

Chapter 5

Part A: Effects of Cationic Side Chains on Polyamide-DNA Interactions

5.1 BACKGROUND

DNA is composed of two antiparallel strands associated by hydrogen bonds between the four bases adenine: (A), thymine (T), cytosine (C), and guanine (G).¹ The most common structural form, B-DNA displays a wide and shallow major groove and a deep and narrow minor groove (**Figure 5.1**).² DNA bases recognize their pairing partner by specific, complementary hydrogen-bonding patterns (**Figure 5.2**). The major groove is a typical target for DNA binding, where DNA-binding proteins frequently form hydrogen bonds, van der Waals interactions, and electrostatic interactions with DNA.³ In contrast, polyamides represent a class of synthetic DNA-binding ligands that sequence specifically recognizes the minor groove of DNA.⁴



Figure 5.1. B-form Double Helix DNA. Antiparallel strands are in dark and light gray. (left) space filling CPK model, (right) ribbon representation.

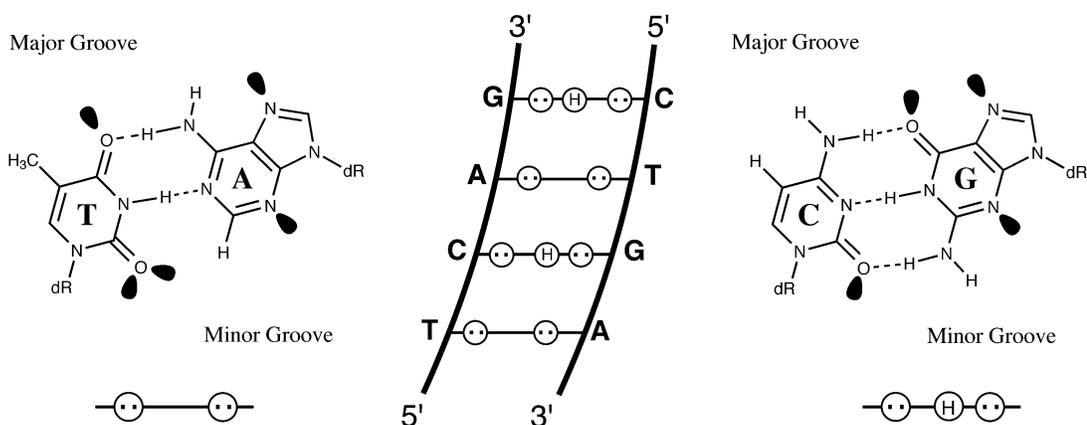


Figure 5.2. A Schematic Model of Minor Groove Recognition. Hydrogen bond donors represented by (H) and hydrogen bond acceptors as two dots.

Recent Advances in Polyamide Studies

Since this chapter was written in 2002, several recent advances have been made by Dervan and co-workers.⁵ For example, polyamide-intercalator conjugates have been shown to inhibit transcription factor-binding *in vitro*,⁶ extended polyamide dimers can now target longer DNA sequences, and U-pin and H-pin motifs are striving to target pure GC sequences.⁵ Importantly, researchers have now achieved nuclear uptake of polyamide conjugates, concluding that nuclear localization depends on polyamide-dye conjugate composition and on the mammalian cell line.^{5,7,8} In addition, genomic-wide analysis of the impact of polyamides on transcription regulation using GeneChips is being conducted.⁵

5.2 INTRODUCTION

Small molecules with the ability to specifically target any predetermined DNA sequence would be useful tools in molecular biology and, potentially, in human medicine. Polyamides containing *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methyl-3-

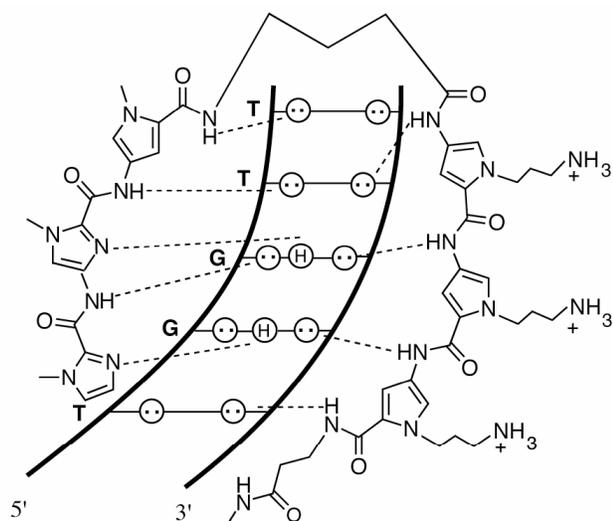
hydroxypyrrole (Hp) amino acids bind to specific predetermined sequences in the minor groove of DNA with affinities and specificities comparable to naturally occurring DNA-binding proteins.⁴ DNA recognition depends on side-by-side aromatic amino acid pairings oriented N to C with respect to the 5' to 3' direction of the DNA helix in the minor groove. A pairing of imidazole opposite pyrrole (Im/Py) recognizes a G•C base pair, while a Py/Im combination recognizes C•G.⁹⁻¹¹ An Hp/Py pair discriminates T•A over A•T.^{12,13} A Py/Py pair is degenerate and recognizes either A•T or T•A bases.^{10,14} Hairpin polyamides contain an alkyl amino acid, either γ -aminobutyric acid (γ) or amine-functionalized derivative (R)-2,4-diaminobutyric acid, which serves as the covalent linker region between the antiparallel strands and is specific for A•T or T•A base pairs.⁴

Polyamides offer a potentially general approach to gene regulation, provided that polyamide motifs can be developed to target a variety of DNA sequences with the subnanomolar affinities necessary to compete with DNA-binding proteins. The development of tools that enhance polyamide affinity for DNA sequences, while maintaining specificity, should prove beneficial.

The addition of a positively charged aminoalkyl moiety to several known synthetic DNA-binding ligands results in an increase in DNA binding affinity.¹⁵⁻¹⁸ Polyamides, traditionally monocationic or dicationic ligands, contain a dimethylaminopropylamine group at the C-terminus. Previous studies show the position exchange of the charged C-terminal dimethylaminopropyl (Dp) moiety with the methyl group at the N-1 position of a pyrrole residue enhances polyamide-DNA affinity by approximately 10-fold with only a moderate loss in specificity.¹⁹ Possible explanations of these results include the alleviation of steric strain between the Dp moiety and the floor of the minor groove and

the favorable electrostatic interactions of the cationic Dp side chain with the polyanionic DNA phosphate backbone.¹⁹ This type of general, affinity-enhancing modification may prove a powerful tool for several applications, including increasing the affinities of weak-binding polyamide-DNA complexes and developing polyamides with new functionality.

The current study explores the effects of incorporating multiple aminoalkyl side chains on the DNA recognition properties of hairpin polyamides (**Figure 5.3**). Derived from a previously studied six-ring hairpin ImImPy- γ -PyPyPy- β -Dp²⁰ (**1**), a series of compounds (**1-8**, **Figure 5.4**) was synthesized, incorporating variation in the number, relative spatial distribution, and linker length of aminoalkyl side chains. In addition, a polyamide containing alkyl guanidine side chains, compound **9** (**Figure 5.4**), was prepared to study the effects of charge distribution in the cationic head group. Quantitative DNase I footprint titration experiments were performed to compare the relative binding affinities of these compounds for their cognate match and mismatch sites, in order to probe the influence of multiple charges, charge distribution, and linker length on DNA binding-affinity and sequence selectivity.



ImImPy- γ -Py(C₃N)Py(C₃N)Py(C₃N)- β -Me

5' - A T G G T T -3'

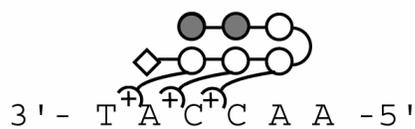


Figure 5.3. Binding Model of the Polyamide-DNA Complex between six-ring hairpin polyamide ImImPy- γ -Py(C₃N)Py(C₃N)Py(C₃N)- β -Me (**4**) and a 5'- ATGGTT -3' site. (top) Circles with two dots represent the lone pairs of N₃ purines and O₂ of pyrimidines. Circles containing an H represent the N₂ hydrogens of guanine. Putative hydrogen bonds are illustrated by dotted lines. (bottom) A ball and stick representation of polyamide **4** with DNA. Filled circles denote imidazole while open circles represent pyrrole. The diamond represents β -alanine, the curved line connecting two circles represents γ -aminobutyric acid, and the lines ending in plus signs represent aminopropyl side chains.

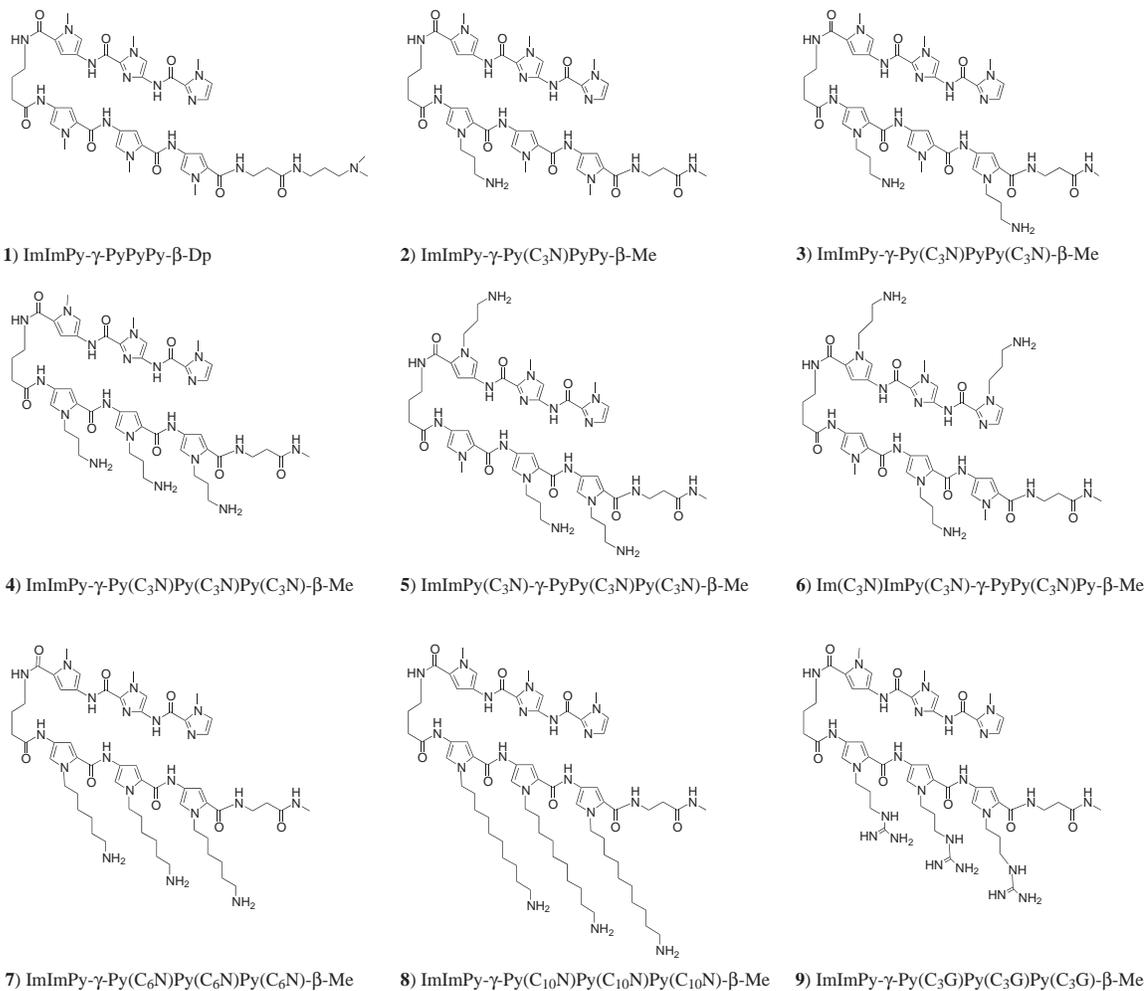


Figure 5.4. Structures of Polyamides 1-9.

