Recognition of 5′-(A,T)GG(A,T)$_2$-3′ Sequences in the Minor Groove of DNA by Hairpin Polyamides

Michelle E. Parks, Eldon E. Baird, and Peter B. Dervan*

Contribution from the Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received March 6, 1996[⊗]

Abstract: A series of four hairpin pyrrole−imidazole polyamides, ImImPy-γ-PyPyPy-β-Dp, PyPyPy-γ-ImImPy-β-Dp, AcImImPy-γ-PyPyPy-β-Dp, and AcPyPyPy-γ-ImImPy-β-Dp (Im = *N*-methylimidazole-2-carboxamide, Py = *N*-methylpyrrole-2-carboxamide, Dp = *N*,*N*-dimethylaminopropylamide, γ = γ-aminobutyric acid, β = β-alanine, and Ac = acetyl), designed for recognition of 5′-(A,T)GG(A,T)$_2$-3′ sequences in the minor groove of DNA were synthesized using solid phase methodology and analyzed with respect to DNA binding affinity and sequence specificity. Quantitative DNase I footprint titration experiments reveal that the optimal polyamide ImImPy-γ-PyPyPy-β-Dp binds a designated 5′-TGGTT-3′ match site with an equilibrium association constant of $K_a = 1.0 \times 10^8$ M^{-1} and the single base pair mismatch sites, 5′-TGTTA-3′ and 5′-GGGTA-3′, with 50-fold and 100-fold-lower affinity, respectively (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$, pH 7.0 and 22 °C). Polyamides of sequence composition AcImImPy-γ-PyPyPy-β-Dp and AcPyPyPy-γ-ImImPy-β-Dp, which differ only by the position of the γ-linker, bind with similar affinities and specificities. Recognition of sequences containing contiguous G·C base pairs expands the sequence repertoire available for targeting DNA with pyrrole−imidazole polyamides.

Introduction

Pyrrole−imidazole polyamides offer a general method for the design of non-natural molecules for sequence-specific recognition in the minor groove of DNA.[1−3] Within the 2:1 polyamide−DNA model, an imidazole (Im) on one ligand opposite a pyrrolecarboxamide (Py) on the second ligand recognizes a G·C base pair, while a pyrrolecarboxamide/imidazole combination targets a C·G base pair.[1,3] A pyrrolecarboxamide/pyrrolecarboxamide pair is partially degenerate for A·T or T·A base pairs.[1−3] On the basis of this model, the recognition of the sequences 5′-(A,T)G(A,T)C(A,T)-3′,[1] 5′-(A,T)G(A,T)$_3$-3′,[3] (A,T)$_2$G(A,T)$_2$-3′,[4] and 5′-(A,T)GCGC(A,T)-3′ [5] has been achieved. However, sequences containing *contiguous* G·C base pairs are notably absent from this list.

Formation of a hairpin polyamide by covalently linking a polyamide heterodimer with a γ-aminobutyric acid (γ) residue provides an approximate 300-fold enhancement in affinity over the unlinked polyamides, ImPyPy- and PyPyPy-.[6] Moreover, the specificity of the hairpin is greatly improved. The initial placement of the γ-amino acid turn was chosen for synthetic ease and was not varied. With the development of solid phase methodology for polyamide synthesis, we now assess the effect of varying the position of the γ-turn monomer.[7]

In order to explore the recognition of 5′-(A,T)GG(A,T)$_2$-3′ sequences, a series of four head-to-tail linked hairpin polyamides containing *neighboring* imidazole rings, ImImPy-γ-PyPyPy-β-Dp (**1**), PyPyPy-γ-ImImPy-β-Dp (**2**), AcImImPy-γ-PyPyPy-β-Dp (**3**), and AcPyPyPy-γ-ImImPy-β-Dp (**4**), were prepared using solid phase methods (Figures 1 and 2).[7] The polyamides are all synthesized with Boc-β-alanine-Pam-resin, previously shown as optimal for polyamides.[8] Each imidazole is expected to form a specific hydrogen bond with a guanine amino group allowing the recognition of contiguous G·C base pairs (Figure 1). In addition, the linker turn position is varied within the nonacetylated and acetylated pairs of polyamides to determine the effect on the sequence specificity and binding affinity. We report here the binding specificity and affinity of the polyamides as determined by the complementary techniques, MPE·Fe^{II} footprinting[9] and quantitative DNase I footprinting.[10] MPE·Fe^{II} footprinting verifies that sequence-specific recognition of the expected 5′-TGGTT-3′ target site has been achieved. In addition, DNase I quantitative footprint titration experiments reveal that the position of the γ-linker does not dramatically affect either affinity or specificity of polyamides, especially the pair containing acetylated N-termini.

Results

Synthesis of Polyamides. The polyamides ImImPy-γ-PyPyPy-β-Dp (**1**), PyPyPy-γ-ImImPy-β-Dp (**2**), AcImImPy-γ-PyPyPy-β-Dp (**3**), and AcPyPyPy-γ-ImImPy-β-Dp (**4**) were

[⊗] Abstract published in *Advance ACS Abstracts,* June 15, 1996.

(1) (a) Wade, W. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 1574−1575. (b) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783. (c) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci., U.S.A.* **1992**, *89*, 7586. (d) Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11385.

(2) (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723. (b) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393. (c) Chen, X.; Ramakrishnan, B.; Rao, S. T.; Sundaralingham, M. *Nature Struct. Biol.* **1994**, *1*, 169.

(3) (a) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572. (b) Geierstanger, B. H.; Jacobsen, J.-P.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Biochemistry* **1994**, *33*, 3055.

(4) Geierstanger, B. H.; Dwyer, T. J.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 4474.

(5) (a) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Science* **1994**, *266*, 646−650. (b) Mrksich, M.; Dervan, P. B.; *J. Am. Chem. Soc.* **1995**, *117*, 3325.

(6) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.

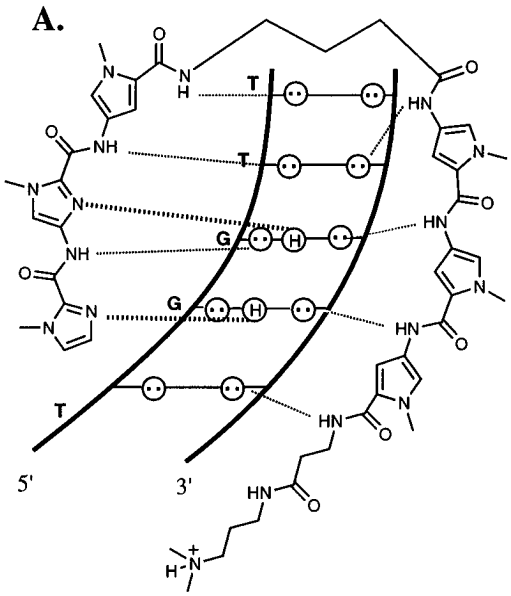
(7) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141−6146.

