Orientation Preferences of Pyrrole−Imidazole Polyamides in the Minor Groove of DNA

Sarah White, Eldon E. Baird, and Peter B. Dervan*

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

Received May 14, 1997[⊗]

Abstract: In order to determine whether there is an orientation preference of pyrrole−imidazole (Py−Im) polyamide dimers with respect to the 5′−3′ direction of the backbone in the DNA helix, equilibrium association constants (K_a) were determined for a series of six-ring hairpin polyamides which differ with respect to substitution at the N and C termini. Affinity cleaving experiments using hairpin polyamides of core sequence composition ImPyPy-γ-PyPyPy with an EDTA·Fe(II) moiety at the C-terminus reveal a single binding orientation at each formal match site, 5′-(A,T)G(A,T)$_3$-3′ and 5′-(A,T)C(A,T)$_3$-3′. A positive charge at the C-terminus and no substitution at the N-terminus imidazole affords the maximum binding orientation preference, calculated from K_a(5′-TGTTA-3′)/K_a(5′-TCTTA-3′), with the N-terminal end of each three-ring subunit located toward the 5′ side of the target DNA strand. Removal of the positive charge, rearrangement of the positive charge to the N-terminus or substitution at the N-terminal imidazole decreases the orientation preference. These results suggest that second generation design principles superimposed on the simple pairing rules can further optimize the sequence-specificity of Py-Im polyamides for double helical DNA.

Polyamides containing pyrrole (Py) and imidazole (Im) amino acids bind cooperatively as antiparallel dimers in the minor groove of the DNA helix.[1,2] Sequence-specificity depends on the side-by-side pairings of *N*-methylpyrrole and *N*-methylimidazole amino acids.[1] A pairing of Im opposite Py targets a G·C base-pair, while Py opposite Im targets a C·G base-pair.[1] A pyrrole/pyrrole combination is degenerate and targets both T·A and A·T base-pairs.[2] Py−Im polyamides have been shown to be cell permeable and to inhibit the transcription of genes in cell culture.[3] This provides impetus to develop second generation polyamide design rules that provide for enhanced sequence-specificity and perhaps optimal biological regulation.

Although the polyamides bind DNA antiparallel to each other, the "pairing rules" do not distinguish whether there should be any energetic preference for alignment of each polyamide (N−C) with respect to the backbone (5′−3′) of the DNA double helix (Figure 1). In a formal sense the homodimer $(ImPyPy)_2$ could bind 5′-WGWCW-3′ or 5′-WCWGW-3′ and still not violate the binary code. Remarkably, even in the first report on the binding specificity of the three ring polyamide ImPyPy-Dp there were qualitative data to suggest that there was indeed a binding preference 5′-WGWCW-3′ > 5′-WCWGW-3′.[1a,4] This suggested that pyrrole−imidazole polyamide dimers align N−C with the 5′−3′ direction of the DNA strand. This orientation preference superimposed on the pairing rules confers added specificity by breaking a potential degeneracy for recognition. It would be useful to find out whether this preference is general and which aspects of the ligand design control the energetics of orientation preference. Therefore we describe here a study to address the influence on orientation of (1) positive charge or lack of, (2) position of the positive charge at the N- or C-terminus, and (3) substitution of the terminal imidazole.

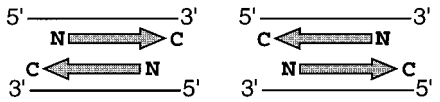

Figure 1. Antiparallel polyamide subunits are depicted as filled arrows. Arrowheads correspond to C-terminal end of the polyamide: (a) polyamide binding with N-terminal end located toward 5′-side of the targeted DNA strand and (b) binding with the C-terminal end of the polyamide located toward 5′-side of binding site.

Three-ring polyamide subunits covalently coupled by a γ-aminobutyric acid linker form six-ring hairpin structures that bind to 5-bp target sequences with enhanced affinity and specificity relative to the unlinked polyamide pair.[5] In principle, a hairpin polyamide:DNA complex can form at two different DNA sequences depending on the N−C alignment of the polyamide with the walls of the minor groove of DNA (5′−3′). A six-ring hairpin polyamide of core sequence composition ImPyPy-γ-PyPyPy which places the N-terminus of each three-ring polyamide subunit at the 5′-side of each recognized DNA strand would bind 5′-TGTTA-3′. Placement of the polyamide N-terminus at the 3′ side of each recognized strand would result in targeting of a 5′-TCTTA-3′ sequence (Figure 2).