5.3 RESULTS

Polyamide Synthesis

Polyamides **1-9** were synthesized using solid phase methods.²¹ Compound **6** incorporates a 1-(*N*-Boc-3-aminopropyl)-2-Im-CO₂H monomer, which allows the aminopropyl group to be located on the terminal imidazole.²² Polyamides **7** and **8**, containing aminoalkyl side chains of varying length, required the preparation of new *N*-aminohexyl and *N*-aminodecyl pyrrole monomers (**13** and **14**, **Figure 5.5a**). Monomers **13** and **14** were prepared from alkylation of silyl-protected pyrrole **10**²³ with *N*-(6-bromohexyl)phthalimide or *N*-(10-bromodecyl)phthalimide to yield the aminoalkyl pyrrole esters **11** and **12**, respectively. Removal of the trimethylsilylethyl group with tetrabutylammonium fluoride provided the functionalized monomers. To facilitate solid-phase synthesis of target polyamides, ImIm-CO₂Na dimer **17** was prepared from imidazole ester **16** in two steps (**Figure 5.5b**). Treatment of **16** with 2-(trichloroacetyl)-1-methylimidazole **15**²⁴ and base afforded an ester that was subsequently hydrolyzed to provide dimer **17**. Derivatives (**1-8**) of ImImPy- γ -PyPyPy- β -Dp²⁰ were synthesized in a stepwise manner from Boc- β -alanine-PAM resin using solid-phase methodology²¹ in 14 steps (**Figure 5.6**), followed by cleavage with methylamine and purification by reverse-phase HPLC. Polyamide **9** was prepared by perguanidinylation of **4** with excess pyrazole-1-carboxamide and sodium carbonate.²⁵

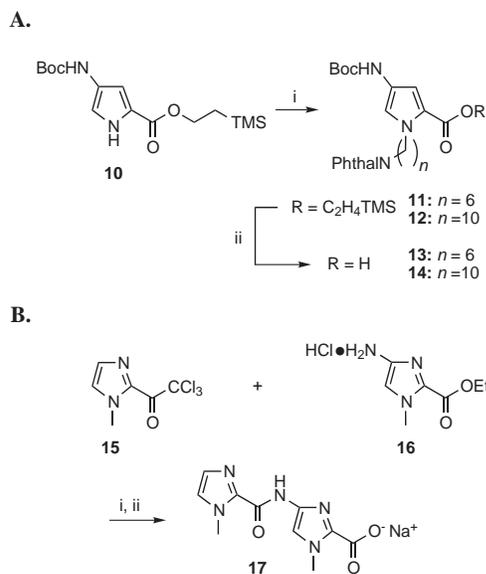


Figure 5.5. Monomer Synthesis. (A) Synthesis of Boc-Py(C_n N)OH monomers from **10**. i) *N*-(6-bromohexyl)phthalimide ($n = 6$) or *N*-(10-bromodecyl)phthalimide ($n = 10$), Bu_4NI , K_2CO_3 , 4 Å mol. sieves, CH_3CN ; ii) TBAF, THF, $0^\circ\text{C} \rightarrow \text{rt}$. (B) Synthesis of ImIm- CO_2Na dimer **17**. i) DIEA, EtOAc, 35°C ; ii) NaOH (aq), MeOH, 60°C .

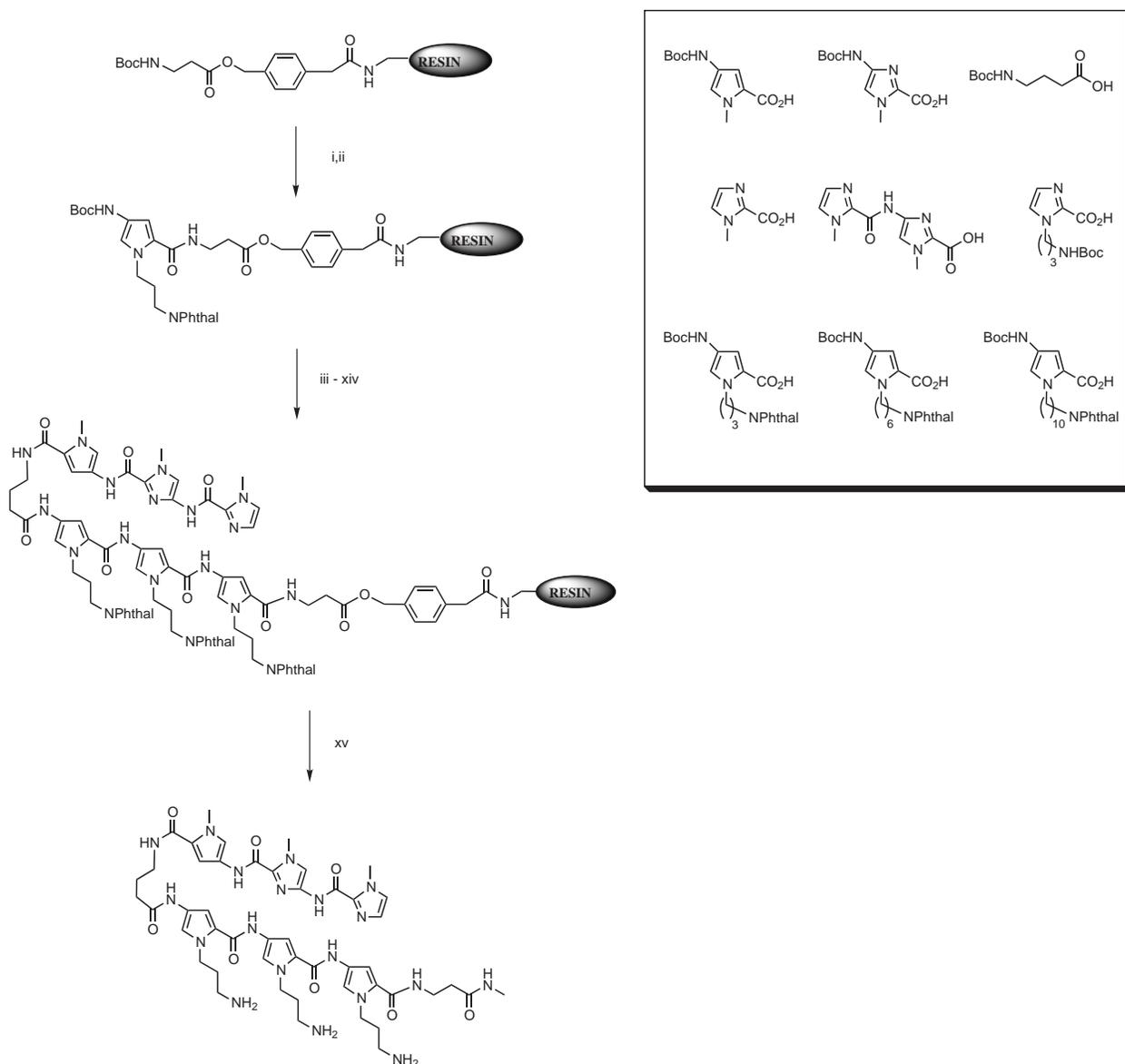


Figure 5.6. Solid-phase Synthetic Scheme exemplified for ImImPy- γ -Py(C₃N)Py(C₃N)Py(C₃N)- β -Me (**4**) starting from commercially available Boc- β -PAM resin: (i) TFA (ii) Boc-N-(3-phthalimidopropyl)pyrrole-acid, HOBT, DCC, DMF, DIEA. (iii) TFA (iv) Boc-N-(3-phthalimidopropyl)pyrrole-acid, HOBT, DCC, DMF, DIEA. (v) TFA (vi) Boc-N-(3-phthalimidopropyl)pyrrole-acid, HOBT, DCC, DMF, DIEA. (vii) TFA (viii) Boc- γ -aminobutyric-acid, HOBT, DCC, DMF, DIEA. (ix) TFA (x) Boc-pyrrole-OBt, DMF, DIEA. (xi) TFA (xii) Boc-Imidazole-acid, HOBT, DCC, DMF, DIEA. (xiii) TFA (xiv) Imidazole-acid, HBTU, DMF, DIEA. (xv) Methylamine, 55 °C, 12 h.

Quantitative DNase I Footprinting Titrations²⁶

Quantitative DNase I footprint titrations were conducted on the 5'-³²P-end-labeled 197 bp PCR product from pALC1 containing a 5'- TGGTT -3' match and a 5'- TGTAT -3' single base pair mismatch site (**Figure 5.7**). These experiments, performed for the six-ring hairpin polyamides **1-9** (**Figure 5.8**, **Figure 5.9**), revealed an increase in binding affinity through incorporation of aminopropyl side chains at the N-1 position of polyamide rings relative to the parent polyamide, **1**²⁰ (**Table 5.1**).

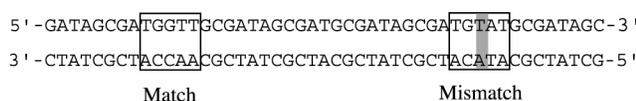


Figure 5.7. Portion of the 197-bp PCR Product Derived from pALC1. The targeted 5-bp match and single base pair mismatch recognition sites are indicated.

Polyamide **1** bound the match and mismatch sites with moderate affinity [$K_a = 1.3 \times 10^8 \text{ M}^{-1}$ and $K_a = 1.5 \times 10^6 \text{ M}^{-1}$, respectively] and excellent specificity, showing an 82-fold preference for the match site. Removal of the Dp tail and substitution of a single aminopropyl side chain in compound **2** increased the affinity approximately 10-fold relative to **1** [$K_a = 8.5 \times 10^8 \text{ M}^{-1}$] and revealed a 42-fold preference over the mismatch site [$K_a = 2.1 \times 10^7 \text{ M}^{-1}$]. The incorporation of a second aminopropyl side chain in compound **3** further increased the affinity to approximately 20-fold greater than **1** [$K_a = 2.7 \times 10^9 \text{ M}^{-1}$], with a 14-fold preference over the mismatch site [$K_a = 1.9 \times 10^8 \text{ M}^{-1}$]. Compound **4**, with three contiguous aminopropyl side chains, bound the match site with approximately 100-fold higher affinity [$K_a = 1.6 \times 10^{10} \text{ M}^{-1}$] than the parent and with a 15-fold

preference over the mismatch site [$K_a = 1.0 \times 10^9 \text{ M}^{-1}$]. In compounds **5** and **6**, derivatives of **4**, the spacing of the three aminopropyl side chains throughout the six-ring hairpin is increased, further enhancing binding affinity approximately 1000-fold relative to **1** [$K_a = 1.3 \times 10^{11} \text{ M}^{-1}$ and $1.6 \times 10^{11} \text{ M}^{-1}$, respectively] with an 11-fold preference over the mismatch site.

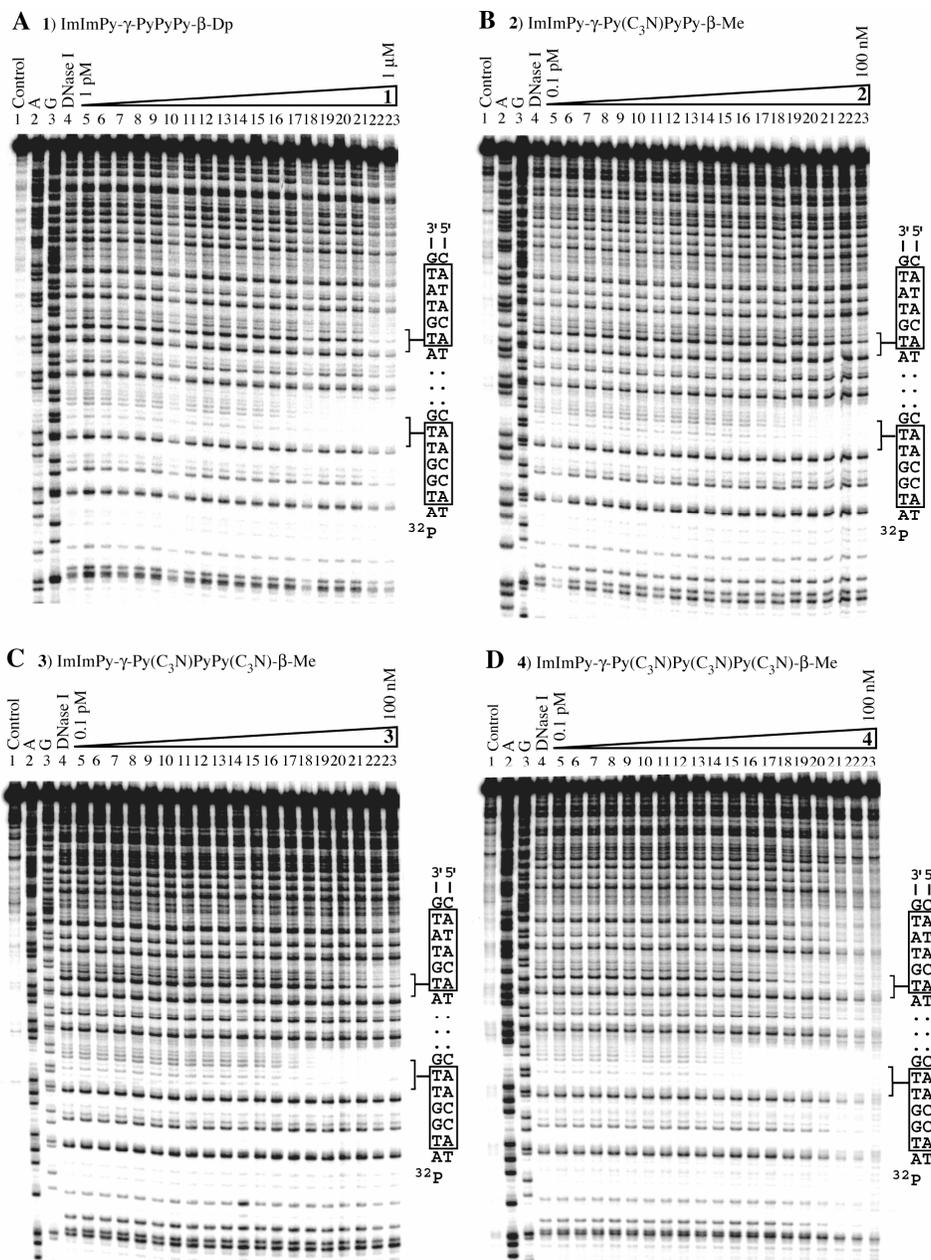


Figure 5.8. Quantitative DNase I Footprint Experiments with (A) ImImPy- γ -PyPyPy- β -Dp (1), (B) ImImPy- γ -Py(C₃N)PyPy- β -Me (2), (C) ImImPy- γ -Py(C₃N)PyPy(C₃N)- β -Me (3), and (D) ImImPy- γ -Py(C₃N)Py(C₃N)Py(C₃N)- β -Me (4) on the 5'-³²P-end-labeled 197 bp PCR product from pALC1. All reactions, performed at 22 °C, contained 10 kcpm restriction fragment 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂. For compound 1: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 μ M polyamide, respectively. For compounds 2-4: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 0.1 pM, 0.2 pM, 0.5 pM, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM polyamide, respectively. The positions of the match (5'-TGGTT-3') and single base pair mismatch (5'-TGTAT-3') sites are indicated.

