate cancer cells.** We conducted cytotoxicity assays with compounds **1** and **2** in three different prostate cancer cell lines. LNCaP is a tissue culture model of hormone-sensitive prostate cancer (26). LNAR is an engineered LNCaP cell line containing a stably incorporated construct that increases AR expression roughly 5-fold (27). In tumor xenograft studies in mice, this upregulation was sufficient to induce resistance to first-generation anti-androgens. The third cell line, DU145, expresses very low levels of AR and does not produce PSA or exhibit androgen-dependent growth: it is considered hormone insensitive (28). The compounds displayed time- and dose-dependent cytotoxicity at 72 and 96 h as measured by sulfarhodamine B staining in all three cell lines, regardless of AR status (Table 6.1). There was no correlation between AR status

and  $IC_{50}$  value, suggesting

that the observed cytotoxicity occurred via an AR-independent mechanism.

Compound **2** had

approximately tenfold higher

potency than compound **1** at

both time points, which is

consistent with its previously



reported increased potency

against AR-driven gene expression. This may represent uptake differences between the

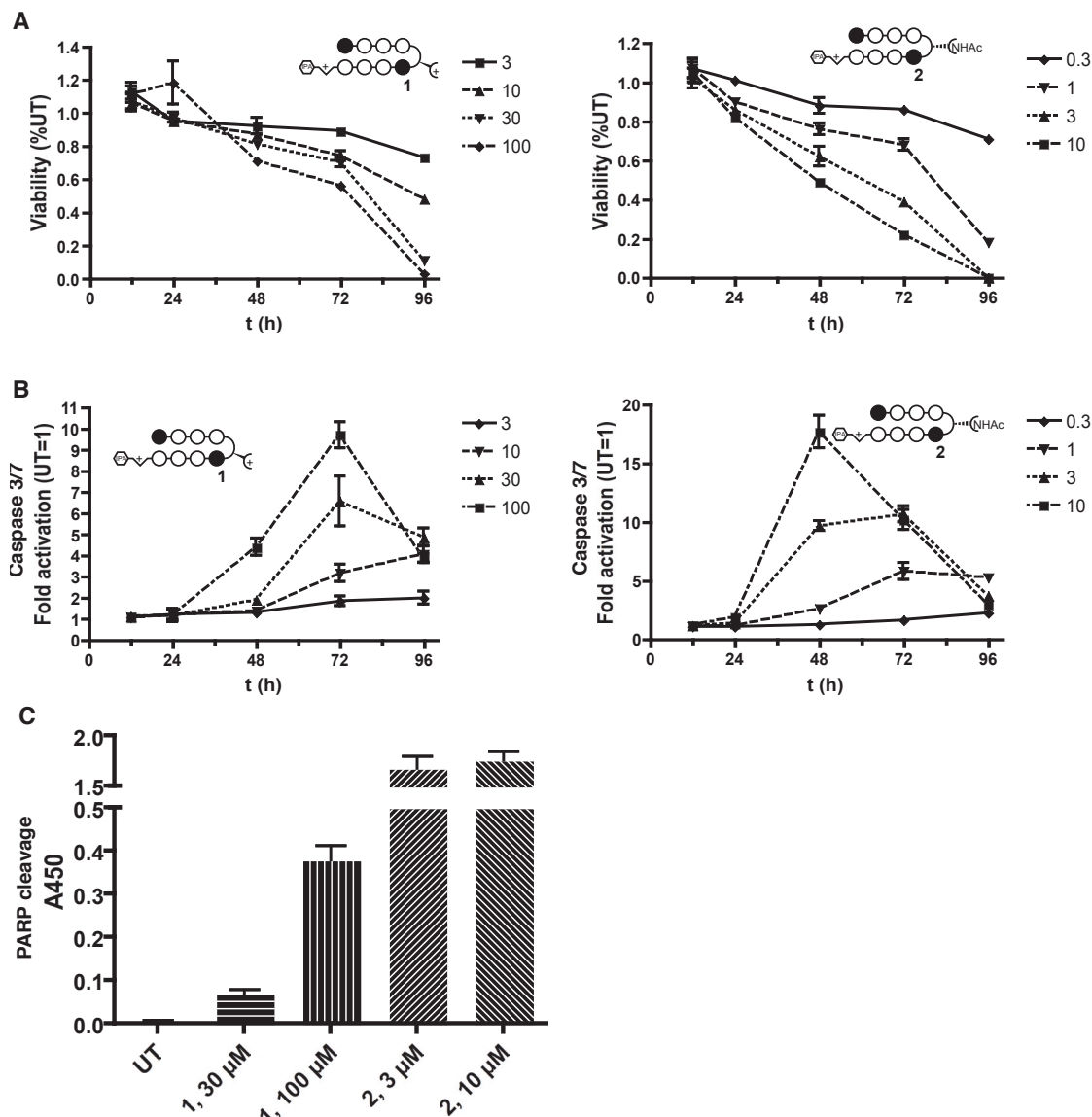
two compounds, as small structural changes have been observed to dramatically change the

cellular uptake properties of polyamides (29).