Four six-ring hairpin polyamides, ImPyPy-γ-PyPyPy-β-Dp **1**, ImPyPy-γ-PyPyPy-β-EtOH **2**, Ac-ImPyPy-γ-PyPyPy-β-Dp **3**, and Dp-ImPyPy-γ-PyPyPy-β-Me **4**, were synthesized by solid phase methods (Figure 3).[6] The corresponding EDTA analogs ImPyPy-γ-PyPyPy-β-Dp-EDTA **1-E**, ImPyPy-γ-PyPyPy-β-C7-

[⊗] Abstract published in *Advance ACS Abstracts,* September 15, 1997.

(1) (a) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783. (b) 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. (c) Wade, W. S.; Mrksich, M.; Dervan, P. B. *Biochemistry* **1993**, *32*, 11385. (d) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572 (e) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559.

(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. (d) White, S.; Baird, E. E.; Dervan, P. B. *Biochemistry* **1996**, *35*, 12532.

(3) Gottesfield, J. M.; Nealy, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202.

(4) See Table 1, ref 1a.

(5) (a) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983. (b) Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147. (c) de Claire, R. P. L.; Geierstanger B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7906.

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

S0002-7863(97)01569-2 CCC: $14.00

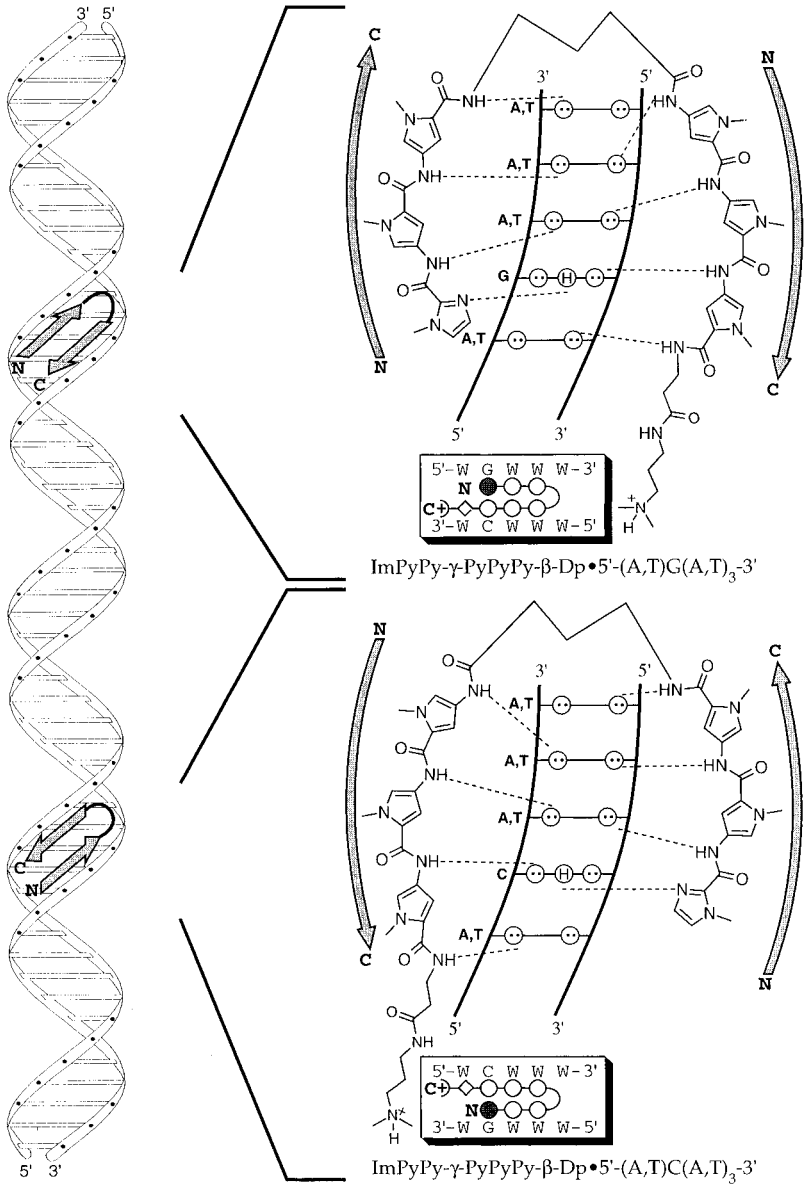

Figure 2. Binding models for ImPyPy-γ-PyPyPy-β-Dp in complex with 5′-WGWWW-3′ (top) and 5′-WCWWW-3′ (bottom) (W = A or T). Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. For schematic binding models, the imidazole and pyrrole rings are represented as shaded and unshaded spheres, respectively, the curved line represents γ-aminobutyric acid, and the β-alanine residue is represented as an unshaded diamond. Shaded arrows represent the orientation of individual polyamide subunits. Arrowheads represent the polyamide C-terminus. (left) A ribbon model depicting the hairpin structure bound in the minor groove of the DNA helix with either the N or C-terminus located at the 5′ side of the binding site.