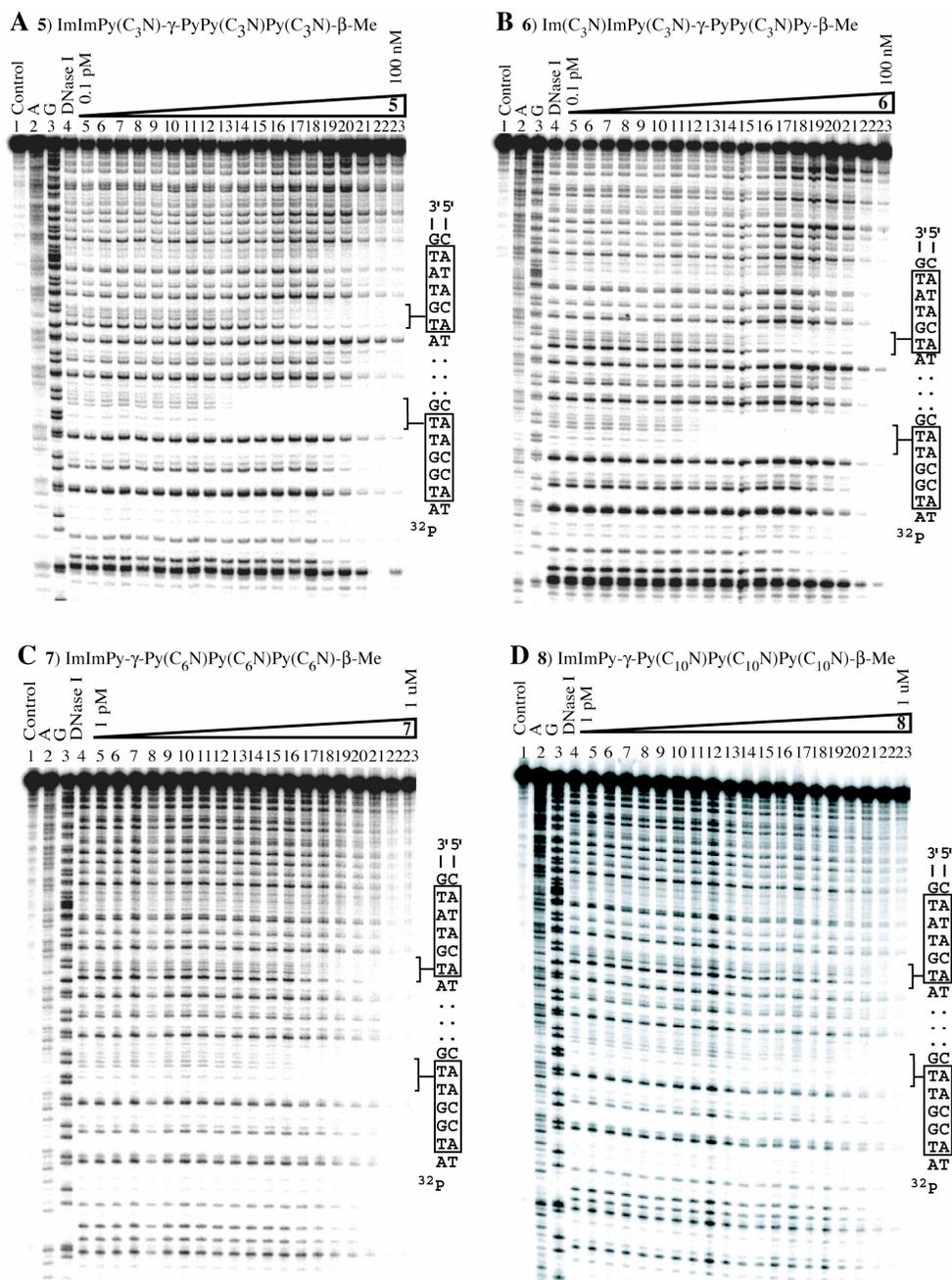


Figure 5.9. Quantitative DNase I Footprint Experiments with (A) ImImPy(C₃N)-γ-PyPy(C₃N)Py(C₃N)-β-Me (5), (B) Im(C₃N)ImPy(C₃N)-γ-PyPy(C₃N)Py-β-Me (6), (C) ImImPy-γ-Py(C₆N)Py(C₆N)Py(C₆N)-β-Me (7), and (D) ImImPy-γ-Py(C₁₀N)Py(C₁₀N)Py(C₁₀N)-β-Me (8) on the 5'-³²P-end-labeled 197 bp PCR product from pALC1. All reactions, performed at 22 °C, contained 10 kcpm restriction fragment 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂. For compounds 5 and 6: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 0.1 pM, 0.2 pM, 0.5 pM, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM polyamide, respectively. For compounds 7 and 8: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 μM polyamide, respectively. The positions of the match (5'- TGGTT -3') and single base pair mismatch (5'- TGTAT -3') sites are indicated.

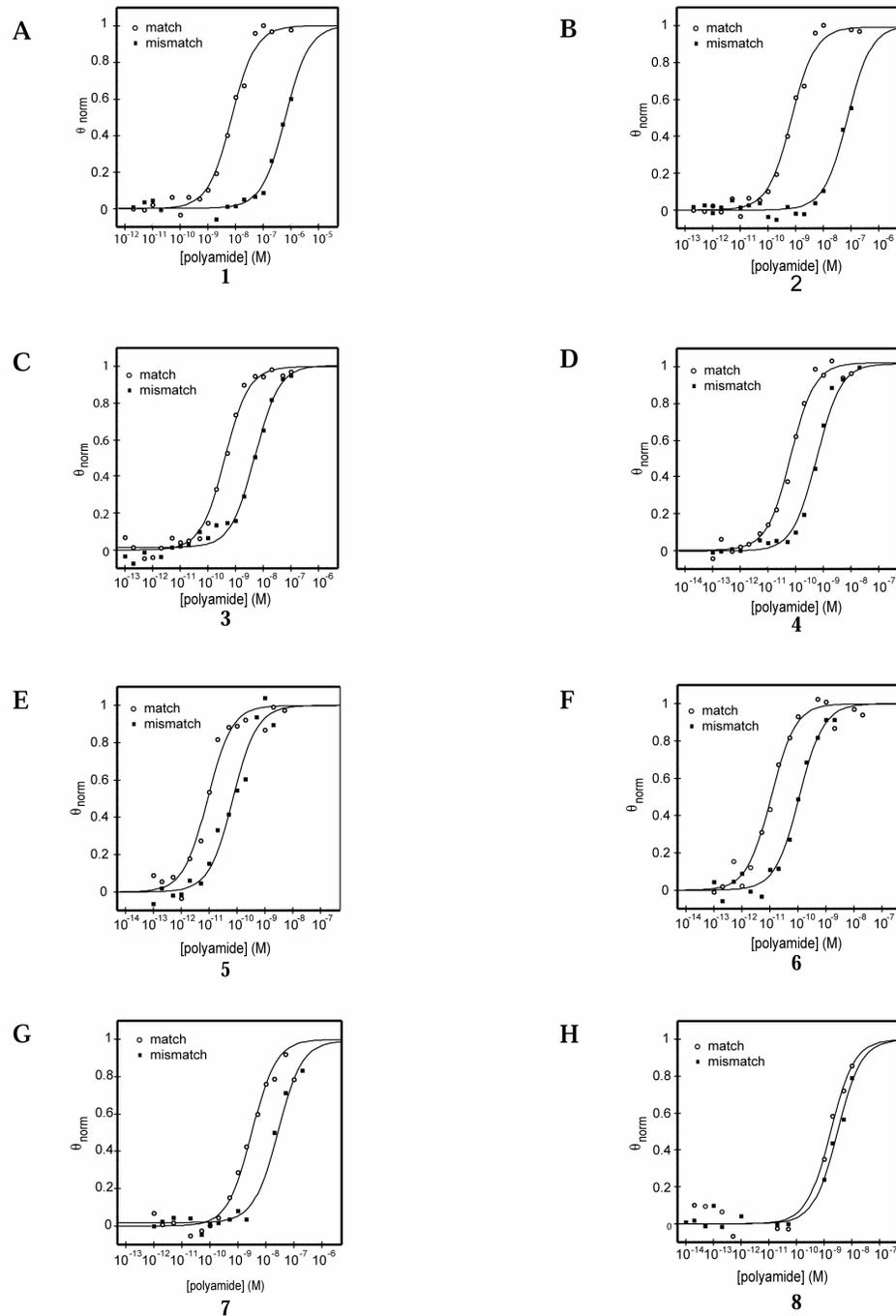


Figure 5.10. Binding Isotherms Derived from DNase Footprinting Gels on the match 5'-ATGGTT-3' and mismatch 5'-ATGTAT-3' sites. A) Polyamide 1. B) Polyamide 2. C) Polyamide 3. D) Polyamide 4. E) Polyamide 5. F) Polyamide 6. G) Polyamide 7. H) Polyamide 8.

Two additional derivatives of **4** were examined to determine the effects of cationic-side chain length on binding affinity. Binding characteristics of polyamides **7** and **8**, containing three contiguous aminohexyl and aminodecyl side chains respectively, were examined. Compound **7** bound with a decreased affinity for the match site [$K_a = 2.8 \times 10^8 \text{ M}^{-1}$] compared to **4** and with an approximately 10-fold preference over the mismatch site [$K_a = 4.3 \times 10^7 \text{ M}^{-1}$]. Compound **8** displayed poor binding with similar affinity for match and mismatch sites [$K_a \leq 6 \times 10^7 \text{ M}^{-1}$].

To examine the effect of size and charge distribution within the cationic head group, the perguanidinylated derivative of **4**, compound **9**, was synthesized. Polyamide **9** binds with high match site affinity [$K_a = 1.0 \times 10^{10} \text{ M}^{-1}$] and good specificity over the mismatch site [$K_a = 7.0 \times 10^{10} \text{ M}^{-1}$]. Perguanidinylated derivative **9** displays DNA-binding properties nearly identical to the aminopropyl-substituted parent **4**.

Polyamide	Match	Mismatch
	5'- TGGTT -3'	5'- TGTAT -3'
1) ImImPy- γ -PyPyPy- β -Dp	$1.3 (\pm 0.2) \times 10^8$	$1.5 (\pm 0.7) \times 10^6$
2) ImImPy- γ -Py(C ₃ N)PyPy- β -Me	$8.5 (\pm 0.9) \times 10^8$	$2.1 (\pm 0.7) \times 10^7$
3) ImImPy- γ -Py(C ₃ N)PyPy(C ₃ N)- β -Me	$2.7 (\pm 1.4) \times 10^9$	$1.9 (\pm 0.5) \times 10^8$
4) ImImPy- γ -Py(C ₃ N)Py(C ₃ N)Py(C ₃ N)- β -Me	$1.6 (\pm 0.4) \times 10^{10}$	$1.0 (\pm 0.6) \times 10^9$
5) ImImPy(C ₃ N)- γ -PyPy(C ₃ N)Py(C ₃ N)- β -Me	$1.3 (\pm 0.5) \times 10^{11}$	$1.2 (\pm 0.3) \times 10^{10}$
6) Im(C ₃ N)ImPy(C ₃ N)- γ -PyPy(C ₃ N)Py- β -Me	$1.1 (\pm 0.4) \times 10^{11}$	$1.0 (\pm 0.2) \times 10^{10}$
7) ImImPy- γ -Py(C ₆ N)Py(C ₆ N)Py(C ₆ N)- β -Me	$2.8 (\pm 0.7) \times 10^8$	$4.3 (\pm 1.8) \times 10^7$
8) ImImPy- γ -Py(C ₁₀ N)Py(C ₁₀ N)Py(C ₁₀ N)- β -Me	$\leq 6 \times 10^7$	$\leq 6 \times 10^7$
9) ImImPy- γ -Py(C ₃ G)Py(C ₃ G)Py(C ₃ G)- β -Me	$1.0 (\pm 0.2) \times 10^{10}$	$7.0 (\pm 2.6) \times 10^8$

Table 5.1. Equilibrium Association Constants, K_a (M^{-1}). Values reported are the mean values obtained from at least three DNase I footprint titration experiments with standard deviation for each data set in parentheses. The assays were carried out at 22°C at pH 7.0 in the presence of 10 mM Tris•HCl, 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂.

5.4 DISCUSSION

Incorporation of Multiple Aminopropyl Side Chains

Increasing the number of aminopropyl side chains incorporated into a six-ring hairpin results in a corresponding enhancement of DNA-binding affinity. Previous studies demonstrate that moving the dimethylaminopropyl group from the C-terminal tail of a polyamide to the N-1 position of an aromatic ring residue enhances the polyamide-DNA match site binding affinity, while only moderately decreasing specificity. The current study reveals a dramatic affinity-enhancing trend for the incremental incorporation of multiple primary aminopropyl substituents. The placement of an aminopropyl group on a single pyrrole residue in **2** results in an approximately 10-fold increase in binding affinity, while compound **4**, with three contiguous charges, displays a further affinity enhancement—approximately 100-fold over the parent. These observed affinity enhancements may result from favorable electrostatic interactions between the aminopropyl groups and phosphate oxygens of the DNA backbone. The incremental incorporation of multiple aminopropyl side chains results in an exponential enhancement of polyamide affinity, while maintaining good specificity.

Effects of Cationic Side Chain Spacing

The next series of compounds examines the effects of distributing three aminopropyl residues throughout a six-ring polyamide. Polyamide **4** contains three contiguous cationic side chains, while **5** and **6** contain an increased spatial distribution of the three side chains. The spacing of aminopropyl side chains in compounds **5** and **6** enhances the match-site affinity approximately 10-fold relative to compound **4**, while maintaining

specificity. The distribution of three aminopropyl side chains throughout a polyamide results in an approximately 1000-fold affinity enhancement relative to the parent polyamide **1**.