**6.2.2. Polyamide treatment induces apoptosis.** The cell death response to many DNA-

Cell line	AR	Cytotoxicity $IC_{50}$ values ( $\mu$ M)			
					
		72h	96h	72h	96h
LNCaP	+	18±4	7±3	1.8±0.9	0.6±0.2
LNAR	+++	40±10	36±14	3±1	1.5±0.2
DU145	-	14±4	8±4	1.5±0.5	0.76±0.06

**Table 6.1.** Summary of cytotoxicity  $IC_{50}$  values of WG-WWCW-targeted polyamides in AR-expressing (+, LNCaP), AR-overexpressing (+++, LNAR), and AR-negative (-, DU145) cancer cell lines. Cells were treated continuously with test compounds for 72 or 96 h. before fixation and staining. Values represent the mean  $\pm$  s.d. of three replicates.



**Figure 6.2.** Polyamide treatment induces apoptosis in DU145 cells. (A) Cell viability assay. Cells were treated in quadruplicate with a range of polyamide concentrations for up to 96 h and then assayed for bioreductive capacity with WST-1 reagent. The data are normalized to the untreated condition. (B) Caspase 3/7 activity assay. Cells were treated in triplicate for the indicated time and then homogenized in guanidinium lysis buffer containing a proluminescent Caspase 3/7 substrate. The data are normalized to the untreated condition. (C) ELISA for cleaved PARP formation. Cells were treated with compounds for 72 h before assaying the lysates by sandwich ELISA using an HRP-conjugated secondary antibody and a chromogenic substrate. The data are presented as the background-corrected absorbance values at 450 nm. All data are represented as the mean  $\pm$  s.d. of experiments conducted in triplicate or quadruplicate.

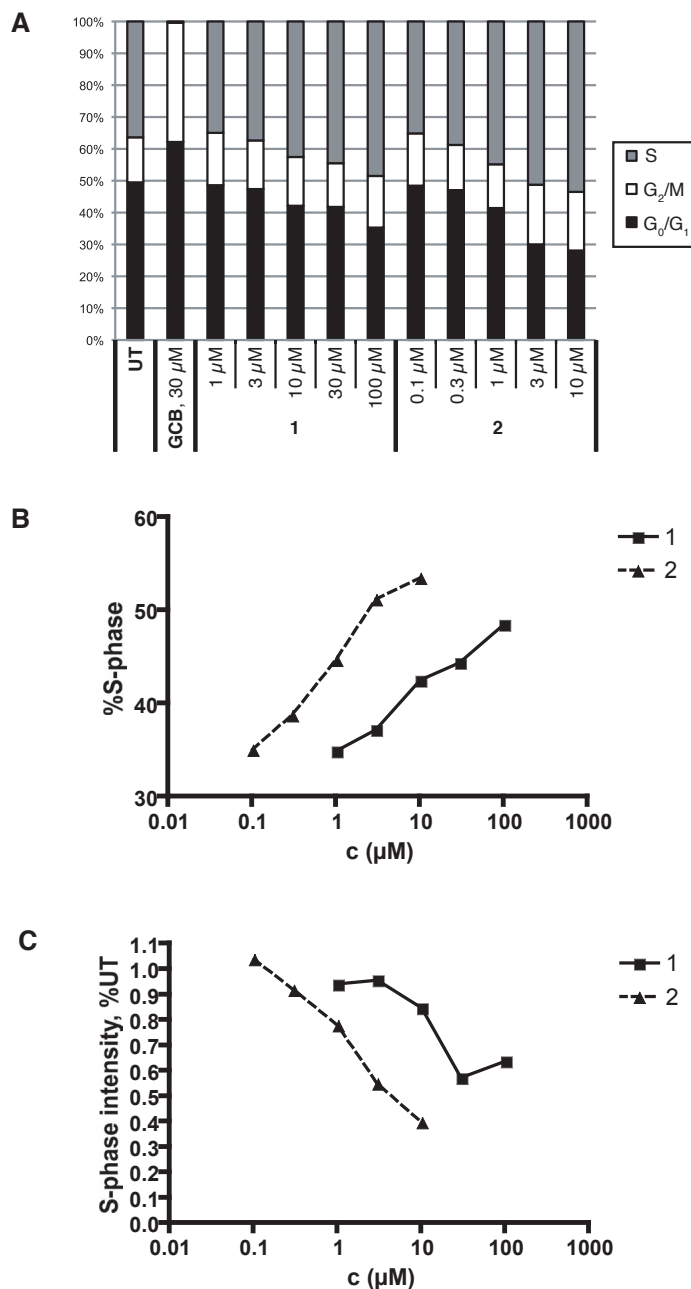
binding agents involves induction of apoptosis, two hallmarks of which are caspase activation and subsequent cleavage of poly(ADP-ribosyl) transferase (PARP) (30). To assess the involvement of the apoptotic pathway in the cell death response to polyamide treatment, we treated DU145 cells with the polyamides and assayed for decreased cell viability and

Caspase 3/7 over the same time course (Figure 6.2). Viability was measured with WST-1 reduction; Caspase 3/7 activation was measured with a proluminescent peptide substrate. Steep declines in viability did not occur until after 48 h of treatment with polyamide **1** or 72 h. of treatment with compound **2**. Cell viability decreased concomitantly with induction of Caspase 3/7 activity in a time- and concentration-dependent manner. With compound **1**, elevated levels of Caspase 3/7 activity were detected over the entire dose range (3-100  $\mu$ M), with peak activity measured at approximately tenfold activation with 100  $\mu$ M **1** after 72 h. The results were similar with **2**, albeit with earlier caspase induction at approximately ten-fold lower concentration than was observed with **1**. Caspase activation was also accompanied by PARP cleavage, as measured after 72 h. continuous exposure by sandwich ELISA (Figure 6.2C), confirming that the cell death response to polyamide treatment involves apoptosis.