(8) Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147−6152.

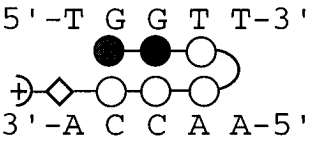
(9) (a) Van Dyke, M. W.; Dervan, P. B. *Biochemistry* **1983**, *22*, 2373. (b) Van Dyke, M. W.; Dervan, P. B. *Nucleic Acids Res.* **1983**, *11*, 5555.

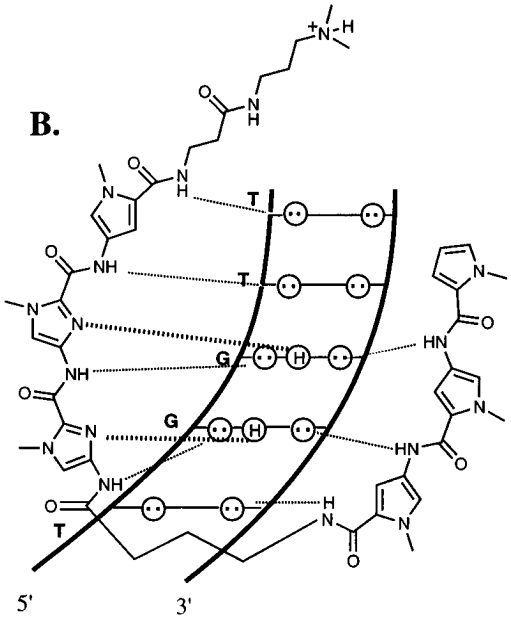
(10) (a) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132. (b) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8462. (c) Senear, D. F.; Brenowitz, M.; Shea, M. A.; Ackers, G. K. *Biochemistry* **1986**, *25*, 7344.

S0002-7863(96)00728-7 CCC: $12.00



ImImPy-γ-PyPyPy-β-Dp • TGGTT





PyPyPy-γ-ImImPy-β-Dp • TGGTT

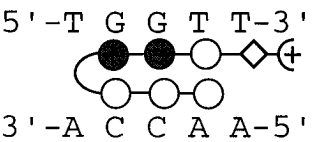
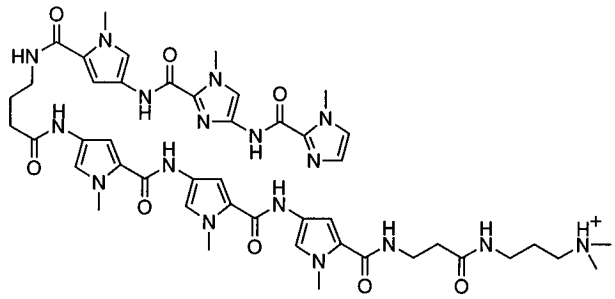
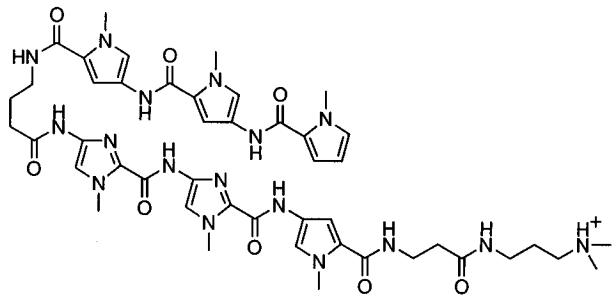


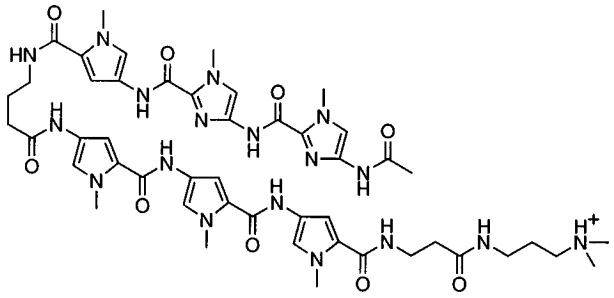
Figure 1. Binding model for the complexes formed between polyamides ImImPy-γ-PyPyPy-β-Dp (**1**) (A) and PyPyPy-γ-ImImPy-β-Dp (**2**) (B) and a 5′-TGGTT-3′ sequence. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. Ball and stick models are also shown. Shaded and nonshaded circles denote imidazole and pyrrole carboxamides, respectively. Nonshaded diamonds represent a β-alanine residue.



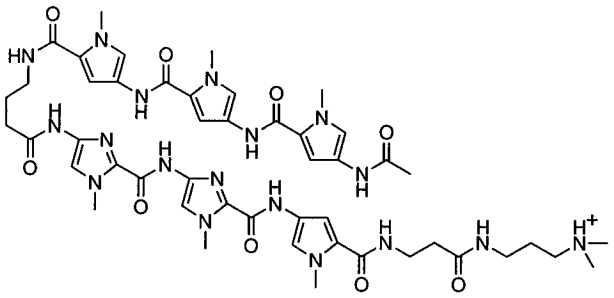
ImImPy-γ-PyPyPy-β-Dp (1)



PyPyPy-γ-ImImPy-β-Dp (2)



AcImImPy-γ-PyPyPy-β-Dp (3)



AcPyPyPy-γ-ImImPy-β-Dp (4)

Figure 2. Series of polyamides synthesized using solid phase methodology.[7]

prepared by solid phase methodology (Figure 2). Four unique pyrrole and imidazole building blocks were combined in a stepwise manner on a solid support using Boc-chemistry protocols (Figure 3). For example, polyamide **2,** PyPyPy-γ-ImImPy-β-Dp, was prepared in 14 steps on the resin, and then cleaved with a single-step aminolysis reaction (Figure 3). All polyamides were found to be soluble up to at least 1 mM concentration in aqueous solution.