Impact of Aminoalkyl Chain Length

Compounds **7** and **8**, extended linker derivatives of compound **4**, were employed to probe the impact of alkyl-linker length on DNA-binding affinity and specificity. Polyamides containing three contiguous aminopropyl (**4**), aminohexyl (**7**), and aminodecyl (**8**) side chains display a dramatic decrease in affinity with increasing alkyl side chain length. The aminohexyl side chains of compound **7** act to decrease the affinity to that of the parent **1**, while maintaining good specificity. The aminodecyl side chains of compound **8** confer poor binding properties, further decreasing the match site affinity to a value below that of the parent (**1**). These unfavorable effects seen with increased alkyl chain lengths of six and ten methylene units may be attributed to the increased hydrophobicity of the longer alkyl chains. Therefore, aminopropyl side chains provide a more optimal linker length affording polyamides with high affinity and good specificity for the DNA match site.

Effects of Cation Alteration

Converting primary amines to guanidines alters the size and the charge distribution within the cationic head group. To probe the DNA-binding properties of a polyamide containing multiple guanidines, *N*-aminopropyl-substituted polyamide **4** was perguanidinylated to provide compound **9**. Polyamide **9** displays DNA-binding affinity

and specificity nearly identical to **4**. The DNA-binding properties of the extended linker derivatives **7** and **8** were also unaltered by guanidinylation (see Chapter 2). Therefore, converting amines to guanidines, despite the difference in size and charge density, does not appear to significantly alter polyamide-DNA interactions. This observation provides opportunities for exploring new polyamide motifs with potentially interesting cellular uptake properties, since certain classes of guanidinylated small molecules have been shown to traffic to the interior of cells.²⁵

5.5 CONCLUSIONS

The incorporation of multiple aminopropyl side chains to hairpin polyamides has been shown to efficiently enhance polyamide-DNA affinity, while maintaining good specificity. With proper spatial distribution, multiple aminopropyl side chains can be incorporated to dramatically enhance polyamide-DNA affinity by approximately 1000-fold. This general modification may prove useful for future applications, including enhancement of affinities for weakly binding polyamide-DNA complexes and the development of polyamides with new functions. This new class of hairpin polyamides may display potentially interesting cellular uptake properties as well. The proposed electrostatic interactions of multiple polyamide cationic side chains with the DNA phosphate backbone may neutralize a portion of DNA, causing DNA to relax and bend toward the minor groove. As well, these electrostatic interactions may potentially function to inhibit DNA binding of major groove proteins through competition with

protein side chains for electrostatic contacts or through DNA deformation caused by charge neutralization of the DNA backbone.

5.6 MATERIALS AND METHODS

General

All synthetic reagents were prepared as previously described²¹ or obtained from Aldrich or Trans World Chemicals. Semi-automated synthesis was performed on a Quest 210 (Argonaut Technologies Inc.) manual synthesizer. HPLC analysis was performed on a Beckman *Gold Nouveau* system using a RAINEN C₁₈, Microsorb MV, 5 μm, 300 x 4.6 mm reverse-phase column in 0.1 % (*wt v⁻¹*) TFA:H₂O with acetonitrile as eluent and a flow rate of 1.0 mL min⁻¹, gradient elution 1.25 % acetonitrile min⁻¹. Preparatory reverse-phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 x 100 mm, 100 μm C₁₈ column equipped with a guard, 0.1 % (*wt v⁻¹*) TFA, 0.25 % acetonitrile min⁻¹. Milli-Q water was obtained from a Millipore Milli-Q water purification system, and all buffers were 0.2 μm filtered.

Monomer Synthesis

1-(*N*-Boc-3-aminopropyl)-2-imidazole carboxylate. Prepared as previously described.²² Di-*tert*-butyl dicarbonate (17.5 g, 80.0 mmol) and K₂CO₃ (20% aq., 100 mL) were added to a solution of 1-(3-aminopropyl)imidazole in EtOAc (120 mL). After 8 h stirring at room temperature, the organic layer was separated and washed with Na₂CO₃ (sat. aq., 50 mL) and brine (50 mL), and then evaporated *in vacuo* affording a colorless oil (15.1 g, 83% yield). To a cooled (-72 °C) solution of 1-(*N*-(*t*-

Butoxycarbonyl)-3-aminopropyl)imidazole (13.8 g, 61.3 mmol) in dry THF (150 mL) under Ar was added *n*-Butyllithium (61 mL, 2.0 M in hexanes). After 1.5 h stirring, CO₂ (g) was bubbled through the reaction mixture for 2 h. The reaction was acidified with HCl (1 N aq., 61 mL) to pH 8. The aqueous layer was separated, frozen, and lyophilized to yield a white solid (11.4 g). The crude solid was purified on reverse phase C₁₈ silica gel with 0.1% (*w/v*) TFA and acetonitrile to yield a white solid (5.7 g, 35% yield). ¹H NMR (DMSO-*d*₆) δ 7.16 (s, 1H), 7.01 (bt, 1H), 6.82 (s, 1H), 4.45 (t, 2H), 2.82 (d of d, 2H), 1.75 (m, 2H), 1.37 (s, 9H).

(2-Trimethylsilyl)ethyl 4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidohexyl)-pyrrole-2-carboxylate (11). To a solution of **10**²³ (8.1 g, 24.8 mmol) in anhydrous acetonitrile (41 mL) were added activated molecular sieves (3 Å, finely ground, 1.5 g), K₂CO₃ (finely ground, 5.14 g, 37.2 mmol), tetrabutylammonium iodide (1.83 g, 4.96 mmol), and *N*-(6-bromohexyl)phthalimide (10 g, 32.2 mmol). The resulting suspension was stirred at 90 °C for 45 h. Chloroform (300 mL) was added, and the mixture was filtered on a fritted funnel. This solution was washed with water (3 x 400 mL) and brine (400 mL), and then dried (MgSO₄), filtered, and evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel (elution with 20% EtOAc-hexanes) affording **11** (12.4 g, 22.2 mmol, 90% yield) as a yellow oil. ¹H NMR (DMSO-*d*₆) δ 9.10 (s, 1H), 7.82 (m, 4H), 7.09 (s, 1H), 6.60 (s, 1H), 4.18 (m, 4H), 3.52 (t, 2H, *J* 7.2), 1.55 (m, 4H), 1.41 (s, 9H), 1.22 (m, 4H), 0.98 (t, 2H, *J* 8.3), 0.00 (s, 9H); ¹³C NMR (DMSO-*d*₆) δ 167.7, 160.0, 152.5, 134.2, 131.5, 123.0, 122.8, 118.0 (m, 2C), 107.5, 78.4, 61.3,

47.9, 37.3, 31.1, 28.2, 27.9, 25.9, 25.6, 17.0, -1.3; HRMS Calcd for M⁺ (C₂₉H₄₁N₃O₆Si): 555.2765, found 555.2782.

(2-Trimethylsilyl)ethyl 4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidodecyl)pyrrole -2-carboxylate (12). To a solution of **10**²³ (3.42 g, 10.5 mmol) in anhydrous acetonitrile (50 mL) were added activated molecular sieves (3 Å, finely ground, 700 mg), K₂CO₃ (finely ground, 2.18 g, 15.8 mmol), tetrabutylammonium iodide (776 mg, 2.1 mmol), and N-(10-bromodecyl)phthalimide (5 g, 13.6 mmol). The resulting suspension was stirred at 90 °C for 26 h, and then additional N-(10-bromodecyl)phthalimide (2.7 g, 7.4 mmol) was added. After an additional 40 h stirring at reflux, the mixture was concentrated by removing solvent (30 mL) via a Dean-Stark trap, with subsequent stirring at 90 °C for an additional 16 h. Chloroform (100 mL) was added, and the mixture was filtered on a fritted funnel. This solution was washed with water (2 x 100 mL) and brine (50 mL), and then dried (MgSO₄), filtered, and evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel (elution with 15% EtOAc-hexanes) affording **12** (4.31 g, 7.04 mmol, 66% yield) as a yellow oil. ¹H NMR (DMSO-*d*₆) δ 9.09 (s, 1H), 7.82 (m, 4H), 7.07 (s, 1H), 6.58 (s, 1H), 4.20 (t, 2H, *J* 8.2), 4.18 (t, 2H, *J* 7.1), 3.52 (t, 2H, *J* 7.1), 1.54 (m, 4H), 1.41 (s, 9H), 1.22-1.17 (m, 12H), 0.98 (t, 2H, *J* 8.2), 0.00 (s, 9H); ¹³C NMR (DMSO-*d*₆) δ 167.7, 160.0, 152.5, 134.2, 131.4, 123.0, 122.9, 118.1, 118.0, 107.4, 78.4, 61.2, 47.9, 37.4, 31.2, 28.9, 28.8, 28.6 (m, 2C), 28.2, 27.9, 26.3, 25.9, 17.0, -1.3; HRMS Calcd for M⁺ (C₃₃H₄₉N₃O₆Si): 611.3391, found 611.3396.

4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidohexyl)pyrrole-2-carboxylic acid (13). Molecular sieves (3 Å, 5 g) were added to a cooled (0 °C) solution of **11** (12.4 g, 22.2 mmol) in anhydrous THF (90 mL). Tetrabutylammonium fluoride (1 M, 45 mL, 45 mmol) was then added via syringe under a positive pressure of Ar, and the solution was stirred 20 min before being allowed to slowly warm to room temperature. After an additional 12 h, the reaction was quenched with 10% aqueous citric acid (60 mL), and then filtered through a fritted funnel. The filtrate was extracted with EtOAc (3 x 200 mL), and the combined organic layers were washed with brine (200 mL), dried (MgSO₄), filtered, and evaporated *in vacuo*. The residue was purified by flash chromatography on silica gel (elution with 1:20:80 AcOH-EtOAc-hexanes) and azeotroped with benzene (3 x 100 mL) affording **13** (6.34 g, 13.9 mmol, 79% yield) as a yellow solid. ¹H NMR (DMSO-*d*₆) δ 12.02 (s, 1H), 9.05 (s, 1H), 7.83 (m, 4H), 7.05 (s, 1H), 6.55 (s, 1H), 4.16 (t, 2H, *J* 7.2), 3.52 (t, 2H, *J* 7.2), 1.56 (m, 4H), 1.41 (s, 9H), 1.22 (m, 4H); ¹³C NMR (DMSO-*d*₆) δ 167.7, 161.4, 152.5, 134.2, 131.5, 122.8, 122.7, 118.7, 117.6, 107.7, 78.3, 47.7, 37.3, 31.1, 28.2, 27.9, 25.9, 25.6; HRMS Calcd for M⁺ (C₂₄H₂₉N₃O₆): 455.2056, found 455.2061.

4-[(*t*-Butoxycarbonyl)amino]-1-(phthalimidodecyl)pyrrole-2-carboxylic acid (14). Molecular sieves (3 Å, 2 g) were added to a cooled (0 °C) solution of **12** (4.31 g, 7.04 mmol) in anhydrous THF (32 mL). Tetrabutylammonium fluoride (1 M, 10.6 mL, 10.6 mmol) was then added via syringe under a positive pressure of Ar, and the solution was stirred 30 min before being allowed to slowly warm to room temperature. After 10 h, more tetrabutylammonium fluoride (1 M, 10.6 mL, 10.6 mmol) was added. After an

additional 4 h, the reaction was quenched with 10% aqueous citric acid (35 mL), then filtered through a fritted funnel. The filtrate was extracted with EtOAc (3 x 100 mL), and the combined organic layers were washed with brine (100 mL), dried (MgSO₄), filtered, and evaporated *in vacuo*. The resulting yellow oil was purified by flash chromatography on silica gel (elution with 1:15:85 AcOH-EtOAc-hexanes) and azeotroped with benzene (3 x 50 mL) affording **14** (2.55 g, 4.98 mmol, 66% yield) as a light yellow solid. ¹H NMR (DMSO-*d*₆) δ 9.05 (s, 1H), 7.83 (m, 4H), 7.05 (s, 1H), 6.55 (s, 1H), 4.16 (t, 2H, *J* 7.2), 3.53 (t, 2H, *J* 7.2), 1.57 (m, 4H), 1.41 (s, 9H), 1.19 (m, 12H); ¹³C NMR (DMSO-*d*₆) δ 167.7, 161.5, 152.5, 134.2, 131.5, 122.9, 122.7, 118.7, 117.6, 107.6, 78.3, 47.8, 37.4, 31.2, 28.9, 28.8, 28.6, 28.5, 28.2, 27.9, 26.3, 26.0; HRMS Calcd for M⁺ (C₂₈H₃₇N₃O₆): 511.2682, found 511.2688.