**6.2.3. Polyamide treatment induces replication stress.** Many DNA-binding agents are known to cause cell cycle disturbances. To examine the effects of polyamides **1** and **2** on the cell cycle, we pulse-labeled polyamide-treated DU145 cells for 30 min. with ethynyldeoxyuridine (EdU) after 24 h of compound treatment at various concentrations (Figure 6.3). The cells were harvested, and the incorporated, EdU-labeled nucleotides were conjugated with an azido-Alexa488 fluorophore using copper-catalyzed cycloaddition chemistry. Total DNA content was determined with 7-aminoactinomycin D staining, and the cells were subjected to two-color flow cytometric analysis. Both compounds produced a dose-dependent increase in the percentage of cells in S-phase, with a corresponding drop in the percentage of  $G_0/G_1$  cells. The proportion of  $G_2/M$  phase cells was not affected in a dose-dependent fashion. In addition, although the percentage of cells in S-phase increased, the average intensity of EdU staining decreased, suggesting that the treated cells were replicating their DNA more slowly.

#### 6.2.4. Polyamide treatment does not cause DNA damage.

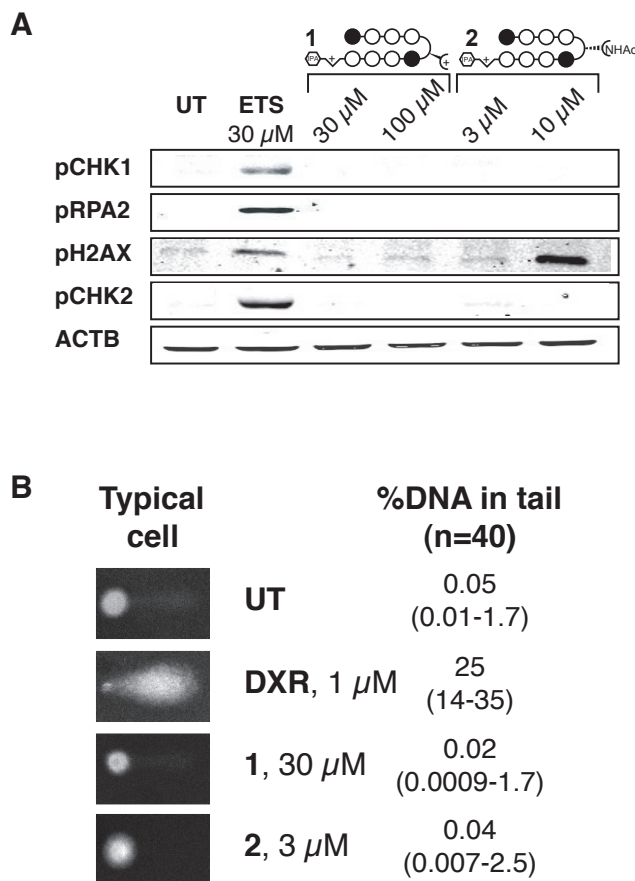
S-phase arrest subsequent to treatment with a DNA binding compound is suggestive of S-phase checkpoint activation in response to DNA damage or replication stress. In this pathway, DNA lesions and stalled replication forks activate ATM/ATR kinases. These proteins phosphorylate many substrates to activate checkpoints that slow cell cycle progression and allow time for DNA damage repair (31). In the presence of DNA damage, ATR phosphorylates RPA2 at S4 and S8, and CHK1 at S345 (32,33). ATM phosphorylates CHK2 at T68 after DNA damage by ionizing radiation. However, CHK2 is also phosphorylated in an ATM-independent manner after replication stress caused by depletion of nucleotide pools by hydroxyurea (34,35).



**Figure 6.3.** Polyamide treatment causes S-phase arrest. (A) Cell cycle distribution of DU145 cells treated with compounds or controls for 24h as measured by two-color flow cytometric evaluation of EdU pulse-labeled cells stained for DNA content with 7AAD. (B) %S-phase data from (A) shown as a function of increasing polyamide dose showing S-phase arrest. (C) Dose-dependent decrease in average EdU incorporation indicative of slowed DNA synthesis in response to polyamide treatment. 20,000 cells were counted for each sample. The experiment was conducted in duplicate.

To look for activation of the S-phase checkpoint by DNA damage or replication stress, we treated DU145 cells with toxic levels of compounds **1** and **2** (30 and 3  $\mu\text{M}$ , respectively) for 18 h and then performed Western blotting on total cellular lysates using phospho-specific antibodies against RPA2, CHK1, and CHK2 (Figure 6.4A). Etoposide, a DNA-binding compound that poisons Topoisomerase II (TOPII) to produce DNA double-strand breaks, was included as a positive control. In the polyamide-treated samples, none of the markers of DNA damage during S-phase were activated; no phospho-CHK1, -CHK2, or -RPA2 was detected.