EDTA **2-E**, Ac-ImPyPy-γ-PyPyPy-β-Dp-EDTA **3-E**, and Dp-ImPyPy-γ-PyPyPy-β-C7-EDTA **4-E** were also constructed in order to confirm a single orientation of each hairpin:DNA complex. We report here the DNA-binding affinity, orientation, and sequence-selectivity of the four polyamides for the two match five base-pair binding sites, 5′-TGTTA-3′ and 5′-TCTTA-3′. Three separate techniques are used to characterize the DNA-binding properties of the polyamides: affinity cleaving[7] and MPE·Fe(II)[8] and DNase I[9] footprinting. Affinity cleavage studies determine the specific binding orientation and stoichiometry of each hairpin:DNA complex. Binding site size is accurately determined by MPE·Fe(II) footprinting, while quantitative DNase I footprint titration is more suitable for measurement of equilibrium association constants (K_a) for the polyamide binding to designated sequences.

Results

Synthesis. Polyamides were synthesized from Boc-β-alanine-Pam-resin (0.2 mmol/g substitution) using stepwise solid phase methods (Figure 4).[6] A C-terminal β-alanine residue facilitates solid phase synthesis and increases hairpin-polyamide DNA-binding affinity and sequence specificity.[5b] The synthesis of polyamides **1** and **3** has been described.[5b] A sample of ImPyPy-γ-PyPyPy-β-Pam-resin was cleaved by a single-step aminolysis reaction with neat ethanolamine (55 °C, 16 h) to provide polyamide **2** after HPLC purification. Cleavage of Dp-ImPyPy-γ-PyPyPy-β-resin with a saturated solution of methylamine in DMF (55 °C, 80 psi, 48 h) provided polyamide **4** after HPLC purification. For the synthesis of analogs modified with EDTA, a sample of resin was cleaved with 3,3′-diamino-*N*-methyl-

(7) (a) Taylor, J. S.; Schultz, P. G.; Dervan, P. B. *Tetrahedron* **1984**, *40*, 457. (b) Dervan, P. B. *Science* **1986**, *232*, 464.

(8) (a) Van Dyke, M. W.; Dervan, P. B. *Nucl. Acids Res.* **1983**, *11*, 5555. (b) Van Dyke, M. W.; Dervan, P. B. *Science* **1984**, *225*, 1122.

(9) (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.