Sodium 1-Methyl-4-[(1-methylimidazol-2-yl)carbonylamino]imidazole-2-carboxylate (17). To a solution of 2-(trichloroacetyl)-1-methylimidazole **15**²⁴ (10 g, 44.0 mmol) in EtOAc (104 mL) was added ethyl 4-amino-1-methylimidazole-2-carboxylate hydrochloride **16**²¹ (7.5 g, 36.6 mmol) followed by DIEA (6.4 mL, 36.6 mmol), and the mixture was stirred at 35 °C for 9 h. After standing at 4 °C for 12 h, the resulting precipitate was collected by vacuum filtration, and the filtrate was poured into cold (0 °C) water (200 mL). The resulting precipitate was collected by vacuum filtration, and the combined solid material was washed with cold water (2 x 20mL) and then dried *in vacuo* to provide crude ethyl 1-methyl-4-[(1-methylimidazol-2-yl)carbonylamino]imidazole-2-carboxylate. ¹H NMR (DMSO-*d*₆) δ 10.13 (s, 1H), 7.69 (s, 1H), 7.43 (s, 1H), 7.06 (s, 1H), 4.28 (q, 2H, *J* 7.2), 3.98 (s, 3H), 3.94 (s, 3H), 1.29 (t, 3H, *J* 7.2). To the crude ester

dissolved in MeOH (40 mL) was added NaOH (1 M, 40 mL, 40 mmol), and the resulting mixture was stirred for 11 h at 60 °C. After standing at -20 °C for 1 h, the resulting precipitate was collected by vacuum filtration and then dried *in vacuo*, affording **17** (6.33 g, 27.4 mmol, 64% yield) as an off-white powder. ¹H NMR (DMSO-*d*₆) δ 10.47 (s, 1H), 7.42 (s, 1H), 7.24 (s, 1H), 7.05 (s, 1H), 4.00 (s, 3H), 3.93 (s, 3H); ¹³C NMR (DMSO-*d*₆) δ 161.7, 155.7, 140.6, 138.3, 134.5, 127.1, 126.6, 110.6, 35.3, 35.0.

Polyamide Synthesis

Reagents and protocols for solid-phase polyamide synthesis were as previously described.²¹ The crude polyamides were redissolved in 1:3 acetonitrile-0.1% (*wt v⁻¹*) TFA-H₂O (8 mL), filtered to remove resin, and purified by reverse-phase preparatory HPLC to yield an orange powder upon lyophilization.

Polyamides 2-6. Polyamides **2-6** were synthesized using semi-automated (Quest 210, Argonaut Technologies Inc.) solid-phase protocols from Boc-β-PAM resin (200-300 mg, 0.26-0.59 mmol g⁻¹). Resin was cleaved with methylamine in a Parr apparatus (55 °C, 12 h). After allowing the methylamine to evaporate, compounds were redissolved and purified as above. **ImImPy-γ-PyPyPy(C₃N)-β-Me. (2).** 4% recovery. MALDI-TOF-MS Calcd for C₄₄H₅₅N₁₇O₈ (M+H): 950.4, found 950.5. **ImImPy-γ-Py(C₃N)PyPy(C₃N)-β-Me. (3).** 3% recovery. MALDI-TOF-MS Calcd for C₄₆H₆₀N₁₈O₈ (M+H): 993.5, found 993.6. **ImImPy-γ-Py(C₃N)Py(C₃N)Py(C₃N)-β-Me. (4).** 2% recovery. MALDI-TOF-MS Calcd for C₄₈H₆₅N₁₉O₈ (M+H): 1036.5, found 1036.6. **ImImPy(C₃N)-γ-PyPy(C₃N)Py(C₃N)-β-Me. (5).** 9% recovery. MALDI-TOF-MS Calcd

for $C_{48}H_{65}N_{19}O_8$ (M+H): 1036.5, found 1036.7. **Im(C₃N)ImPy(C₃N)- γ -PyPy(C₃N)Py- β -Me (6)**. 5% recovery. MALDI-TOF-MS Calcd for $C_{48}H_{65}N_{19}O_8$ (M+H): 1036.5, found 1036.6.

Polyamides 7, 8. Polyamides **7** and **8** were synthesized using manual solid phase synthesis protocols from Boc- β -PAM-resin (250 mg, 0.25 mmol g⁻¹). Resin was swelled with DMF (0.5 mL) and cleaved with methylamine (5.6 M in MeOH) for 12 h at 22 °C, followed by rotary evaporation of solvent. Compounds were redissolved and purified as above. **ImImPy- γ -Py(C₆N)Py(C₆N)Py(C₆N)- β -Me (7)**. 3% recovery. MALDI-TOF-MS Calcd for $C_{57}H_{83}N_{19}O_8$ (M+Na): 1184.7, found 1184.9. **ImImPy- γ -Py(C₁₀N)Py(C₁₀N)Py(C₁₀N)- β -Me (8)**. 0.6% recovery. MALDI-TOF-MS Calcd for $C_{69}H_{107}N_{19}O_8$ (M+H): 1330.9, found 1331.1.

Polyamide 9. To polyamide **4** (385 nmol) was added Na₂CO₃ (150 mM aq., 155 μ L), and pyrazole-1-carboxamide•HCl (300 mM aq., 155 μ L), and the solution was heated at 50 °C for 15 h. The crude mixture was then diluted with acetonitrile (1 mL) and 0.1% (wt/v) TFA (4 mL), and the resulting solution was purified by preparative HPLC affording **ImImPy- γ -Py(C₃G)Py(C₃G)Py(C₃G)- β -Me (9)**. (0.2 mg, 44% yield). MALDI-TOF-MS Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5.

Quantitative DNase I footprinting. As previously described.²⁶ We note explicitly the final buffer concentrations: Tris-HCl buffer (10 mM, pH 7.0), KCl (10 mM), MgCl₂ (10 mM), CaCl₂ (5 mM), and 10 kcpm 5'-radiolabeled DNA.

Construction of plasmid DNA and 5' end-labeling.²⁶ Oligodeoxynucleotides were synthesized by DNA synthesis facility at the California Institute of Technology. Plasmid pALC1 was constructed by hybridization of a complementary set of synthetic oligonucleotides: **A)** 5'- GATCGCGATAGCGAGCTCAGCGATAGCGATGCGATAGCGATGGTTGCGATAGCGATGCGATAGCGATGTATGCGATAGCGAT-3', **B)** 5'- CTATCGCATAACATCGCTATCGCATCGCTATCGCAACCATCGCTATCGCATCGCTATCGCTGAGCTCGCTATCGC-3', **C)** 5'- GCGATAGCGTAGGGTGGCGATAGCGATGCGATAGCGATAGCGAGCTCAGCGATAGCGAGCTCAGCGATAGCGATGCGTGCA-3', **D)** 5'- CGCATCGCTATCGCTGAGCTCGCTATCGCTGAGCTCGCTATCGCTGAGCTCGCTATCGCTATCGCATCGCTATCGCACCTACGCTATCGCATCG-3'. Oligonucleotides **B)** and **C)** were phosphorylated with dATP and T4 PNK, and then annealed to their respective complementary strands **A)** and **D)**. The two sets of oligonucleotides were ligated and inserted into *XbaI/PstI*-linearized pUC19 using T4 DNA ligase. Plasmid purification was performed with Wizard Plus Midipreps DNA purification system. Fluorescent sequencing, performed at the DNA-sequencing facility at the California Institute of Technology, was used to verify the presence of the desired insert. Two 21-mer primers were synthesized for PCR amplification: forward primer 5'- TTCGAGCTCGGTACCCGGGGA -3' and reverse primer 5'- AGCTTGCATGCCTGCACGCAT -3'. PCR amplification generated the 5' end-labeled 197 bp product from pALC1.

Chapter 5

Part B: Effects of Cationic Side Chains on Nuclear Uptake

5.7 INTRODUCTION

Cell-permeable small molecules with the ability to specifically target a predetermined DNA sequence would be useful tools in molecular biology and medicine. Polyamides containing *N*-methylpyrrole (Py), *N*-methylimidazole (Im), and *N*-methyl-3-hydroxypyrrole (Hp) amino acids bind to specific predetermined sequences in the minor groove of DNA with affinities and specificities comparable to naturally occurring DNA-binding proteins.⁴ DNA recognition depends on side-by-side aromatic amino acid pairings oriented N to C with respect to the 5' to 3' direction of the DNA helix in the minor groove. A pairing of imidazole opposite pyrrole (Im/Py) recognizes a G•C base pair, while a Py/Im combination recognizes C•G.⁹⁻¹¹ An Hp/Py pair discriminates T•A over A•T.^{12,13} A Py/Py pair is degenerate and recognizes either A•T or T•A bases.^{10,14} Hairpin polyamides contain an alkyl amino acid, either γ -aminobutyric acid (γ) or amine-functionalized derivative (R)-2,4-diaminobutyric acid ($H_2N\gamma$), which serves as the covalent linker region between the antiparallel strands and is specific for A•T or T•A base pairs.⁴

Polyamides offer a potentially general approach to gene regulation. In cell-free systems, polyamides have been shown to inhibit several classes of transcription factors²⁷ as well as to induce gene activation.^{28,29} Current efforts focus on examining the potential involvement of polyamides in gene regulation in cell culture assays.³⁰⁻³² Recent studies utilize confocal microscopy to examine the cellular uptake properties of polyamide-Bodipy dye conjugates in a variety of living cells.³³ Preliminary studies demonstrate nuclear localization in CEM human T-cells, consistent with previous success in cell culture assays. The majority of live cells tested, however, do not exhibit nuclear

localization. Instead, the polyamide conjugates localize mainly in the cytoplasm.³³ These results indicate that nuclear uptake is an obstacle for *in vivo* polyamide gene regulation. Therefore, achieving nuclear localization of polyamides remains a key focus.

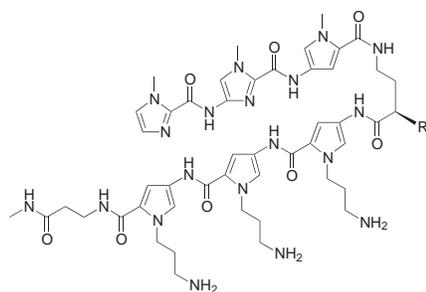
Several techniques have been developed to enable cellular uptake of drugs or molecular probes. In particular, short peptides of positively charged residues such as lysine and arginine have been shown to effect nuclear localization of natural proteins.³⁴ For example, a short peptide sequence from the HIV-Tat protein, TAT₄₉₋₅₇ (RKKRRQRRR), has been shown to successfully deliver functional biomolecules into cells. Wender and co-workers probed short arginine-rich molecular transporters derived from TAT₄₉₋₅₇, including polyguanidine peptoid derivatives composed of oligo-glycine backbones and arginine-like side chains on the amide nitrogen.²⁵ Interestingly, guanidine peptoids containing a six-methylene linker between the cationic head group and the backbone, rather than a three-methylene linker, display enhanced cellular uptake properties.

The current study explores these broadly applicable cellular uptake approaches by investigating polyamides substituted with cationic primary amine and guanidine side chains with varying linker lengths at the N-1 position of pyrrole rings. Derived from the previously studied six-ring hairpin ImImPy- γ -PyPyPy- β -Dp²⁰ (**1**), a series of compounds (**4** and **7-11**, Figure 1) was synthesized for DNase I footprinting experiments. Polyamide-Bodipy conjugates (**4B** and **7B-11B**) were synthesized for confocal microscopy experiments. The aminoalkyl side chain lengths and cationic head groups were varied to study the effects of these alterations on DNA binding and cellular uptake. Quantitative DNase I footprint titration experiments were performed to compare the

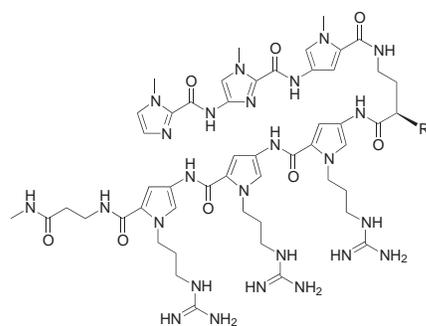
relative binding affinities of these compounds for their cognate match and mismatch sites, in order to probe the influence of these modifications, while confocal microscopy was utilized to examine cellular uptake in a variety of cell types.

Recent Advances in Polyamide Nuclear Uptake

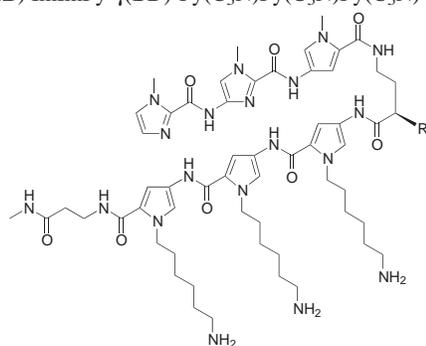
Since this chapter was written in 2002, several recent advances have been made towards achieving nuclear uptake of polyamide conjugates.^{7,8} Work by Bashkin and co-workers discovered that conjugation with Bodipy may contribute to the localization of the polyamide-BODIPY conjugates into vesicles. Instead, they discovered conjugation with fluorescein onto polyamides enabled polyamide-fluorescein conjugates to localize to the nucleus when in the presence of verapamil, a P-glycoprotein inhibitor.³⁵ Best and Edelson in the Dervan lab soon after discovered that specific combinations of polyamide linkers and fluorescein dyes enable nuclear uptake in the absence of verapamil.⁷ Researchers concluded that the nuclear uptake depends on polyamide composition and the mammalian cell line.⁸ A noteworthy positive factor for nuclear localization is a cationic amino alkyl moiety on the hairpin turn of polyamide-fluorescein conjugates.⁸



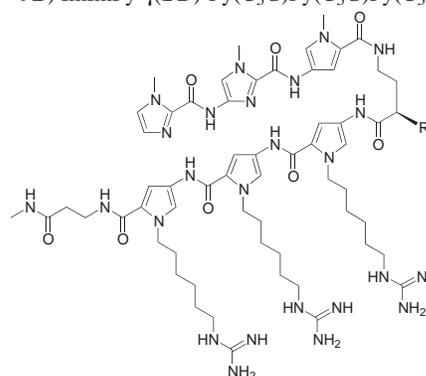
4) ImImPy- γ -Py(C₃N)Py(C₃N)Py(C₃N)- β -Me
4B) ImImPy- γ (BD)-Py(C₃N)Py(C₃N)Py(C₃N)- β -Me