Footprinting. MPE·Fe^{II} footprinting on a 3′- or 5′-^{32}P-end-labeled 266 base pair *Eco*RI/*Pvu*II restriction fragment from plasmid pMEPGG (25 mM Tris-acetate, 100 μM bp calf thymus DNA, 10 mM NaCl) reveals that the synthetic polyamides **1**−**4**, at 10 μM concentration, bind the designated target site 5′-

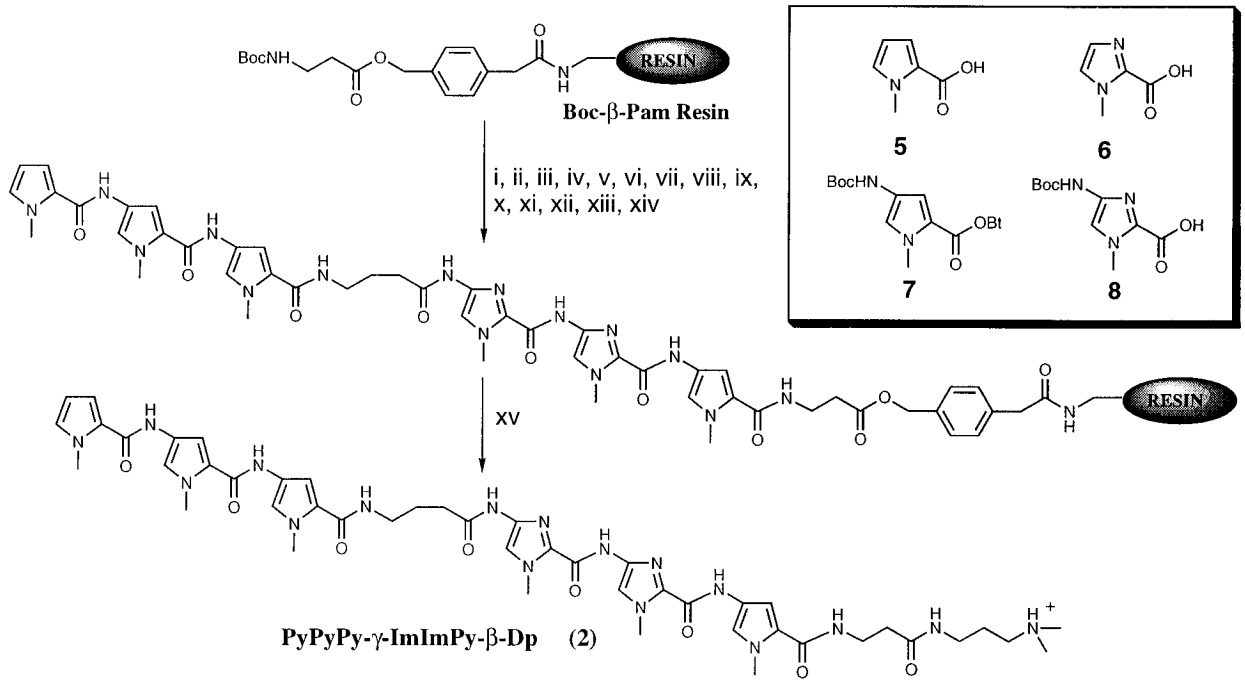


Figure 3. Solid phase synthetic scheme for PyPyPy-γ-ImImPy-β-Dp starting from commercially available Boc-β-Pam-resin: (i) 80% TFA/DCM, 0.4 M PhSH; (ii) Boc-Py-OBt, DIEA, DMF; (iii) 80% TFA/DCM, 0.4 M PhSH; (iv) Boc-Im-OBt (DCC/HOBt), DIEA, DMF; (v) 80% TFA/DCM, 0.4 M PhSH; (vi) Boc-Im-OBt (DCC/HOBt), DIEA, DMF; (vii) 80% TFA/DCM, 0.4 M PhSH; (viii) Boc-γ-aminobutyric acid (HBTU, DIEA); (ix) 80% TFA/DCM, 0.4 M PhSH; (x) Boc-Py-OBt, DIEA, DMF; (xi) 80% TFA/DCM, 0.4 M PhSH; (xii) Boc-Py-OBt, DIEA, DMF; (xiii) 80% TFA/DCM, 0.4 M PhSH; (xiv); pyrrole-2-carboxylic acid (HBTU/DIEA); (xv) (*N*,*N*- dimethylamino)propylamine, 55 °C. (Inset) Pyrrole and imidazole monomers for synthesis of all compounds described here: pyrrole-2-carboxylic acid **5**, imidazole-2-carboxylic acid **6**,[1b] Boc−pyrrole−OBt ester **7**,[7] and Boc−imidazole acid **8**.[7]

TGGTT-3′ (Figures 4 and 5). In addition, several single base pair mismatch sites are bound with lower affinity. Quantitative DNase I footprint titration experiments (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$, pH 7.0 and 22 °C) were performed to determine the equilibrium association constants of the four polyamides **1**−**4** for a designated match site, 5′-TGGTT-3′, as well as for two single base pair mismatch sites, 5′-TGTTA-3′ and 5′-GGGTA-3′ (Table 1).[10] The polyamide ImImPy-γ-PyPyPy-β-Dp binds the target site 5′-TGGTT-3′ with the highest affinity (association constant $K_a = 1.0 \times 10^8$ M^{-1}) (Figures 6 and 7). The remaining polyamides have lower but approximately equal association constants of $K_a = \sim 2 \times 10^7$ M^{-1} for the target site. The nonacetylated polyamides in this series are >50-fold specific for the 5′-TGGTT-3′ match site over either of the single base pair mismatch sites analyzed. The acetylated pair of polyamides exhibit lower sequence specificity for the analyzed sites.

Discussion

Each polyamide within this series specifically binds the five base pair designated target sequence 5′-TGGTT-3′, as shown by MPE·Fe^{II} footprinting experiments, providing the first example of contiguous G·C recognition in the polyamide−DNA motif. Interestingly, the polyamides prefer different mismatch sequences, indicating that the position of the turn alters sequence selectivity, although only for the mismatches.

Quantitative DNase I footprint titration experiments reveal that ImImPy-γ-PyPyPy-β-Dp (**1**) is optimal within this series of four polyamides. This hairpin binds a 5′-TGGTT-3′ match site with an equilibrium association constant of $K_a = 1 \times 10^8$ M^{-1}, while the corresponding hairpin PyPyPy-γ-ImImPy-β-Dp (**2**), which differs only in the position of the γ turn, shows lower affinity ($K_a = \sim 2 \times 10^7$ M^{-1}) for the 5′-TGGTT-3′ site. Both unacetylated polyamides demonstrate good specificity (>50-fold) for the target match site over the single base pair mismatch sites. The acetylated polyamides are similar in affinity to PyPyPy-γ-ImImPy-β-Dp (**2**), but exhibit lower specificity. AcImImPy-γ-PyPyPy-β-Dp (**3**) and AcPyPyPy-γ-ImImPy-β-Dp (**4**) are virtually indistinguishable from each other on the basis of affinity and specificity for the analyzed target sequences, indicating little preference for turn position.