The absence of downstream signals from ATM/ATR suggests that the S-phase arrest seen in polyamide-treated cells is not due to DNA damage from stalled replication forks. Histone H2AX, a marker of DNA double-strand breaks (36), was observed to be phosphorylated when the cells were treated with 10  $\mu\text{M}$  **2**. Although ATM rapidly phosphorylates H2AX in response to DNA damage, the absence of any other signals of DNA damage suggests



**Figure 6.4:** Polyamide treatment does not induce DNA damage or activate the DNA-damage induced S-phase checkpoint. (A) Western blot of S-phase and DNA damage response proteins using phospho-specific antibodies to probe DU145 lysates from cells treated with polyamides or controls for 18 h. (B) Single cell gel electrophoretic analysis of DU145 cells treated with polyamides or controls for 24 h. Photomicrographs of a representative cell nucleus from each treatment condition are shown on the left. Migration of DNA from the centroid is proportional to the amount of DNA damage. Images of 40 randomly selected cells from each treatment condition were analyzed with image processing software to determine the percentage of DNA in each 'tail.' The data are expressed as the median and 95% confidence values.

an alternative, independent mechanism for its phosphorylation in this case. Namely, the observed phospho-H2AX at high concentrations of **2** may be a result of apoptotic activation of caspase cleavage of DNA.

To look for DNA damage as a result of polyamide treatment more directly, we subjected DU145 cells to cytotoxic levels of the two polyamides for 24 h and then assayed them by semiquantitative single cell alkaline gel electrophoresis (comet assay) (Figure 6.4B). Migration of the DNA from the centroid into the 'comet tail' is proportional to the amount of single- and double-strand breakage that has occurred. Cells treated with doxorubicin, a known DNA-damaging agent, had a median value of 25% of total DNA in the tail. Polyamide-treated cells were indistinguishable from the untreated control, with median tail % values of 0.02 and 0.04 for compounds **1** and **2**, respectively.

### 6.3. Discussion

This study seeks to define the effects of two 8-ring, hairpin, Py-Im polyamide antagonists of AR-driven gene expression on the growth and viability of prostate cancer cells. We show that polyamides **1** and **2** cause cytotoxic effects and cell cycle disturbance at concentrations known to interfere with the induction of some AR-target genes in LNCaP cells in response to androgen stimulus (10,25). AR signaling is mitogenic in LNCaP cells, but the set of the genes involved is still being defined with massively parallel sequencing techniques (37). Nevertheless, anti-androgen steroid antagonists like bicalutamide slow the growth of this cell line, producing arrest in the  $G_0/G_1$  phase of the cell cycle and apoptosis (38,39). The mechanism of cell death produced by AR-targeted polyamides also involves apoptosis, but there is no dependence on intact AR signaling to produce this effect; the polyamides are equitoxic in LNCaP (AR+, hormone-sensitive) and DU145 (AR-, hormone-insensitive) cells. In addition, the growth arrest due to polyamide treatment occurs in S-phase, not in  $G_0/G_1$ . The polyamide AR antagonists retard the growth of cancer cells independent of their effects on AR-regulated mRNA expression.

The cytotoxic effects of the polyamides used in this study are observed in the same concentration range and on the same timescale as their effects on AR-driven gene expression. These data suggest that it may not be possible to divorce their specific effects on AR-ARE binding from their non-AR-related induction of programmed cell death based on dose titration. Compounds **1** and **2** both bind the DNA sequence 5'-WGWWCW-3', a relatively short and degenerate motif that occurs millions of times in the human genome. The AR consensus DNA binding sequence is also degenerate, but transcription factors can take advantage of a wealth of additional interactions including protein-protein contacts and allosteric modulation of the DNA topology to restrict their binding to a defined set of response elements (40). Future efforts to disrupt transcription factor-DNA binding with polyamides may need to focus on compounds with higher specificity. For example,

polyamides with higher imidazole content would have fewer binding sites genome wide, as would larger compounds with 10 or 12 ring pairs that could specify 8 or 10 DNA base pairs. In fact, larger hairpin compounds containing beta-alanine residues to increase flexibility have been employed to inhibit the binding of the AP-1 transcription factor in order to downregulate MMP9 expression in cell culture and mouse models of metastatic colorectal cancer (41).

The S-phase growth arrest produced by polyamides **1** and **2** may be the signal event that produces the eventual cytotoxicity. The effects on the cell cycle are large and occur in advance of the drop in cell viability. S-phase arrest is apparent after 18 h of continuous treatment, whereas the cell viability does not drop sharply until at least 24 h of treatment (Figure 6.2). In addition to the temporal considerations, one should note that the 10-fold difference in potency between the two compounds in the cytotoxicity assay is also present in the cell cycle data. Compound **2** produces cell cycle disturbance at 10-fold lower concentrations than compound **1**. Compound **2** also produces a greater maximal inhibition of EdU incorporation than compound **1**. The size of the cell cycle disturbance, its relatively rapid onset, and its correlation with the cytotoxicity  $IC_{50}$  values suggest that inhibition of DNA synthesis is a primary mechanism of polyamide-induced cytotoxicity.

The fact that the accumulation of cells in S-phase produced by the polyamides occurs in the absence of any observable DNA damage suggests activation of a damage-independent intra-S phase checkpoint. The absence of CHK1, RPA2 and H2AX phosphorylation tends to rule out DNA damage, while the absence of CHK2 phosphorylation suggests that although replication is impeded, the replication forks have not collapsed to form double stranded breaks. Activation of the intra-S phase checkpoint can arise after phosphorylation of the replisome proteins including MCM2-7, MCM10 and DNA polymerases (42). In metazoans, conditions of replicative stress or DNA damage led to phosphorylation of MCM2-7, the replicative helicase, by ATM/ATR to activate a checkpoint (43-45). Examining other substrates of ATM/ATR phosphorylated after replicative stress could help



define the mechanism of polyamide-induced S-phase arrest.