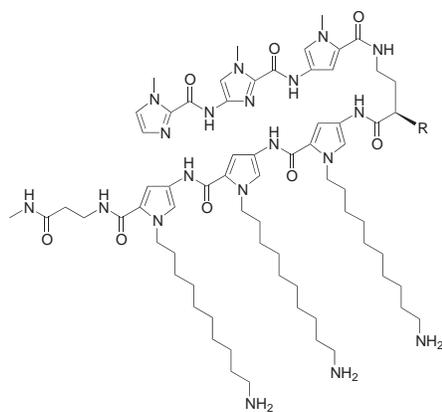
9) ImImPy- γ -Py(C₃G)Py(C₃G)Py(C₃G)- β -Me
9B) ImImPy- γ (BD)-Py(C₃G)Py(C₃G)Py(C₃G)- β -Me



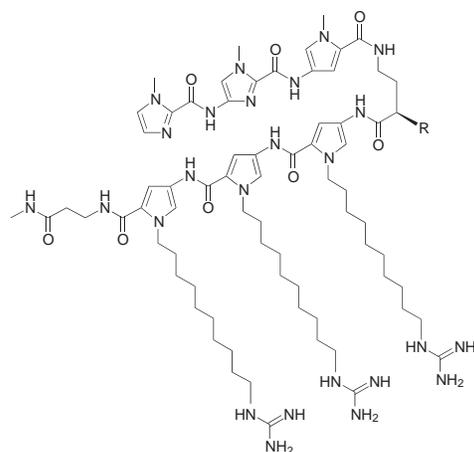
7) ImImPy- γ -Py(C₆N)Py(C₆N)Py(C₆N)- β -Me
7B) ImImPy- γ (BD)-Py(C₆N)Py(C₆N)Py(C₆N)- β -Me



10) ImImPy- γ -Py(C₆G)Py(C₆G)Py(C₆G)- β -Me
10B) ImImPy- γ (BD)-Py(C₆G)Py(C₆G)Py(C₆G)- β -Me



8) ImImPy- γ -Py(C₁₀N)Py(C₁₀N)Py(C₁₀N)- β -Me
8B) ImImPy- γ (BD)-Py(C₁₀N)Py(C₁₀N)Py(C₁₀N)- β -Me



11) ImImPy- γ -Py(C₁₀G)Py(C₁₀G)Py(C₁₀G)- β -Me
11B) ImImPy- γ (BD)-Py(C₁₀G)Py(C₁₀G)Py(C₁₀G)- β -Me

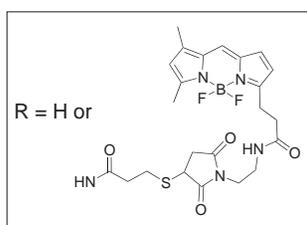


Figure 5.11. Structures of Polyamides 4 and 7-11. “R” represents hydrogen for compounds 4 and 7-11. “R” represents Bodipy for compounds **4B** and **7B-11B**.

5.8 RESULTS

Polyamide Synthesis

Polyamides **4**, **4B**, **7-11**, and **7B-11B** (**Figure 5.11**) were synthesized using solid phase methods.²¹ ImImPy- γ -PyPyPy- β -Dp²⁰ derivatives (**4**, **7**, **8**) and tritylprotected precursors of **4B**, **7B**, and **8B** were synthesized in a stepwise manner from Boc- β -alanine-PAM resin using solid-phase methodology,²¹ followed by cleavage with methylamine and purification by reverse-phase HPLC. Polyamides **9-11** were prepared by perguanidinylation of **4**, **7**, and **8**, respectively, with excess pyrazole-1-carboxamide and sodium carbonate.²⁵ Precursors to polyamide-Bodipy conjugates **4B**, **7B**, and **8B** were synthesized incorporating a protected thiol group at the turn position.²¹ Polyamide-Bodipy conjugates **4B**, **7B**, and **8B** were prepared by TFA deprotection of the corresponding precursors and by treating the resulting free thiol polyamide with Bodipy-FL *N*-(2-aminoethyl)maleimide (**Figure 5.12**). Polyamide-Bodipy conjugates **9B-11B** were prepared by perguanidinylation, followed by TFA deprotection and conjugation with Bodipy-FL *N*-(2-aminoethyl)maleimide.

Quantitative DNase I Footprinting Titrations²⁶

Quantitative DNase I footprint titrations were conducted on the 5'-³²P-end-labeled 197 bp PCR product from pALC1 containing a 5'- TGGTT -3' match and a 5'- TGTAT -3' single-base-pair mismatch site (**Table 5.2**). DNA-binding properties were examined for derivatives of **4** containing three contiguous aminohexyl (**7**) and aminodecyl (**8**) side chains (see Chapter 5A). Compound **7** bound with a decreased affinity for the match site [$K_a = 2.8 \times 10^8 \text{ M}^{-1}$] compared to **4** and with an approximately 10-fold preference over the

mismatch site [$K_a = 4.3 \times 10^7 \text{ M}^{-1}$]. Compound **8** displayed poor binding with similar affinity for match and mismatch sites [$K_a \leq 6 \times 10^7 \text{ M}^{-1}$].

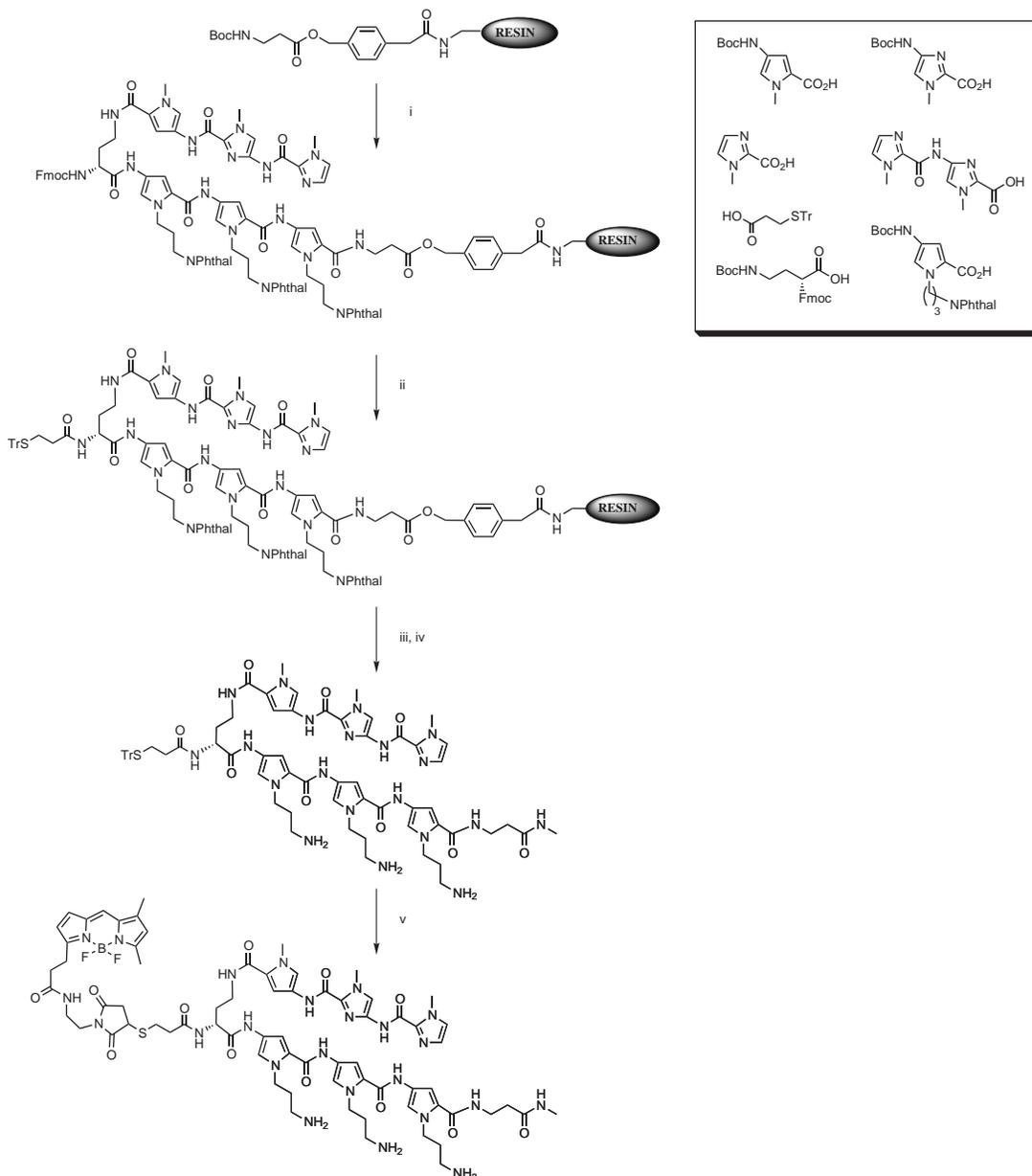


Figure 5.12. Solid-phase Synthesis Scheme for 4B. i) 14-step solid-phase synthesis incorporating H₂Nγ at the turn position (see Chapter 1). ii) Piperidine-DMF (4:1), then β-(Tritylthio)propionic acid, HOBT, DCC, NMP, DIEA. iii) Methylamine, 12 h. iv) TFA, then Et₃SiH. v) 1:1 DMF-Na₂HPO₄ aq. Bodipy-FL N-(2-aminoethyl)maleimide, tris-carboxyethylphosphine•HCl.

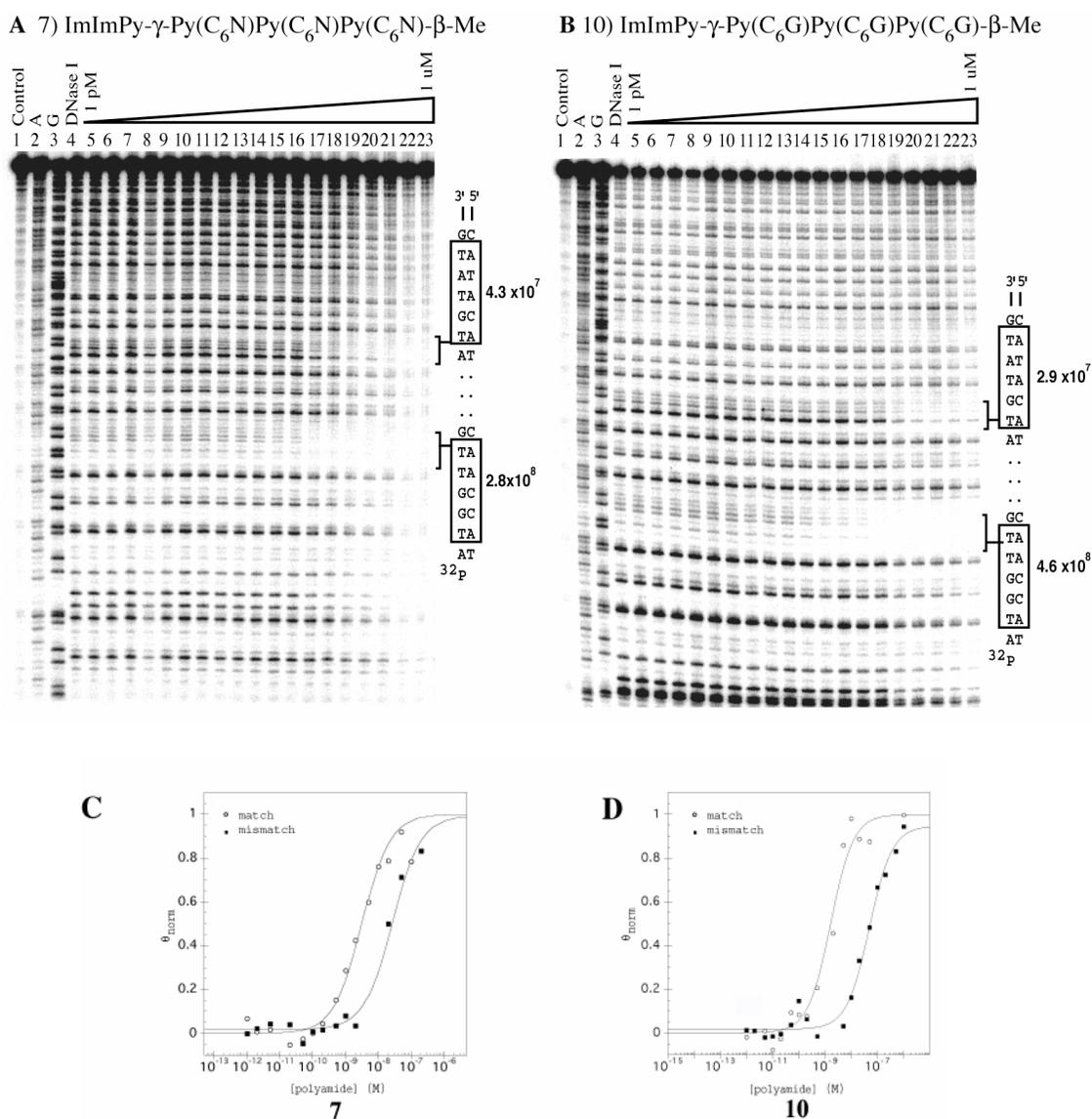


Figure 5.13. Quantitative DNase I Footprint Experiments with A) ImImPy- γ -Py(C₆N)Py(C₆N)Py(C₆N)- β -Me (**7**) and B) ImImPy- γ -Py(C₆G)Py(C₆G)Py(C₆G)- β -Me (**10**) on the 5' ³²P – end labeled 197 bp PCR product from pALC1. All reactions, performed at 22 °C, contained 10 kcpm restriction fragment, 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂. For compounds **7** and **10**: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5-23, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 μ M polyamide, respectively. The positions of the match (5'- TGGTT -3') and single-base-pair mismatch (5'- TGTAT -3') sites are indicated. C) Binding isotherms derived from the DNase I footprinting gels of polyamide **7** on the match 5'-ATGGTT-3' and mismatch 5'-ATGTAT-3' sites. D) Binding isotherms derived from the DNase I footprinting gels of polyamide **10** on the match 5'-ATGGTT-3' and mismatch 5'-ATGTAT-3' sites.