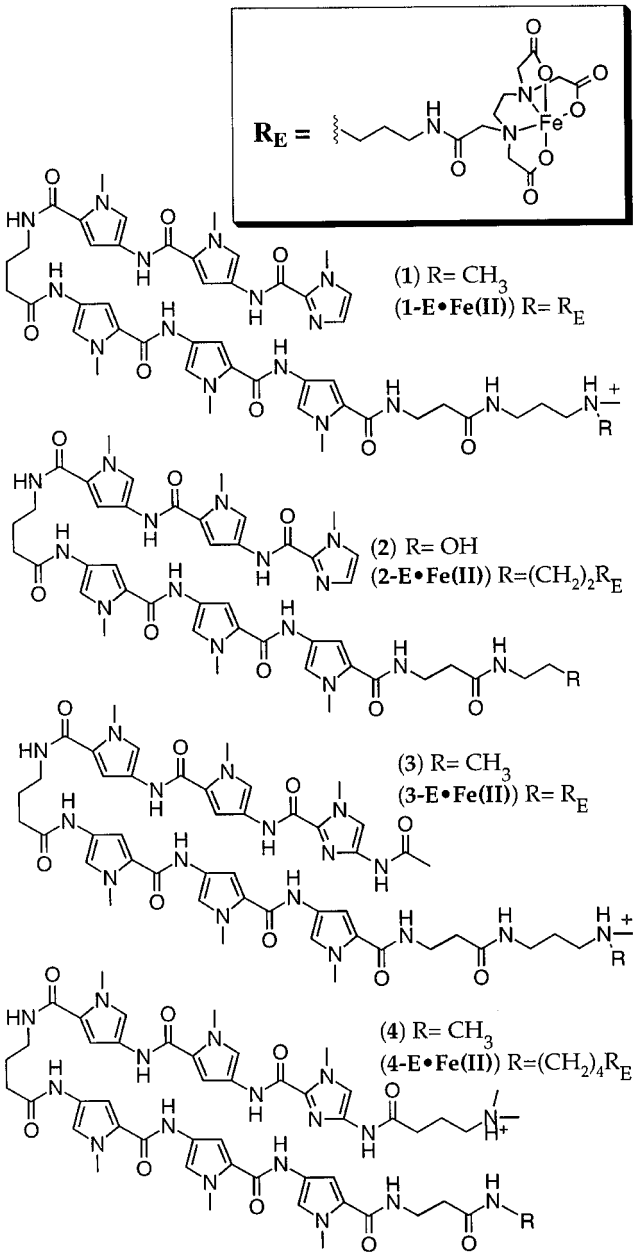

Figure 3. Structure of the hairpin polyamides **1**−**4** and the corresponding EDTA-modified derivatives: ImPyPy-γ-PyPyPy-β-Dp (**1**), ImPyPy-γ-PyPyPy-β-Dp-EDTA (**1-E**), ImPyPy-γ-PyPyPy-β-EtOH (**2**), ImPyPy-γ-PyPyPy-β-C7-EDTA (**3-E**), AcImPyPy-γ-PyPyPy-β-Dp (**3**), AcImPyPy-γ-PyPyPy-β-Dp-EDTA (**3-E**), Dp-ImPyPy-γ-PyPyPy-β-Me (**4**), Dp-ImPyPy-γ-PyPyPy-β-C7-EDTA (**4-E**).

dipropylamine or 1,7-diaminoheptane (55 °C, 18 h) for polyamides **1-E** and **3-E** and polyamides **2-E** and **4-E**, respectively. The amine modified polyamides were purified by reverse phase HPLC and then treated with an excess of the dianhydride of EDTA (DMSO/NMP, DIEA, 55 °C, 30 min) and the remaining anhydride hydrolyzed (0.1 M NaOH, 55 °C, 10 min). The EDTA modified polyamides **1-E, 2-E, 3-E**, and **4-E** were then isolated by reverse phase HPLC. Polyamides were characterized by a combination of analytical HPLC, ^{1}H NMR spectroscopy, and MALDI-TOF mass spectroscopy.

Identification of Binding Sites by MPE·Fe(II) Footprinting. MPE·Fe(II) footprinting experiments were used to determine precise binding site location and size.[8] Footprinting on the 3′- and 5′-^{32}P end labeled 266 base pair *Eco*RI/*Pvu*II restriction fragment from the plasmid pDEH2 (25 mM tris-acetate, 10 mM NaCl, 100 μM bp calf thymus DNA, 5 mM DTT, pH 7.0 and 22 °C) reveals that the polyamides are binding to the two match binding sites, 5′-TGTTA-3′ and 5′-TCTTA-3′ (Figures 5 and 6). Polyamides **1** and **2** at 5 μM concentration bind to the 5′-TGTTA-3′ site and bind weakly to the 5′-TCTTA-3′ site. Polyamides **3** and **4** at 10 μM concentration bind to both the 5′-TGTTA-3′ site and the 5′-TCTTA-3′ site.[10] The size of the footprint cleavage protection patterns for the polyamides are consistent with the expected binding location and 5-bp binding site size.