This series of contiguous imidazole-containing polyamides is remarkably similar in affinity and specificity to the single imidazole-containing hairpin polyamide, ImPyPy-γ-PyPyPy-β-Dp, indicating little or no energetic penalty in this system for *adjacent* imidazoles.[8] Importantly, the position of the hairpin turn does not significantly affect the recognition of the target 5′-TGGTT-3′ match site, although single base pair mismatch relative affinities are altered.

Implications for the Design of Minor Groove Binding Molecules. The 2:1 motif has been used to specifically target several sequences: 5′-TGTCA-3′,[1] 5′-TGTTA-3′,[3] 5′-AAGTT-3′,[4] and 5′-TGCGCA-3′.[5] The results reported herein add sequences containing two contiguous G·C base pairs to the list, expanding the sequence repertoire for DNA recognition by polyamides. Furthermore, turn position showed minimal effects on the specificity and affinity of the polyamides, indicating a new degree of flexibility within the 2:1 motif. The expansion of the polyamide sequence repertoire through contiguous G·C recognition coupled with solid phase synthetic advances allowing the rapid assembly and characterization of polyamides brings the goal of sequence-specific recognition of any DNA sequence by designed molecules closer to fruition.

Experimental Section

Materials. Boc-glycine-(4-carbonylaminomethyl)-benzyl-ester-co-poly(styrene−divinylbenzene) resin (Boc-G-Pam-resin) (0.2 mmol/g) 0.2 mmol/g Boc-β-alanine-(4-carbonylaminomethyl)-benzyl-ester-co-

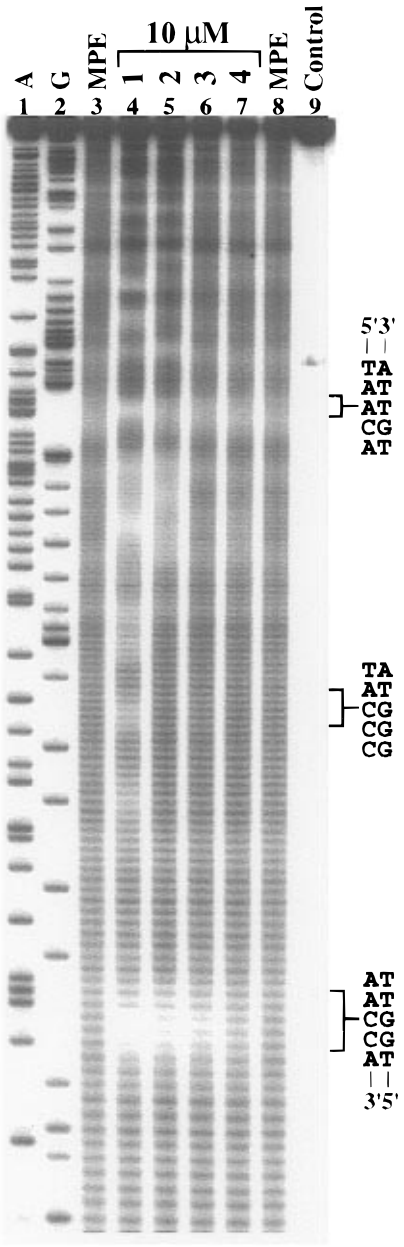


Figure 4. MPE·Fe^{II} footprinting experiment on a 3′-^{32}P-labeled 266 bp *Eco*RI/*Pvu*II restriction fragment from plasmid pMEPGG. The 5′-TGGTT-3′, 5′-GGGTA-3′, and 5′-TGTTA-3′ sites are shown on the right side of the autoradiogram. All reactions contain 10 kcpm restriction fragment, 25 mM Tris-acetate, 10 mM NaCl, 100 μM calf thymus DNA (bp), and 5 mM DTT. Lane 1: A reaction. Lane 2: G reaction. Lanes 3 and 8: MPE·Fe^{II} standard. Lane 4: 10 μM ImImPy-γ-PyPyPy-β-Dp (**1**). Lane 5, 10 μM PyPyPy-γ-ImImPy-β-Dp (**2**). Lane 6: 10 μM AcImImPy-γ-PyPyPy-β-Dp (**3**). Lane 7: 10 μM AcPyPyPy-γ-ImImPy-β-Dp (**4**). Lane 9; intact DNA.

poly(styrene−divinylbenzene) resin (Boc-β-Pam-resin), dicyclohexylcarbodiimide (DCC), hydroxybenzotriazole (HOBt), 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), Boc-glycine, and Boc-β-alanine were purchased from Peptides International. *N*,*N*-Diisopropylethylamine (DIEA), *N*,*N*-dimethylformamide (DMF), *N*-methylpyrrolidone (NMP), DMSO/NMP, and acetic anhydride (Ac_2O) were purchased from Applied Biosystems. Boc-γ-aminobutyric acid was from NOVA Biochem, dichloromethane (DCM) and triethylamine (TEA) were reagent grade from EM, thiophenol (PhSH) and (dimethylamino)propylamine were from Aldrich, trifluoroacetic acid (TFA) was from Halocarbon. All reagents were used without further purification.

^{1}H NMR spectra were recorded in DMSO-d_6 on a GE 300 instrument operating at 300 MHz. Chemical shifts are reported in parts per million relative to the solvent residual signal. UV spectra were measured on a Hewlett-Packard Model 8452A diode array spectrophotometer. Matrix-assisted, laser desorption/ionization time of flight mass spectrometry (MALDI-TOF-MS) was carried out at the Protein and Peptide Microanalytical Facility at the California Institute of Technology. HPLC analysis was performed either on a HP 1090M analytical HPLC or on a Beckman Gold system using a Rainin C_{18}, Microsorb MV, 5 μm, 300 × 4.6 mm reversed-phase column in 0.1% (w/v) TFA with acetonitrile as eluent and a flow rate of 1.0 mL/min, gradient elution 1.25% acetonitrile/min. Preparatory HPLC was carried out on a Beckman HPLC using a Waters DeltaPak 25 × 100 mm, 100 μm C_{18} column equipped with a guard, 0.1% (w/v) TFA, 0.25% acetonitrile/min. 18MΩ water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered. Reagent-grade chemicals were used unless otherwise stated.