The effects of incorporating guanidine head groups were determined by examining the perguanidinylated derivatives **9**, **10**, and **11**. Polyamide **9** bound with high match-site affinity [$K_a = 1.0 \times 10^{10} \text{ M}^{-1}$] and good specificity over the mismatch site [$K_a = 7.0 \times 10^{10} \text{ M}^{-1}$]. Compound **10** bound with similar affinities to the match site [$K_a = 4.6 \times 10^8 \text{ M}^{-1}$] and mismatch site [$K_a = 2.9 \times 10^7 \text{ M}^{-1}$] compared to **7** (**Figure 5.13**). Polyamide **11** bound with similar affinity for the match site and mismatch site [$K_a \leq 6 \times 10^7 \text{ M}^{-1}$] compared to **8**. Thus, perguanidinylated derivatives displayed DNA-binding properties nearly identical to the aminoalkyl-substituted parents.

Polyamide	Match	Mismatch
	5'- TGGTT -3'	5'- TGTAT -3'
4) ImImPy- γ -Py(C ₃ N)Py(C ₃ N)Py(C ₃ N)- β -Me	$1.6 (\pm 0.4) \times 10^{10}$	$1.0 (\pm 0.6) \times 10^9$
7) ImImPy- γ -Py(C ₆ N)Py(C ₆ N)Py(C ₆ N)- β -Me	$2.8 (\pm 0.7) \times 10^8$	$4.3 (\pm 1.8) \times 10^7$
8) ImImPy- γ -Py(C ₁₀ N)Py(C ₁₀ N)Py(C ₁₀ N)- β -Me	$< 6 \times 10^7$	$< 6 \times 10^7$
9) ImImPy- γ -Py(C ₃ G)Py(C ₃ G)Py(C ₃ G)- β -Me	$1.0 (\pm 0.2) \times 10^{10}$	$7.0 (\pm 2.6) \times 10^8$
10) ImImPy- γ -Py(C ₆ G)Py(C ₆ G)Py(C ₆ G)- β -Me	$4.6 (\pm 1.3) \times 10^8$	$2.9 (\pm 2.7) \times 10^7$
11) ImImPy- γ -Py(C ₁₀ G)Py(C ₁₀ G)Py(C ₁₀ G)- β -Me	$< 6 \times 10^7$	$< 6 \times 10^7$

Table 5.2. Equilibrium Association Constants (M^{-1}). Assays were carried out at 22°C, 10 mM Tris•HCl (pH 7.0), 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂.

Confocal Microscopy³³

Confocal microscopy was utilized to study the distribution of polyamide-Bodipy conjugates **4B** and **7B-11B** in nine cell lines (**Table 5.3**). The polyamide-dye conjugates were mainly observed in the cytoplasm, not in the nucleus. Consistent with previous results, all compounds displayed nuclear uptake in CEM-CCL cultured T-cells.³³ Interestingly, polyamide **4B**, **7B**, and **9B** were also observed in the nuclei of Kc

Drosophila cells (**Figure 5.14A**). Polyamides **8B** and **10B** displayed nuclear uptake in NB4 human leukemia cells (**Figure 5.14B**).

	NB4	CEM-CCL	MEL	SKBR-3	293	LnCAP	PC3	Kc	SF-9
Conjugate	Human Leukemia	Cultured T-cell	Human Erythroid Cancer	Human Breast Cancer	Human Kidney Fibroblast	Human Prostate Cancer	Human Prostate Cancer	<i>Drosophila</i>	Spodoptera Frugiderpa Insect
4B	Cytoplasm	Nucleus			Cytoplasm			Nucleus	Cytoplasm
7B	Cytoplasm	Nucleus	Cytoplasm	Non-cellular	Cytoplasm	Cytoplasm	Cytoplasm	Nucleus	Cytoplasm
8B	Nucleus	Nucleus	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
9B	Nuclear, Inconsistent	Nucleus	Cytoplasm	Non-cellular	Cytoplasm	Cytoplasm	Cytoplasm	Nucleus	Cytoplasm
10B	Nucleus	Nucleus	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm	Cytoplasm
11B	Cytoplasm	Nucleus		Cytoplasm	Cytoplasm			Cytoplasm	Nuclear, Inconsistent

Table 5.3. Cellular Localization of Polyamide-Bodipy Conjugates. Localization of polyamide-Bodipy conjugates in live cells as determined by confocal microscopy. The designation “Nucleus” indicates observation of fluorescence in the interior of the nucleus. The designation “Cytoplasm” indicates cellular, non-nuclear fluorescence. Cells were imaged and prepared as previously described.³³

5.9 DISCUSSION

Impact of Aminoalkyl Chain Length on DNA-Binding Properties

Compounds **7** and **8**, extended linker derivatives of compound **4**, were probed to determine the impact of alkyl linker length on DNA-binding affinity and specificity. Polyamides containing three contiguous aminopropyl (**4**), aminohexyl (**7**), and aminodecyl (**8**) side chains display a dramatic decrease in affinity with increasing alkyl side-chain length. Polyamides **4** and **7** display high affinity and good specificity for their cognate match and mismatch sites, while **8** displays poor binding. Thus aminopropyl and aminohexyl side chains are potential candidates for further studies.

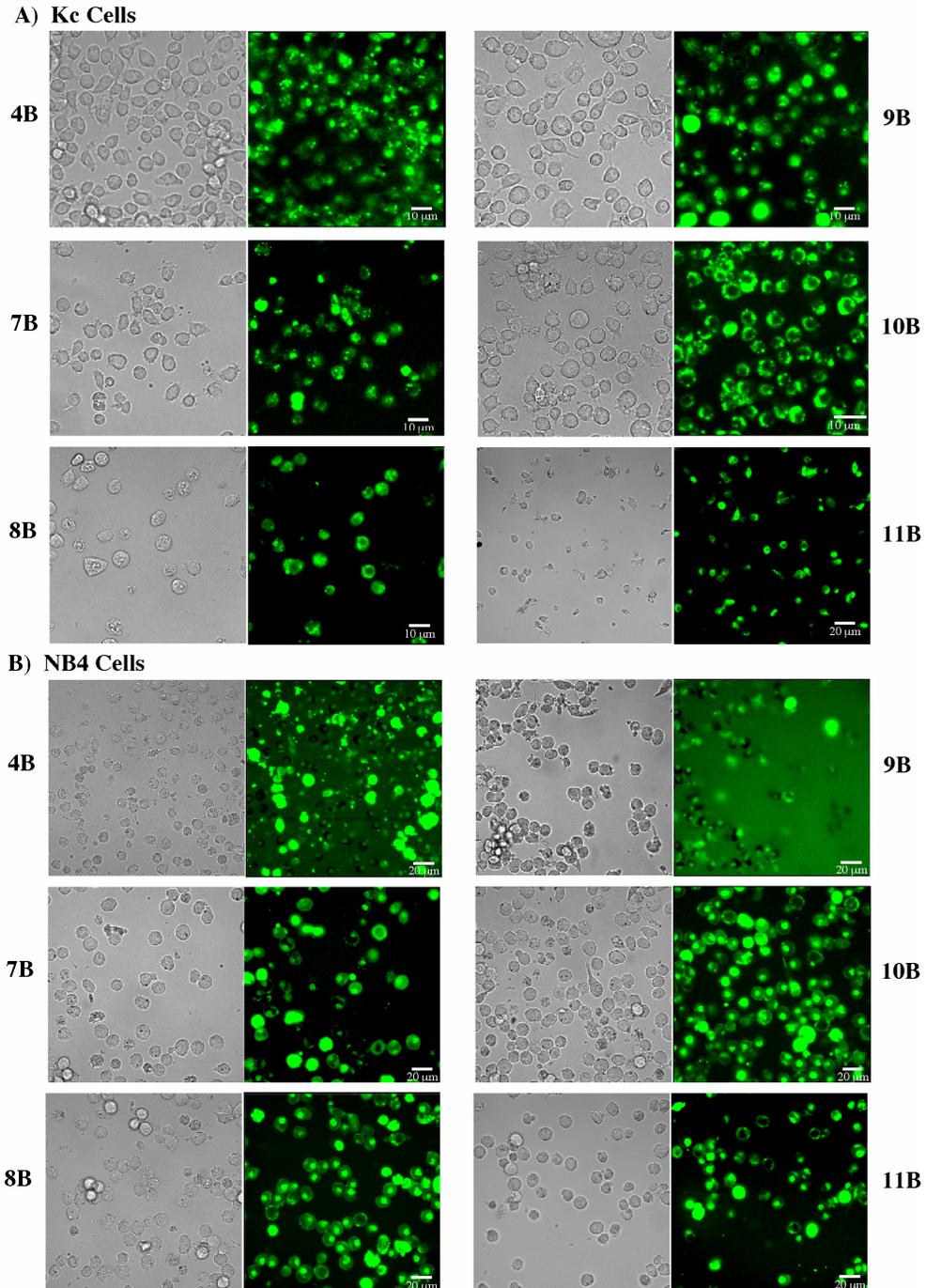


Figure 5.14. Confocal Microscopy. Conjugates **4B** and **7B-11B** in **A)** Kc Cells and **B)** NB4 cells. The transmitted light image is on the left; fluorescent image is on the right. The brightest cells are most likely dead or membrane-compromised, and should be disregarded as they do not reflect uptake of live cells. **A)** Staining throughout cells indicates uptake to the cytoplasm and nucleus (**4B**, **7B**, **8B**), while prominent black circles indicate nuclear exclusion (**9B**, **10B**, **11B**). **B)** Staining throughout the cell indicates uptake to the cytoplasm and nucleus (**8B**, **10B**), while prominent black circles indicate nuclear exclusion (**4B**, **7B**, **9B**, **11B**). A bright spot next to the cell membrane is present in many cells, and does not appear to be associated with the nucleus.

Effects of Cation Alteration on DNA-Binding Properties

Certain classes of guanidinylated small molecules have been shown to display interesting cellular uptake properties.²⁵ Converting primary amines to guanidines alters the size and the charge distribution within the cationic head group, and it was unclear if the introduction of multiple guanidines to a polyamide would alter its DNA-binding properties. To probe these effects *N*-aminoalkyl-substituted polyamides **4**, **7**, and **8** were perguanidinylated to provide compound **9**, **10**, and **11**, respectively. Perguanidinylated derivatives displayed DNA-binding properties nearly identical to the aminoalkyl-substituted parents. Therefore, the conversion of amines to guanidines does not appear to significantly alter polyamide-DNA interactions. This observation provides opportunities for exploring new polyamide motifs with potentially interesting cellular uptake properties.

Nuclear Uptake Properties

It was anticipated that once polyamides gained entry into cells, their high affinity and fast association kinetics would allow localization of these compounds (MW ~1200) to their DNA targets.³³ Recent studies demonstrate that while most polyamides examined were cell permeable, they were only observed in the nucleus of CEM T-cells.³³ The current study, employing polyamide–Bodipy conjugates with cationic side chains, observes nuclear uptake in CEM cells as well as additional cell lines, NB4 human leukemia cells, and *Kc Drosophila* cells. Polyamide conjugates containing aminopropyl (**4B**), aminohexyl (**7B**), and guanidinopropyl (**9B**) side chains exhibited nuclear uptake in

Kc cells. Conjugates containing aminodecyl (**8B**) and guanidinohexyl (**10B**) displayed nuclear uptake in NB4 cells.

Polyamides **4**, **7**, **9**, and **10** display favorable DNA-binding properties with high affinity and good specificity. Compound **8** with aminodecyl side chains displays nuclear uptake in NB4 cells, but demonstrates poor DNA binding. Thus, aminopropyl, aminohexyl, guanidinopropyl, and guanidinohexyl side chains display good DNA-binding properties and promising results in nuclear localization studies.

5.10 CONCLUSIONS

Aminopropyl, aminohexyl, guanidinopropyl, and guanidinohexyl side chains in six-ring hairpins display good DNA-binding properties and show promising results in nuclear uptake studies. This new class of hairpin polyamides containing cationic side chains demonstrates potentially interesting cellular uptake properties in NB4 human leukemia cells and Kc *Drosophila* cells. These results are significant as *Drosophila* serves as an important system for genetic studies and NB4 cells are a model for gene manipulations in human cancer. We are currently in the process of applying these results towards the study of polyamides that target larger binding sites.

Future Research

While the polycationic six-ring polyamides displayed encouraging nuclear uptake properties, they target a relatively short DNA sequence. Eight-ring polyamides target larger binding sites that will provide a more biologically relevant DNA recognition tool. Future research will explore the effects of incorporating multiple cationic side chains on

the DNA-recognition and nuclear-uptake properties of eight-ring polyamides. We will incorporate the three-methylene unit linker, as the six-ring polyamides with aminopropyl and guanidinopropyl sidechains demonstrated good DNA-binding properties and promising nuclear-uptake properties. This series of eight-ring polyamides will contain three and four aminopropyl and guanidinopropyl side chains (**Figure 5.15**). Each polyamide will be synthesized in two analogous forms: unfunctionalized versions for DNA-binding analyses (**12-15**) and fluorescently labeled conjugates for confocal microscopy studies (**12B-15B**).