One possibility is that the polyamides are interfering with replication by inhibiting topoisomerase function. But as non-covalent minor groove binders, polyamides would be anticipated to inhibit binding of these enzymes to DNA without causing formation of the cleavable complex required to generate DNA damage. In addition, inhibitors of TOPI or TOPII catalytic activity arrest cells in G2/M due to activation of the G2/M checkpoint or decatenation failure (46,47), not in S-phase as observed. An additional possibility is that the polyamides are inhibiting replicative DNA helicase activity, as the structurally related DNA minor groove-binder distamycin is known to do *in vitro*. However, it is unclear if this mode of inhibition would be expected to produce DNA damage, for a synthetic WRN helicase inhibitor was recently shown to cause S-phase delay and H2AX phosphorylation in cell culture (48). Studies of fork progression using *in vitro* reconstitution models would be useful for defining the biochemical target for polyamide cytotoxicity.

## 6.4. Materials and methods

**6.4.1. Chemicals and reagents.** Compounds **1** and **2** were synthesized on solid phase Kaiser oxime resin using previously published protocols (49). Cell culture media was purchased from Invitrogen, and fetal bovine serum from Irvine Scientific. Gemcitabine was purchased from AvaScientific. Etoposide and doxorubicin were purchased from Sigma-Aldrich, as were all other reagents unless otherwise noted.

**6.4.2. Cell culture.** LNCaP, LNAR, and DU145 cells were maintained in RPMI 1640 with 10% FBS at 37 °C under 5% CO<sub>2</sub>. LNCaP and DU145 cells were purchased from ATCC (Manassas, VA). LNAR cells were a gift from C.L. Sawyers at Memorial Sloan-Kettering Cancer Center (NY, NY).

**6.4.3. Cytotoxicity assays.** IC<sub>50</sub> values for cytotoxicity were determined using a previously-described, sulfarhodamine-based, colorimetric assay for cellular protein content in 96-well microplates (50). LNCaP and LNAR cells were plated at 3,000 or 4,000 cells per well for the 72 h and 96 h timepoints, respectively. DU145 cells were plated at 2,000 or 2,500 cells per well. Compounds were added in 100 μL RPMI1640 supplemented with 10% FBS 24 h after plating. Quadruplicate wells were used for each concentration. At the appropriate time, the cells were fixed with 100 μL 10% trichloroacetic acid solution, washed, stained, and dried as described. 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. IC<sub>50</sub> 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=max., B=min., and H=Hill Slope. Three independent trials were averaged; stated IC<sub>50</sub> values represent the mean and standard deviation. These

calculations were performed using Prism 4 (GraphPad) software.

**6.4.4. Apoptosis assays.** DU145 cells were plated in 96-well microplates at 2,000-8,000 cells per well. As above, compounds and controls were added 24 h after plating. Each time point was assayed in triplicate. At harvest, Caspase 3/7 activity was assessed using 100  $\mu$ L of Caspase-Glo reagent (Promega), which contains the proluminescent caspase substrate DEVD-aminoluciferin. Luminescence was measured after 30 min incubation at room temperature. Luminescence data are expressed as a fold difference from untreated controls as measured using a Victor microplate reader (PerkinElmer).

The cell viability of each treatment condition was monitored in a sister plate using a tetrazolium-based assay for mitochondrial bioreductive capacity (51). 10  $\mu$ L WST-1 reagent (Roche) was added to each well and incubated at 37 °C for 30 min before measuring the absorbance at 450 nm. The WST-1 data are corrected for background absorbance and expressed as a percentage of untreated controls.

**6.4.5. PARP cleavage assay.** 400,000 DU145 cells were plated in 10 cm diameter dishes. Compounds were added after 24 h and were allowed to incubate an additional 72 h. At harvest, cells were washed once with PBS then treated with 400  $\mu$ L ice-cold lysis buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 1 mM Na<sub>2</sub>EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 2 mM  $\beta$ -glycerophosphate, 1 mM Na<sub>3</sub>VO<sub>4</sub>, 1  $\mu$ g/mL leupeptin, 1 mM phenylmethanesulfonylfluoride) for 5 min at 5°C. The lysate was sonicated for 15 s and then centrifuged for 10 min at 20,000 x g at 5°C. The supernatant was retained. Protein concentrations were determined by Bradford assay (Bio-Rad) using bovine serum albumin (Bio-Rad) to create a standard curve. PARP cleavage was assayed by sandwich ELISA (Cell Signaling Technology) and performed according to the manufacturer's recommendations. 10  $\mu$ g total protein was loaded into each well of a microplate coated with anti-cleaved PARP (Asp214) mouse mAb and allowed to incubate overnight at 5

°C. Rabbit anti-PARP mAb was then added, followed by anti-rabbit IgG conjugated to HRP. Triplicate wells were included for each condition, and the data are representative of both experimental replicates. The data are expressed as fold change from the untreated condition, showing the mean and standard deviation of each measurement.