Activation of Boc-γ-aminobutyric Acid, Imidazole-2-carboxylic acid, and Pyrrole-2-carboxylic acid. The appropriate amino acid or acid (2 mmol) was dissolved in 2 mL of DMF. HBTU (720 mg, 1.9 mmol) was added followed by DIEA (1 mL) and the solution lightly shaken for at least 5 min.

Activation of Boc−Imidazole Acid. Boc−imidazole acid (257 mg, 1 mmol) and HOBt (135 mg, 1 mmol) were dissolved in 2 mL of DMF, DCC (202 mg, 1 mmol) was then added, and the solution was allowed to stand for at least 5 min.

Typical Manual Synthesis Protocol. PyPyPy-γ-ImImPy-β-Dp. Boc-β-Pam-resin (1.25 g, 0.25 mmol of amine) was shaken in DMF for 30 min and drained. The N-Boc group was removed by washing with DCM for 2 × 30 s, followed by a 1 min shake in 80% TFA/DCM/0.5 M PhSH, draining the reaction vessel, a brief 80% TFA/DCM/0.5 M PhSH wash, and 20 min shaking in 80% TFA/DCM/0.5 M PhSH solution. The resin was washed for 1 min with DCM and 30 s with DMF. A resin sample (8−10 mg) was taken for analysis. The resin was drained completely, Boc-pyrrole-OBt monomer (357 mg, 1 mmol) dissolved in 2 mL of DMF was added, followed by DIEA (1 mL), and the resin was shaken vigorously to make a slurry. The coupling was allowed to proceed for 45 min. A resin sample (8−10 mg) was taken after 40 min to check reaction progress. The reaction vessel was washed with DMF for 30 s and dichloromethane for 1 min to complete a single reaction cycle. Six additional cycles were performed, adding successively Boc-Im-OH (DCC/HOBt), Boc-Im-OH (DCC/HOBt), Boc-γ-aminobutyric acid (HBTU/DIEA), Boc-Py-OBt, Boc-Py-OBt, and pyrrole-2-carboxylic acid (HBTU/DIEA). The resin was washed with DMF, DCM, MeOH, and ethyl ether and then dried *in vacuo*. PyPyPy-γ-ImImPy-β-Pam-resin (180 mg, 29 μmol)[12] was weighed into a glass scintillation vial, 1.5 mL of (*N*,*N*-dimethylamino)-propylamine added, and the mixture heated at 55 °C for 18 h. The resin was removed by filtration through a disposable polypropylene filter and washed with 5 mL of water, the amine solution and the water washes were combined, the solution was loaded on a C_{18} preparatory HPLC column. The polyamide was then eluted in 100 min as a well-defined peak with a gradient of 0.25% acetonitrile/min. The polyamide was collected in four separate 8 mL fractions, and the purity of the individual fractions was verified by HPLC and ^{1}H NMR, to provide purified PyPyPy-γ-ImImPy-β-Dp (**2**) (11.2 mg, 39% recovery): UV λ_{max} 246 (31 100), 312 (51 200); ^{1}H NMR (DMSO-d_6) δ 10.30 (s, 1 H), 10.26 (s, 1 H), 9.88 (s, 1 H), 9.80 (s, 1 H), 9.30 (s, 1 H), 9.2 (br s, 1 H), 8.01 (m, 3 H), 7.82 (br s 1 H), 7.54 (s, 1 H), 7.52 (s, 1 H), 7.20 (d, 1 H, J = 1.3 Hz), 7.18 (d, 1 H, J = 1.2 Hz), 7.15 (d, 1 H, J = 1.3 Hz), 7.01 (d, 1 H, J = 1.4 Hz), 6.96 (d, 1 H, J = 1.4 Hz), 6.92 (d, 1 H, J = 1.8 Hz), 6.89 (m, 2 H), 6.03 (t, 1 H, J = 2.4 Hz), 3.97 (s, 3 H), 3.96 (s, 3 H), 3.85 (s, 3 H), 3.82 (s, 3 H), 3.78 (m, 6 H), 3.37 (m, 2 H), 3.20 (q, 2 H, J = 5.7 Hz), 3.08 (q, 2 H J = 6.6 Hz), 2.94 (q, 2 H J = 5.3 Hz), 2.71 (d, 6 H J = 5.8 Hz), 2.32 (m, 4 H), 1.83 (m, 4 H); MALDI-TOF-MS 978.7 (979.1 calcd for M + H).

ImImPy-γ-PyPyPy-β-Dp (1). Polyamide was prepared by machine-assisted solid phase synthesis protocols,[7] and 900 mg of resin was cleaved and purified to provide **1** as a white powder (69 mg, 48% recovery): UV λ_{max} 246 (43 300), 308 (54 200); ^{1}H NMR (DMSO-d_6)

(11) Kent, S. B. H. *Annu. Rev. Biochem.* **1988**, *57*, 957.

(12) Resin substitution can be calculated as L_{new}(mmol/g) = L_{old}/(1 + L_{old}(W_{new} − W_{old}) × 10^{-3}), where L is the loading (mmol of amine/g of resin), and W is the weight (gmol^{-1}) of the growing polyamide attached to the resin. See: Barlos, K.; Chatzi, O.; Gatos, D.; Stravropoulos, G. *Int. J. Pept. Protein Res.* **1991**, *37*, 513.