Future research will include the exploration of additional polyamide motifs as well as *in vivo* functional assays. Cellular uptake studies indicate that molecular weight may serve as a critical factor in the nuclear uptake of polyamides. Research will probe these effects by incorporating intrinsically fluorescent groups into polyamides. These fluorescent polyamides alleviate the need for conjugation to external dyes, thereby decreasing the molecular weight. A newly developed fluorescent benzimidazole-pyrrole dimer can be incorporated in the polyamide backbone as a DNA-recognition element. This group adds very little molecular weight and has been shown to display similar DNA-binding properties to the parent polyamide.

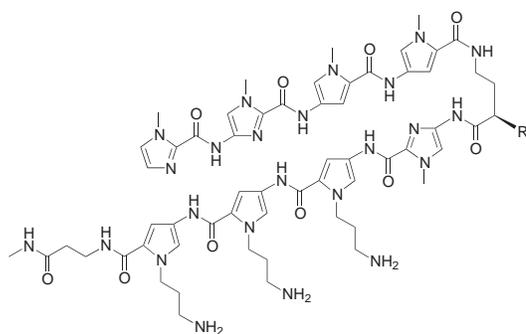
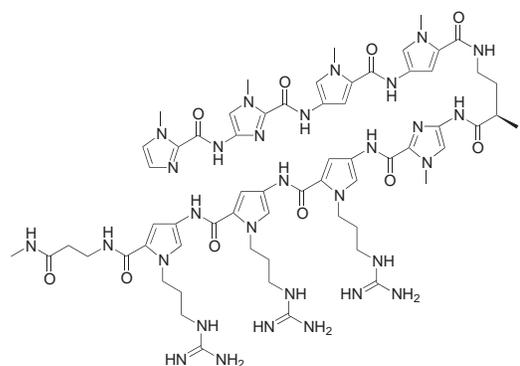
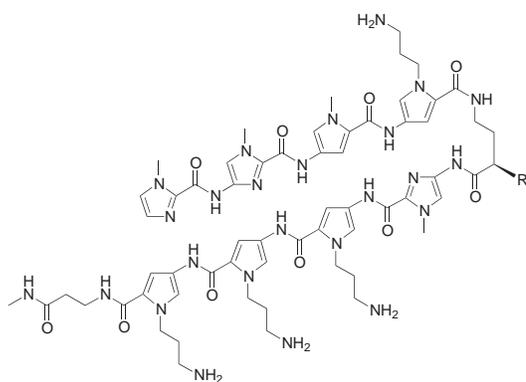
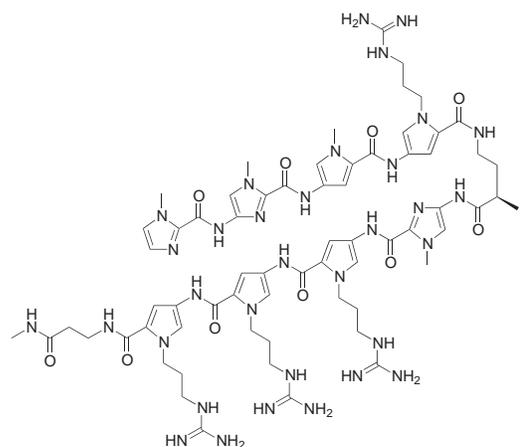
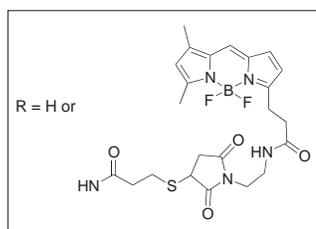
**12** ImImPyPy- γ -ImPy(C₃N)Py(C₃N)Py(C₃N)- β -Me**12B** ImImPyPy- γ (BD)-ImPy(C₃N)Py(C₃N)Py(C₃N)- β -Me**14** ImImPyPy- γ -ImPy(C₃G)Py(C₃G)Py(C₃G)- β -Me**14B** ImImPyPy- γ (BD)-ImPy(C₃G)Py(C₃G)Py(C₃G)- β -Me**13** ImImPyPy(C₃N)- γ -ImPy(C₃N)Py(C₃N)Py(C₃N)- β -Me**13B** ImImPyPy(C₃N)- γ (BD)-ImPy(C₃N)Py(C₃N)Py(C₃N)- β -Me**15** ImImPyPy(C₃G)- γ -ImPy(C₃G)Py(C₃G)Py(C₃G)- β -Me**15B** ImImPyPy(C₃G)- γ (BD)-ImPy(C₃G)Py(C₃G)Py(C₃G)- β -Me

Figure 5.15. Structures of Polyamides 12-15 and 12B-15B. “R” represents hydrogen for compounds 12-15. “R” represents Bodipy for compounds 12B-15B.

Future research will include the exploration of additional polyamide motifs as well as *in vivo* functional assays. Cellular uptake studies indicate that molecular weight may serve as a critical factor in the nuclear uptake of polyamides. Research will probe these effects by incorporating intrinsically fluorescent groups into polyamides. These

fluorescent polyamides alleviate the need for conjugation to external dyes, thereby decreasing the molecular weight. A newly developed fluorescent benzimidazole-pyrrole dimer can be incorporated in the polyamide backbone as a DNA-recognition element. This group adds very little molecular weight and has been shown to display similar DNA-binding properties to the parent polyamide.

In addition, functional assays exploring *in vivo* gene regulation by polyamides could be utilized to monitor the success of nuclear uptake. Such experiments as RNase protection could be utilized to monitor the regulation of genes by polyamides. Fluorescence-activated cell sorting (FACS) could also be performed to monitor the ability of polyamides to regulate gene expression. For example, the activation of green fluorescent protein (GFP) could be studied by FACS to monitor *in vivo* gene activation by polyamides. Reverse transcriptase polymerase chain reaction (RT-PCR) could also be performed. Studies could also be explored on a genomic scale by conducting DNA microarray experiments. These future goals should provide insight toward discovering a solution for polyamide nuclear uptake.

5.11 MATERIALS AND METHODS

Polyamide Synthesis

Reagents and protocols for solid-phase polyamide synthesis were as previously described.²¹ Polyamides **4** and **7-9** were synthesized using semi-automated solid phase protocols as previously described²¹ (see Chapter 5A). Polyamides **9-11** were synthesized by perguanidnylating **4**, **7**, and **8**, respectively (as described in Chapter 5A for polyamide **9**). **ImImPy- γ -Py(C₃G)Py(C₃G)Py(C₃G)- β -Me (9)**. (44% yield). MALDI-TOF-MS

Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5. **ImImPy- γ -Py(C₆G)Py(C₆G)Py(C₆G)- β -Me (10)**. (57% yield). MALDI-TOF-MS Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5. **ImImPy- γ -Py(C₁₀G)Py(C₁₀G)Py(C₁₀G)- β -Me (11)**. (12% yield). MALDI-TOF-MS Calcd for $C_{51}H_{71}N_{25}O_8$ (M+H): 1162.6, found 1162.5. Polyamides **4B** and **7B-11B** were synthesized by Benjamin S. Edelson (data not shown).

Quantitative DNase I Footprinting

Performed as previously described.²⁶ We note explicitly the final buffer concentrations: Tris-HCl buffer (10 mM, pH 7.0), KCl (10 mM), MgCl₂ (10 mM), CaCl₂ (5 mM), and 10 kcpm 5'-radiolabeled DNA. Plasmid pALC1 construction and 5' end-labeling were performed as described (see Chapter 5A).

Confocal Microscopy

Confocal microscopy experiments were conducted on several cell lines as previously described.³³

5.12 REFERENCES

1. Watson, J. D. & Crick, F. H. C. *Nature* **171**, 737 (1953).
2. Dickerson, R. E. et al. *Science* **216**, 475 (1982).
3. Branden, C. & Tooze, J. *Introduction to Protein Structure* (Garland Publishing, Inc., New York and London, 1991).
4. Dervan, P. B. Molecular Recognition of DNA by Small Molecules. *Bioorganic & Medicinal Chemistry* **9**, 2215-2235 (2001).
5. Dervan, P. B., Poulin-Kerstien, A. T., Fechter, E. J. & Edelson, B. S. in *DNA Binders and Related Subjects* 1-31 (2005).
6. Fechter, E. J. & Dervan, P. B. Allosteric inhibition of protein-DNA complexes by polyamide-intercalator conjugates. *Journal of the American Chemical Society* **125**, 8476-85 (2003).
7. Best, T. P., Edelson, B. S., Nickols, N. G. & Dervan, P. B. Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proceedings of the National Academy of Sciences of the USA* **100**, 12063-8 (2003).
8. Edelson, B. S. et al. Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Research* **32**, 2802-18 (2004).
9. Wade, W. S., Mrksich, M. & Dervan, P. B. 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 (1992).
10. White, S., Baird, E. E. & Dervan, P. B. On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chemistry & Biology* **4**, 569-578 (1997).
11. Kielkopf, C. L., Baird, E. E., Dervan, P. D. & Rees, D. C. Structural basis for G center dot C recognition in the DNA minor groove. *Nature Structural Biology* **5**, 104-109 (1998).
12. Kielkopf, C. L. et al. A structural basis for recognition of A center dot T and T center dot A base pairs in the minor groove of B-DNA. *Science* **282**, 111-115 (1998).
13. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E. & Dervan, P. B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **391**, 468-471 (1998).
14. Pelton, J. G. & Wemmer, D. E. Structural Characterization of a 2-1 Distamycin a.D(CGCAAATTGGC) Complex By Two-Dimensional Nmr. *Proceedings of the National Academy of Sciences of the USA* **86**, 5723-5727 (1989).
15. Satz, A. L. & Bruice, T. C. Recognition of Nine Base Pair Sequences in the Minor Groove of DNA at Subpicomolar Concentrations by a Novel Microgonotropen. *Bioorganic & Medicinal Chemistry* **10**, 241-252 (2002).
16. Cuenoud, B. et al. Dual Recognition of Double-Stranded DNA by 2'-Aminoethoxy-Modified Oligonucleotides. *Angewandte Chemie International Edition* **37**, 1288-1291 (1998).

17. Bruice, T. C., Mei, H. Y., He, G. X. & Lopez, V. Rational Design of Substituted Tripyrrole Peptides That Complex With DNA By Both Selective Minor-Groove Binding and Electrostatic Interaction With the Phosphate Backbone. *Proceedings of the National Academy of Sciences of the USA* **89**, 1700-1704 (1992).
18. Blasko, A., Browne, K. A. & Bruice, T. C. Microgonatropens and their interactions with DNA. 5. Structural characterization of the 1:1 complex of d(CGCAAATTTGCG)₂ and the tren-microgonatropen-b by 2D NMR spectroscopy and restrained molecular modeling. *Journal of the American Chemical Society* **116**, 3726-3737 (1994).
19. Bremer, R. E., Wurtz, N. R., Szewczyk, J. W. & Dervan, P. B. Inhibition of major groove DNA binding bZIP proteins by positive patch polyamides. *Bioorganic & Medicinal Chemistry* **9**, 2093-2103 (2001).
20. Parks, M. E., Baird, E. E. & Dervan, P. B. Recognition of 5'-(A,T)GG(A,T)(2)-3' sequences in the minor groove of DNA by hairpin polyamides. *Journal of the American Chemical Society* **118**, 6153-6159 (1996).
21. Baird, E. E. & Dervan, P. B. Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *Journal of the American Chemical Society* **118**, 6141-6146 (1996).
22. Wurtz, N. R., Ph. D. Thesis (2002).
23. Rucker, V. C., Foister, S., Melander, C. & Dervan, P. B., Manuscript in preparation.
24. Nishiwaki, E., Tanaka, S., Lee, H. & Shibuya, M. *Heterocycles* **27**, 1945 (1988).
25. Wender, P. A. et al. The design, synthesis, and evaluation of molecules that enable or enhance cellular uptake: Peptoid molecular transporters. *Proceedings of the National Academy of Sciences of the USA* 13003-13008 (2000).
26. Trauger, J. W. & Dervan, P. B. Footprinting Methods for Analysis of Pyrrole-Imidazole Polyamide/DNA Complexes. *Methods in Enzymology* **340**, 450-466 (2001).
27. Gottesfeld, J. M., Turner, J. M. & Dervan, P. B. Chemical Approaches to Control Gene Expression. *Gene Expression* **9**, 77-91 (2000).
28. Mapp, A. K., Ansari, A. Z., Ptashne, M. & Dervan, P. B. Activation of gene expression by small molecule transcription factors. *Proceedings of the National Academy of Sciences of the USA* **97**, 3930-5 (2000).
29. Ansari, A. Z., Mapp, A. K., Nguyen, D. H., Dervan, P. B. & Ptashne, M. Towards a minimal motif for artificial transcriptional activators. *Chemical Biology* **8**, 583-92 (2001).
30. Dickinson, L. A. et al. Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. *Proceedings of the National Academy of Sciences of the USA* **95**, 12890-12895 (1998).
31. Janssen, S., Durussel, T. & Laemmli, U. K. Chromatin opening of DNA satellites by targeted sequence-specific drugs. *Molecular Cell* **6**, 999-1011 (2000).
32. Janssen, S., Cuvier, O., Muller, M. & Laemmli, U. K. Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to *Drosophila melanogaster*. *Molecular Cell* **6**, 1013-24 (2000).

33. Belitsky, J. M., Leslie, S. J., Arora, P. S., Beerman, T. A. & Dervan, P. B. Cellular Uptake of N-Methylpyrrole/N-Methylimidazole Polyamide-Dye Conjugates. *Bioorganic & Medicinal Chemistry*, Manuscript Submitted. (2002).
34. Alberts, B. et al. *Molecular Biology of the Cell* (ed. Robertson, M.) (Garland Publishing, Inc., New York, 1994).
35. Crowley, K. S. et al. Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorganic and Medicinal Chemistry Letters* **13**, 1565-70 (2003).