**6.4.6. Cell cycle analysis.** 800,000 DU145 cells were plated in 10 cm diameter dishes for 24 h before treatment with test compounds for an additional 24h. 10  $\mu$ M EdU was added 30 min before harvest. The cells were harvested by trypsinization and combined with the cell culture supernatant before pelleting at 300 x g. Following overnight fixation in 70% ethanol, the cells were rehydrated in 1% BSA/PBS and processed with the Click-it EdU Alexa Fluor 488 Flow Cytometry assay kit (Invitrogen) using half the recommended A488 reagent. After overnight treatment with 0.2 mg/mL RNase A in 1% BSA/PBS, the cells were stained for DNA content with the provided 7-aminoactinomycin D and analyzed on a FACSCalibur (Becton-Dickinson) instrument. The data were analyzed using FlowJo v8.8.2 (TreeStar) and are representative of two trials.

**6.4.7. Phospho-specific Western blot.** 800,000 DU145 cells were plated in 10-cm diameter dishes for 24 h. before treatment with test compounds for an additional 18 h. Cells were washed once with ice-cold PBS and lysed in 125  $\mu$ L TBS-Tx buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% Triton X100) containing fresh 1 mM PMSF, protease inhibitors (Roche), and phosphatase inhibitors. Lysates were vortexed, placed on ice for 20 min., and then clarified by brief sonication and centrifugation at 14,000 rpm for 10 min. The samples were quantified by Bradford assay and denatured by boiling in Laemmli buffer. 25  $\mu$ g of total protein was analyzed by SDS-PAGE using a 13% polyacrylamide gel for RPA2 pS4/S8 (~32 kDa),  $\beta$ -actin (~42 kDa), and CHK1 pS345 (~60 kDa). A 15% gel was used for  $\gamma$ H2Ax (~17 kDa) and a 10% gel was used for CHK2 pT68 (~62 kDa). After transfer to the nitrocellulose membrane and blocking with 0.2% Tween, phosphospecific rabbit anti-

human primary antibodies were used to probe CHK1 (Cell Signaling Technologies), CHK2 (Abcam),  $\gamma$ H2Ax (Abcam), and RPA32 pS4/S8 (Bethyl) overnight at 4 °C. The rabbit anti-human  $\beta$ -Actin (Sigma) antibody was used as a loading control. Goat anti-rabbit, near-IR conjugated secondary antibody (Licor) was added and the bands were visualized on an Odyssey infrared imager (Licor). The experiment was conducted in duplicate and the data are representative of both trials.

**6.4.8. Single-cell gel electrophoresis.** The apparatus and reagent kit were purchased from Trevigen (Gaithersburg, MD). The assay was performed according to the manufacturer's recommendations. Briefly, 800,000 DU145 cells were plated in 10 cm diameter dishes for 24 h before treatment with test compounds for an additional 24 h. The cells were harvested with a rubber policeman, pelleted by centrifugation, and washed once with ice-cold PBS before being suspended in 37 °C low melting-point agarose at  $1 \times 10^5$  cells/mL. 50  $\mu$ L of the suspension was placed on a 37 °C glass slide and allowed to cool for 30 min. The slides were bathed in lysis buffer for 30 min followed by a 30 min treatment with alkaline unwinding buffer (200 mM NaOH, 1 mM EDTA) at 5 °C. The slides were subjected to electrophoresis at 21V in a prechilled apparatus and fresh unwinding buffer for 30 min. The slides were washed twice in dH<sub>2</sub>O and once in 70% ethanol, then dried for 30 min at 45 °C. Dried slides were stained with 100  $\mu$ L 1X SYBR Green I in 10 mM Tris-HCl pH 7.5 for 30 min at 5 °C, and excess dye was removed by blotting. Slides were dried and stored at room temperature with desiccant. Comets were visualized using a LSM 5 Pascal confocal microscope with a 5x objective (Zeiss) and scored using CometScore image analysis software (TriTek). A random sampling of 40 cells was analyzed for each condition. The data are displayed as the median and 95% confidence value and are representative of two trials.

## 6.5. References.

1. Dervan, P.B., Poulin-Kerstien, A.T., Fechter, E.J. and Edelson, B.S. (2005) Regulation of gene expression by synthetic DNA-binding ligands. *Top Curr Chem*, **253**, 1-31.
2. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
3. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc Natl Acad Sci U S A*, **100**, 12063-12068.
4. Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B. and Luger, K. (2003) Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J Mol Biol*, **326**, 371-380.
5. Dudouet, B., Burnett, R., Dickinson, L.A., Wood, M.R., Melander, C., Belitsky, J.M., Edelson, B., Wurtz, N., Briehn, C., Dervan, P.B. *et al.* (2003) Accessibility of nuclear chromatin by DNA binding polyamides. *Chem Biol*, **10**, 859-867.
6. Nguyen-Hackley, D.H., Ramm, E., Taylor, C.M., Joung, J.K., Dervan, P.B. and Pabo, C.O. (2004) Allosteric inhibition of zinc-finger binding in the major groove of DNA by minor-groove binding ligands. *Biochemistry*, **43**, 3880-3890.
7. Gearhart, M.D., Dickinson, L., Ehley, J., Melander, C., Dervan, P.B., Wright, P.E. and Gottesfeld, J.M. (2005) Inhibition of DNA binding by human estrogen-related receptor 2 and estrogen receptor alpha with minor groove binding polyamides. *Biochemistry*, **44**, 4196-4203.
8. Muzikar, K.A., Nickols, N.G. and Dervan, P.B. (2009) Repression of DNA-binding dependent glucocorticoid receptor-mediated gene expression. *Proc Natl Acad Sci U S A*, **106**, 16598-16603.
9. Nickols, N.G., Jacobs, C.S., Farkas, M.E. and Dervan, P.B. (2007) Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem Biol*, **2**, 561-571.
10. Nickols, N.G. and Dervan, P.B. (2007) Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci U S A*, **104**, 10418-10423.
11. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G., Jr. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A*, **101**, 16768-16773.
12. Lamont, K.R. and Tindall, D.J. (2010) Androgen regulation of gene expression. *Adv Cancer Res*, **107**, 137-162.
13. Nicolaidis, N.C., Galata, Z., Kino, T., Chrousos, G.P. and Charmandari, E. (2010) The human glucocorticoid receptor: molecular basis of biologic function. *Steroids*, **75**, 1-12.
14. Aranda, A. and Pascual, A. (2001) Nuclear hormone receptors and gene expression. *Physiol Rev*, **81**, 1269-1304.
15. Roche, P.J., Hoare, S.A. and Parker, M.G. (1992) A consensus DNA-binding site