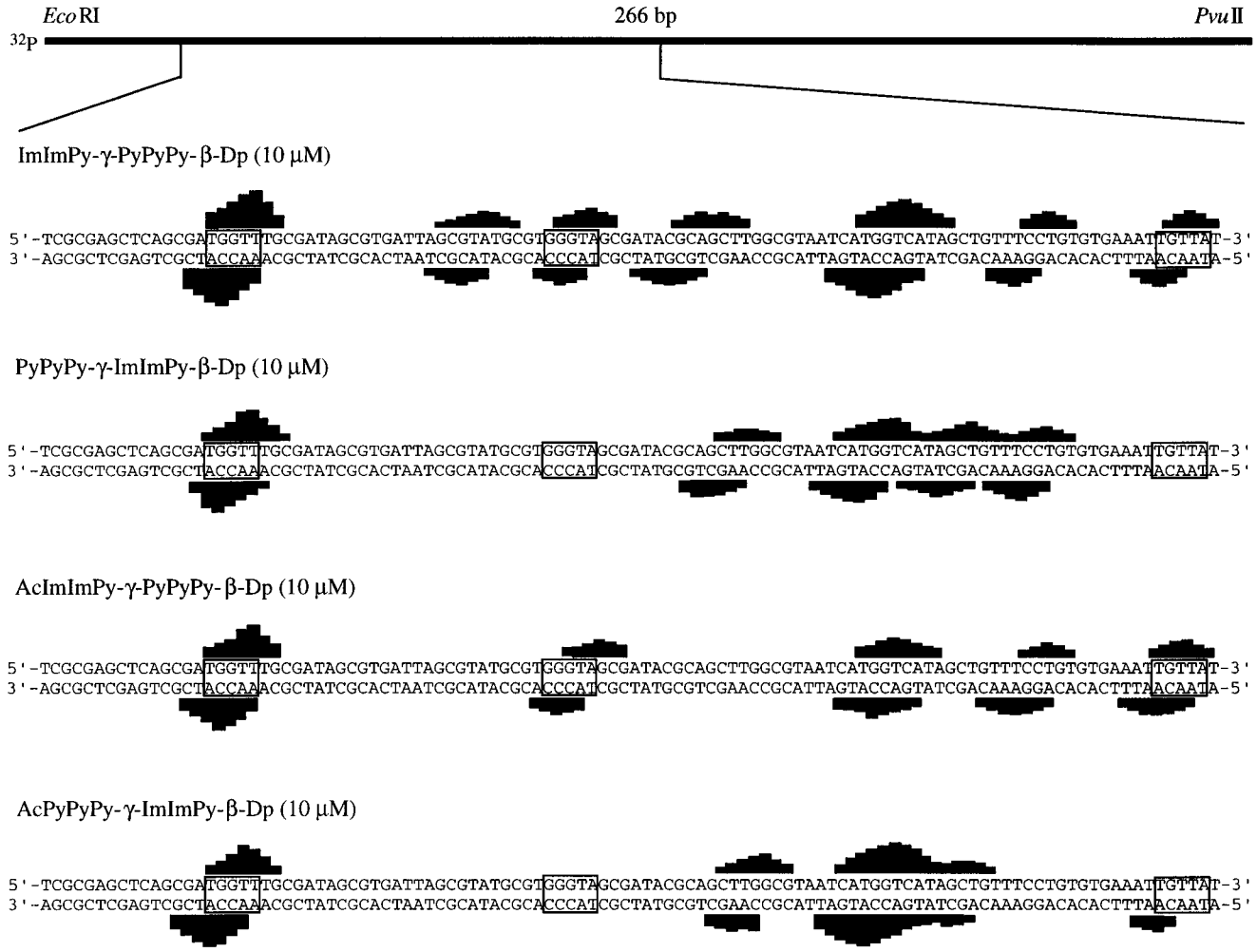


Figure 5. Histograms of cleavage protection (footprinting) data. (Top) Illustration of the 266 bp restriction fragment with the position of the sequence indicated. MPE·Fe^{II} protection patterns for polyamides at 10 μM concentration. Bar heights are proportional to the relative protection from cleavage at each band. Boxes represent equilibrium binding sites determined by the published model.[9] Only sites that were quantitated by DNase I footprint titrations are boxed.

Table 1. Equilibrium Association Constants (M^{-1})[a,b]

polyamide	match site 5′-aTGGTTt-3′	single mismatch sites 5′-tTGTTAt-3′	single mismatch sites 5′-tGGGTAg-3′
ImImPy-γ-PyPyPy-β-Dp	1.0×10^8 (0.1)	1.7×10^6 (0.6)	$\leq 1 \times 10^6$
PyPyPy-γ-ImImPy-β-Dp	1.6×10^7 (0.2)	$< 1 \times 10^5$	$< 1 \times 10^5$
AcImImPy-γ-PyPyPy-β-Dp	1.3×10^7 (0.7)	1.6×10^6 (1.1)	1.3×10^6 (0.8)
AcPyPyPy-γ-ImImPy-β-Dp	2.0×10^7 (0.3)	1.3×10^6 (0.5)	$\leq 1 \times 10^6$

[a] Values reported are the mean values measured from at least three footprint titration experiments, with the standard deviation for each data set indicated in parentheses. [b] The assays were performed at 22 °C at pH 7.0 in the presence of 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

δ 10.31 (s, 1 H), 9.91 (s, 1 H), 9.90 (s, 1 H), 9.85 (s, 1 H), 9.75 (s, 1 H), 9.34 (br s, 1 H), 8.03 (m, 3 H), 7.56 (s, 1 H), 7.46 (s, 1 H), 7.21 (m, 2 H), 7.15 (m, 2 H), 7.07 (d, 1 H J = 1.2 Hz), 7.03 (d, 1 H, J = 1.3 Hz), 6.98 (d, 1 H, J = 1.2 Hz), 6.87 (m, 2 H), 4.02 (m, 6 H), 3.96 (m, 6 H), 3.87 (m, 6 H), 3.75 (q, 2 H, J = 4.9 Hz), 3.36 (q, 2 H, J = 4.0 Hz), 3.20 (q, 2 H, J = 4.7 Hz), 3.01 (q, 2 H, J = 5.1 Hz), 2.71 (d, 6 H, J = 4.8 Hz), 2.42 (m, 4 H), 1.80 (m, 4 H) MALDI-TOF-MS 978.8 (979.1 calcd for M + H).

AcImImPy-γ-PyPyPy-β-Dp (3). Polyamide was prepared by manual solid phase protocols and isolated as a white powder (8 mg, 28% recovery): UV λ_{max} 246 (43 400), 312 (50 200); ^{1}H NMR (DMSO-d_6) δ 10.35 (s, 1 H), 10.30 (s, 1 H), 9.97 (s, 1 H), 9.90 (s, 1 H), 9.82 (s, 1 H), 9.30 (s, 1 H), 9.2 (br s, 1H), 8.02 (m, 3 H), 7.52 (s, 1 H), 7.48 (s, 1 H), 7.21 (m, 2 H), 7.16 (d, 1 H, J = 1.1 Hz), 7.11 (d, 1 H, J = 1.2 Hz), 7.04 (d, 1 H, J = 1.1 Hz), 6.97 (d, 1 H, J = 1.3 Hz), 6.92 (d, 1 H, J = 1.4 Hz), 6.87 (d, 1 H, J = 1.2 Hz), 3.99 (s, 3 H), 3.97 (s, 3 H), 3.83 (s, 3 H), 3.82 (s, 3 H), 3.80 (s, 3 H), 3.79 (s, 3 H), 3.47 (q, 2 H, J = 4.7 Hz), 3.30 (q, 2 H, J = 4.6 Hz), 3.20 (q, 2 H, J = 5.0 Hz), 3.05 (q, 2 H, J = 5.1 Hz), 2.75 (d, 6 H, J = 4.1 Hz), 2.27 (m, 4 H), 2.03 (s, 3 H), 1.74 (m, 4 H); MALDI-TOF-MS 1036.4 (1036.1 calcd for M + H).

AcPyPyPy-γ-ImImPy-β-Dp (4). Polyamide was prepared by machine-assisted solid phase protocols[7] as a white powder (14 mg, 48% recovery): UV λ_{max} 246 (44 400), 312 (52 300); ^{1}H NMR (DMSO-d_6) 10.32 (s, 1 H), 10.28 (s, 1 H), 9.89 (m, 2 H), 9.82 (s, 1 H), 9.18 (s, 1 H), 9.10 (br s, 1 H), 8.03 (m, 3 H), 7.55 (s, 1 H), 7.52 (s, 1 H), 7.21 (d, 1 H, J = 1.1 Hz), 7.18 (d, 1 H, J = 7.16 Hz), 7.15 (d, 1 H, J = 1.0 Hz), 7.12 (d, 1 H, J = 1.0 Hz), 7.02 (d, 1 H, J = 1.0 Hz), 6.92 (d, 1 H, J = 1.1 Hz), 6.87 (d, 1 H, J = 1.1 Hz), 6.84 (d, 1 H, J = 1.0 Hz), 3.97 (s, 3 H), 3.93 (s, 3 H), 3.87 (s, 3 H), 3.80 (s, 3 H), 3.78 (m, 6 H), 3.35 (q, 2 H, J = 5.6 Hz), 3.19 (q, 2 H, J = 5.3 Hz), 3.08 (q, 2 H, J = 5.7 Hz), 2.87 (q, 2 H, J = 5.8 Hz), 2.71 (d, 6 H, J = 4.0 Hz), 2.33 (m, 4 H), 1.99 (s, 3 H), 1.74 (m, 4 H); MALDI-TOF-MS 1036.2 (1036.1 calcd for M + H).

Construction of Plasmid DNA. Using T4 DNA ligase, the plasmid pMEPGG was constructed by ligation of an insert, 5′-GATCGC-

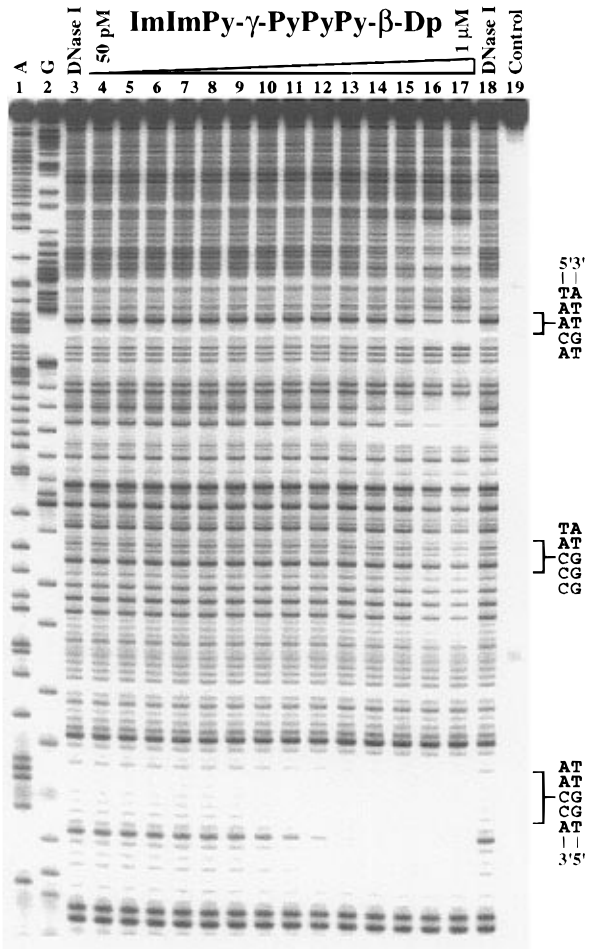


Figure 6. Quantitative DNase I footprint titration experiment with ImImPy-γ-PyPyPy-β-Dp (**1**) on the 3′-^{32}P-labeled 266 base pair *Eco*RI/*Pvu*II restriction fragment from plasmid pMEPGG. Lane 1: A reaction. Lane 2: G reaction. Lanes 3 and 18: DNase I standard. Lanes 4−17: 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 μM ImImPy-γ-PyPyPy-β-Dp (**1**), respectively. Lane 19: intact DNA. The 5′-TGGTT-3′, 5′-GGGTA-3′, and 5′-TGTTA-3′ sites which were analyzed are shown on the right side of the autoradiogram. All reactions contain 10 kpm restriction fragment, 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

GAGCTCAGCGATGGTTTGCGATAGCGTGATTAGC-GTATGCGTGGGTAGCGATACGC-3′ and 5′-GCGTATCGCTAC-CCACGCATACGCTAAT-CACGCTATCGCAAACCATCGCTGA-GCTCGCGATC-3′, into pUC 19 previously cleaved with *Bam*HI and *Hin*dIII. Ligation products were used to transform Epicurian Coli XL 1 Blue competent cells (Stratagene). Colonies were selected for α-complementation on 25 mL Luria-Bertani medium agar plates containing 50 μg/mL ampicillin and treated with XGAL and IPTG solutions. Large-scale plasmid purification was performed with Qiagen purification kits. Plasmid DNA concentration was determined at 260 nm using the relation 1 OD unit = 50 μg/mL duplex DNA. The plasmid was linearized with *Eco*RI, followed by treatment with either Klenow, deoxyadenosine 5′-[α-^{32}P]triphosphate (Amersham), and thymidine 5′-[α-^{32}P]triphosphate for 3′ labeling or calf alkaline phosphatase and subsequent 5’ end labeling with T4 polynucleotide kinase and γ-[^{32}P]dATP. The 3′- or 5′-end-labeled fragment was then digested with *Pvu*II and isolated by nondenaturing gel electrophoresis. The 3′- or 5′-^{32}P-end-labeled 266 base pair *Eco*RI/*Pvu*II restriction fragment was used in all experiments described here. Chemical sequencing reactions were performed according to published protocols.[13] Standard protocols were used for all DNA manipulations.[14]

Identification of Binding Sites by MPE·Fe^{II} Footprinting. All reactions were carried out in a total volume of 40 μL with final concentrations of species as indicated in parentheses. The ligands were added to solutions of radiolabeled restriction fragment (10 000 cpm), calf thymus DNA (100 μM bp), Tris-acetate (25 mM, pH 7.0), and NaCl (10 mM) and incubated for 1 h at 22 °C. A 50 μM MPE·Fe^{II} solution was prepared by mixing 100 μL of a 100 μM MPE solution with a freshly prepared 100 μM ferrous ammonium sulfate solution. Footprinting reactions were initiated by the addition of MPE·Fe^{II} (5 μM), followed 5 min later by the addition of dithiothreitol (5 mM), and allowed to proceed for 15 min at 22 °C. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, and electrophoresed on 8% polyacrylamide denaturing gels (5% cross-link, 7 M urea) at 2000 V for 1 h. The gels were analyzed using storage phosphor technology.