- for the androgen receptor. *Mol Endocrinol*, **6**, 2229-2235.
16. Nelson, S.M., Ferguson, L.R. and Denny, W.A. (2007) Non-covalent ligand/DNA interactions: minor groove binding agents. *Mutat Res*, **623**, 24-40.
  17. Breslauer, K.J., Remeta, D.P., Chou, W.Y., Ferrante, R., Curry, J., Zaunczkowski, D., Snyder, J.G. and Marky, L.A. (1987) Enthalpy-entropy compensations in drug-DNA binding studies. *Proc Natl Acad Sci U S A*, **84**, 8922-8926.
  18. Pelton, J.G. and Wemmer, D.E. (1989) Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc Natl Acad Sci U S A*, **86**, 5723-5727.
  19. Schultz, P.G. and Dervan, P.B. (1984) Distamycin and penta-N-methylpyrrolicarboxamide binding sites on native DNA. A comparison of methidiumpropyl-EDTA-Fe(II) footprinting and DNA affinity cleaving. *J Biomol Struct Dyn*, **1**, 1133-1147.
  20. Beerman, T.A., McHugh, M.M., Sigmund, R., Lown, J.W., Rao, K.E. and Bathini, Y. (1992) Effects of analogs of the DNA minor groove binder Hoechst-33258 on Topoisomerase-II and Topoisomerase-I mediated activities. *Biochim Biophys Acta*, **1131**, 53-61.
  21. Baron, R.M., Lopez-Guzman, S., Riascos, D.F., Macias, A.A., Layne, M.D., Cheng, G., Harris, C., Chung, S.W., Reeves, R., von Andrian, U.H. *et al.* (2010) Distamycin A inhibits HMGA1-binding to the P-selectin promoter and attenuates lung and liver inflammation during murine endotoxemia. *PLoS One*, **5**, e10656.
  22. Zimmer, C., Puschendorf, B., Grunicke, H., Chandra, P. and Venner, H. (1971) Influence of netropsin and distamycin A on the secondary structure and template activity of DNA. *Eur J Biochem*, **21**, 269-278.
  23. Brosh, R.M., Jr., Karow, J.K., White, E.J., Shaw, N.D., Hickson, I.D. and Bohr, V.A. (2000) Potent inhibition of werner and bloom helicases by DNA minor groove binding drugs. *Nucleic Acids Res*, **28**, 2420-2430.
  24. Pezzoni, G., Grandi, M., Biasoli, G., Capolongo, L., Ballinari, D., Giuliani, F.C., Barbieri, B., Pastori, A., Pesenti, E., Mongelli, N. *et al.* (1991) Biological Profile of Fce-24517, a Novel Benzoyl Mustard Analog of Distamycin-A. *Br J Cancer*, **64**, 1047-1050.
  25. Chenoweth, D.M., Harki, D.A., Phillips, J.W., Dose, C. and Dervan, P.B. (2009) Cyclic pyrrole-imidazole polyamides targeted to the androgen response element. *J Am Chem Soc*, **131**, 7182-7188.
  26. Horoszewicz, J.S., Leong, S.S., Kawinski, E., Karr, J.P., Rosenthal, H., Chu, T.M., Mirand, E.A. and Murphy, G.P. (1983) LNCaP model of human prostatic carcinoma. *Cancer Res*, **43**, 1809-1818.
  27. Chen, C.D., Welsbie, D.S., Tran, C., Baek, S.H., Chen, R., Vessella, R., Rosenfeld, M.G. and Sawyers, C.L. (2004) Molecular determinants of resistance to antiandrogen therapy. *Nat Med*, **10**, 33-39.
  28. Alimirah, F., Chen, J., Basrawala, Z., Xin, H. and Choubey, D. (2006) DU-145 and PC-3 human prostate cancer cell lines express androgen receptor: implications for the androgen receptor functions and regulation. *FEBS Lett*, **580**, 2294-2300.
  29. Edelson, B.S., Best, T.P., Olenyuk, B., Nickols, N.G., Doss, R.M., Foister, S., Heckel, A. and Dervan, P.B. (2004) Influence of structural variation on nuclear



- localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res*, **32**, 2802-2818.
30. Tewari, M., Quan, L.T., O'Rourke, K., Desnoyers, S., Zeng, Z., Beidler, D.R., Poirier, G.G., Salvesen, G.S. and Dixit, V.M. (1995) Yama/CPP32 beta, a mammalian homolog of CED-3, is a CrmA-inhibitable protease that cleaves the death substrate poly(ADP-ribose) polymerase. *Cell*, **81**, 801-809.
  31. Dai, Y. and Grant, S. (2010) New insights into checkpoint kinase 1 in the DNA damage response signaling network. *Clin Cancer Res*, **16**, 376-383.
  32. Olson, E., Nievera, C.J., Klimovich, V., Fanning, E. and Wu, X. (2006) RPA2 is a direct downstream target for ATR to regulate the S-phase checkpoint. *J Biol Chem*, **281**, 39517-39533.
  33. Gravel, S., Chapman, J.R., Magill, C. and Jackson, S.P. (2008) DNA helicases Sgs1 and BLM promote DNA double-strand break resection. *Genes Dev*, **22**, 2767-2772.
  34. Matsuoka, S., Huang, M. and Elledge, S.J. (1998) Linkage of ATM to cell cycle regulation by the Chk2 protein kinase. *Science*, **282**, 1893-1897.
  35. Matsuoka, S., Rotman, G., Ogawa, A., Shiloh, Y., Tamai, K. and Elledge, S.J. (2000) Ataxia telangiectasia-mutated phosphorylates Chk2 in vivo and in vitro. *Proc Natl Acad Sci U S A*, **97**, 10389-10394.
  36. Rogakou, E.P., Boon, C., Redon, C. and Bonner, W.M. (1999) Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J Cell Biol*, **146**, 905-916.
  37. Yu, J., Mani, R.S., Cao, Q., Brenner, C.J., Cao, X., Wang, X., Wu, L., Li, J., Hu, M., Gong, Y. *et al.* (2010) An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell*, **17**, 443-454.
  38. Masiello, D., Cheng, S., Buble, G.J., Lu, M.L. and Balk, S.P. (2002) Bicalutamide functions as an androgen receptor antagonist by assembly of a transcriptionally inactive receptor. *J Biol Chem*, **277**, 26321-26326.
  39. Lee, E.C., Zhan, P., Schallhom, R., Packman, K. and Tenniswood, M. (2003) Antiandrogen-induced cell death in LNCaP human prostate cancer cells. *Cell Death Differ*, **10**, 761-771.
  40. Panne, D., Maniatis, T. and Harrison, S.C. (2007) An atomic model of the interferon-beta enhanceosome. *Cell*, **129**, 1111-1123.
  41. Wang, X., Nagase, H., Watanabe, T., Nobusue, H., Suzuki, T., Asami, Y., Shinjima, Y., Kawashima, H., Takagi, K., Mishra, R. *et al.* (2010) Inhibition of MMP-9 transcription and suppression of tumor metastasis by pyrrole-imidazole polyamide. *Cancer Sci*, **101**, 759-766.
  42. Matsuoka, S., Ballif, B.A., Smogorzewska, A., McDonald, E.R., 3rd, Hurov, K.E., Luo, J., Bakalarski, C.E., Zhao, Z., Solimini, N., Lerenthal, Y. *et al.* (2007) ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science*, **316**, 1160-1166.
  43. Cortez, D., Glick, G. and Elledge, S.J. (2004) Minichromosome maintenance proteins are direct targets of the ATM and ATR checkpoint kinases. *Proc Natl Acad Sci U S A*, **101**, 10078-10083.



44. Yoo, H.Y., Shevchenko, A. and Dunphy, W.G. (2004) Mcm2 is a direct substrate of ATM and ATR during DNA damage and DNA replication checkpoint responses. *J Biol Chem*, **279**, 53353-53364.
45. Shi, Y., Dodson, G.E., Mukhopadhyay, P.S., Shanware, N.P., Trinh, A.T. and Tibbetts, R.S. (2007) Identification of carboxyl-terminal MCM3 phosphorylation sites using polyreactive phosphospecific antibodies. *J Biol Chem*, **282**, 9236-9243.
46. Larsen, A.K., Escargueil, A.E. and Skladanowski, A. (2003) Catalytic topoisomerase II inhibitors in cancer therapy. *Pharmacol Ther*, **99**, 167-181.
47. Wu, N., Wu, X.W., Agama, K., Pommier, Y., Du, J., Li, D., Gu, L.Q., Huang, Z.S. and An, L.K. (2010) A novel DNA topoisomerase I inhibitor with different mechanism from camptothecin induces G2/M phase cell cycle arrest to K562 cells. *Biochemistry*, **49**, 10131-10136.
48. Aggarwal, M., Sommers, J.A., Shoemaker, R.H. and Brosh, R.M., Jr. (2011) Inhibition of helicase activity by a small molecule impairs Werner syndrome helicase (WRN) function in the cellular response to DNA damage or replication stress. *Proc Natl Acad Sci U S A*, **108**, 1525-1530.
49. Belitsky, J.M., Nguyen, D.H., Wurtz, N.R. and Dervan, P.B. (2002) Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg Med Chem*, **10**, 2767-2774.
50. Vichai, V. and Kirtikara, K. (2006) Sulforhodamine B colorimetric assay for cytotoxicity screening. *Nat Protoc*, **1**, 1112-1116.
51. Ishiyama, M., Shiga, M., Sasamoto, K., Mizoguchi, M. and He, P.G. (1993) A New Sulfonated Tetrazolium Salt That Produces a Highly Water-Soluble Formazan Dye. *Chem Pharm Bull*, **41**, 1118-1122.