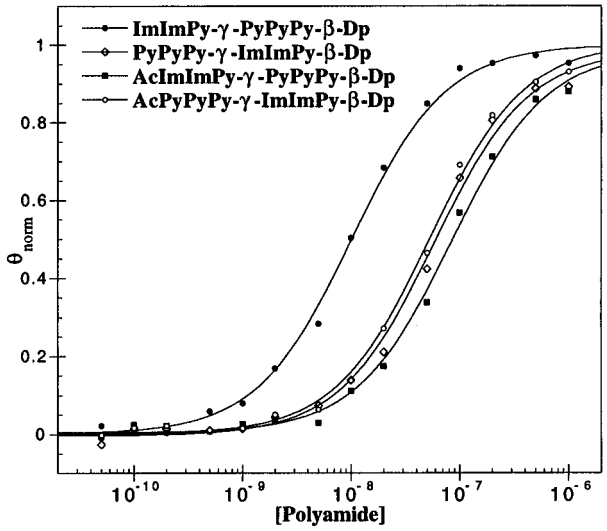


Figure 7. Data for the quantitative DNase I footprint titration experiments for the four polyamides **1**−**4** in complex with the designated 5’-TGGTT-3′ site. The θ_{norm} points were obtained using photostimulable storage phosphor autoradiography and processed as described in the Experimental Section. The data points for ImImPy-γ-PyPyPy-β-Dp (**1**), PyPyPy-γ-ImImPy-β-Dp (**2**), AcImImPy-γ-PyPyPy-β-Dp (**3**), and AcPyPyPy-γ-ImImPy-β-Dp (**4**) are indicated by ●, ◇, ■, and ○, respectively. The solid curves are the best-fit Langmuir binding titration isotherms obtained from a nonlinear least squares algorithm using eq 2.

Analysis of Energetics by Quantitative DNase I Footprint Titration. All reactions were executed in a total volume of 40 μL with final concentrations of each species as indicated. The ligands, ranging from 50 pM to 1 μM, were added to solutions of radiolabeled restriction fragment (10 000 cpm), Tris·HCl (10 mM, pH 7.0), KCl (10 mM), $MgCl_2$ (10 mM), and $CaCl_2$ (5 mM) and incubated for 4 h at 22 °C. Footprinting reactions were initiated by the addition of 4 μL of a stock solution of DNase I (0.025 unit/mL) containing 1 mM dithiothreitol and allowed to proceed for 6 min at 22 °C. The reaction mixtures were stopped by addition of a 3 M sodium acetate solution containing 50 mM EDTA and ethanol precipitated. The reactions were resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer and electrophoresed on 8% polyacrylamide denaturing gels (5% cross-link, 7 M urea) at 2000 V for 1 h. The footprint titration gels were dried and quantitated using storage phosphor technology.

Equilibrium association constants were determined as previously described.[6,10] The data were analyzed by performing volume integrations of the 5′-TGGTT-3′, 5′-TGTTA-3′, and 5′-GGGTA-3′ sites and a reference site. Binding sites are assumed to be independent and noninteracting as they are separated by at least one full turn of the double helix. The apparent DNA target site saturation, θ_{app}, was calculated for each concentration of polyamide using the following equation:

(13) (a) Iverson, B. L.; Dervan, P. B. *Nucleic Acids Res.* **1987**, *15*, 7823−7830. (b) Maxam, A. M.; Gilbert, W. S. *Methods Enzymol.* **1980**, *65*, 499−560.

(14) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.

$$\theta_{app} = 1 - \frac{I_{tot}/I_{ref}}{I_{tot}^{\circ}/I_{ref}^{\circ}} \quad (1)$$

where I_{tot} and I_{ref} are the integrated volumes of the target and reference sites, respectively, and I_{tot}° and I_{ref}° correspond to those values for a DNase I control lane to which no polyamide has been added. The ($[L]_{tot}$, θ_{app}) data points were fitted to a Langmuir binding isotherm (eq 2, $n = 1$) by minimizing the difference between θ_{app} and θ_{fit}, using the modified Hill equation:

$$\theta_{fit} = \theta_{min} + (\theta_{max} - \theta_{min})\frac{K_a^{\,n}[L]_{tot}^{\,n}}{1 + K_a^{\,n}[L]_{tot}^{\,n}} \quad (2)$$

where $[L]_{tot}$ corresponds to the total polyamide concentration, K_a corresponds to the association constant, and θ_{min} and θ_{max} represent the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. The concentration of DNA used for quantitative footprint titrations is ≤50 pM, which justifies the assumption that free ligand concentration is approximately equal to total ligand concentration.[10] Data were fitted using a nonlinear least squares fitting procedure of KaleidaGraph software (version 2.1, Abelbeck software) running on a Power Macintosh 6100/60AV computer with K_a, $\theta_{max,}$ and θ_{min} as the adjustable parameters. The goodness-of-fit of the binding curve to the data points is evaluated by the correlation coefficient, with $R > 0.97$ as the criterion for an acceptable fit. At least three sets of acceptable data were used in determining each association constant. All lanes from each gel were used unless visual inspection revealed a data point to be obviously flawed relative to neighboring points. The data were normalized using the following equation:

$$\theta_{norm} = \frac{\theta_{app} - \theta_{min}}{\theta_{max} - \theta_{min}} \quad (3)$$

Quantitation by Storage Phosphor Technology Autoradiography. Photostimulable storage phosphorimaging plates (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) were pressed flat against gel samples and exposed in the dark at 22 °C for 12−16 h. A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens. The data were analyzed by performing volume integrations of all bands using the ImageQuant version 3.2 software running on an AST Premium 386/33 computer.

Acknowledgment. We are grateful to the National Institutes of Health (GM-27681) for research support, the National Institutes of Health for a research service award to M.E.P., and the Howard Hughes Medical Institute for a predoctoral fellowship to E.E.B.

JA9607289