Identification of Binding Orientation and Stoichiometry by Affinity Cleaving. Affinity cleaving titration experiments[7] using hairpin polyamides modified with EDTA·Fe(II) at the C-terminus were used to determine polyamide binding orientation and stoichiometry. Affinity cleaving experiments were performed on a 3′- and 5′-^{32}P end labeled 266 base *Eco* RI/*Pvu* II restriction fragment from the plasmid pDEH2 (25 mM tris-acetate, 20 mM NaCl, 100 μM bp calf thymus DNA, pH 7, 22 °C, 10 mM DTT, 10 μM Fe(II)) (Figures 7 and 8). The observed cleavage patterns are in all cases 3′-shifted, consistent with minor groove occupancy. A single cleavage locus proximal to the 5′-side of both the 5′-TGTTA-3′ and 5′-TCTTA-3′ binding sites is consistent with a single orientation at each site (Figure 9). A number of additional strong cleavage sites present on the DNA restriction fragment which correspond to sequences of the form 5′-(A,T)G(A,T)$_3$-3′ were not resolved.[2d]

Analysis of Energetics by Quantitative DNase I Footprint Titrations. Quantitative DNase I footprint titration experiments[9] (10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$, pH 7.0, 22 °C) were performed to determine the equilibrium association constants K_a for recognition of the bound sites (Figure 10). The 5′-TGTTA-3′ site is bound by the polyamides with decreasing affinity: ImPyPy-γ-PyPyPy-β-Dp (**1**) > ImPyPy-γ-PyPyPy-β-EtOH (**2**) > AcImPyPy-γ-PyPyPy-β-Dp (**3**) > Dp-ImPyPy-γ-PyPyPy-β-Me (**4**). The 5′-TCTTA-3′ site is bound with decreasing affinity: Dp-ImPyPy-γ-PyPyPy-β-Me (**4**) > AcImPyPy-γ-PyPyPy-β-Dp (**3**) > ImPyPy-γ-PyPyPy-β-EtOH (**2**) > ImPyPy-γ-PyPyPy-β-Dp (**1**). Remarkably, the ratio of association constants for each site varies from 16- to 1-fold between the four polyamides indicating a sensitivity to substitution at the N- and C-terminus (Table 1).

Discussion

Binding Site Size and Orientation. MPE·Fe(II) footprinting reveals that the four polyamides of core sequence composition ImPyPy-γ-PyPyPy bind with high affinity to both match sites 5′-TGTTA-3′ and 5′-TCTTA-3′. Affinity cleavage experiments using polyamides with Fe(II)·EDTA at the carboxy terminus confirm that the four polyamides bind each discreet site with a single orientation (Figures 8 and 9). Asymmetric 3′-shifted cleavage patterns are consistent with the location of the 1:1 polyamide:DNA complex in the minor groove. The observation of a single cleavage locus is consistent only with an oriented 1:1 complex and rules out any 2:1 overlapped or extended binding motifs.[11] A 1:1 oriented but extended motif would require at least an eight base-pair binding site, which is inconsistent with high-resolution MPE footprinting data on both target sites. The hairpin structure is supported by direct NMR structure studies on a six-ring hairpin polyamide of sequence composition ImPyPy-γ-PyPyPy binding a core five base-pair 5′-TGTTA-3′ site.[5c]

A single cleavage locus is observed proximal to the 5′-side of both the 5′-TGTTA-3′ and 5-TCTTA-3′ binding sites,

(10) We note that no conclusion regarding polyamide binding affinities can be drawn from these experiments due to specific and nonspecific interactions of the polyamides with the calf thymus carrier DNA.

(11) (a) Trauger, J. W.; Baird, E. E.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6160. (b) Geierstanger, B. H.; Mrksich M.; Dervan, P. B.; Wemmer, D. E. *Nature Struct. Biol.* **1996**, *3*, 321.

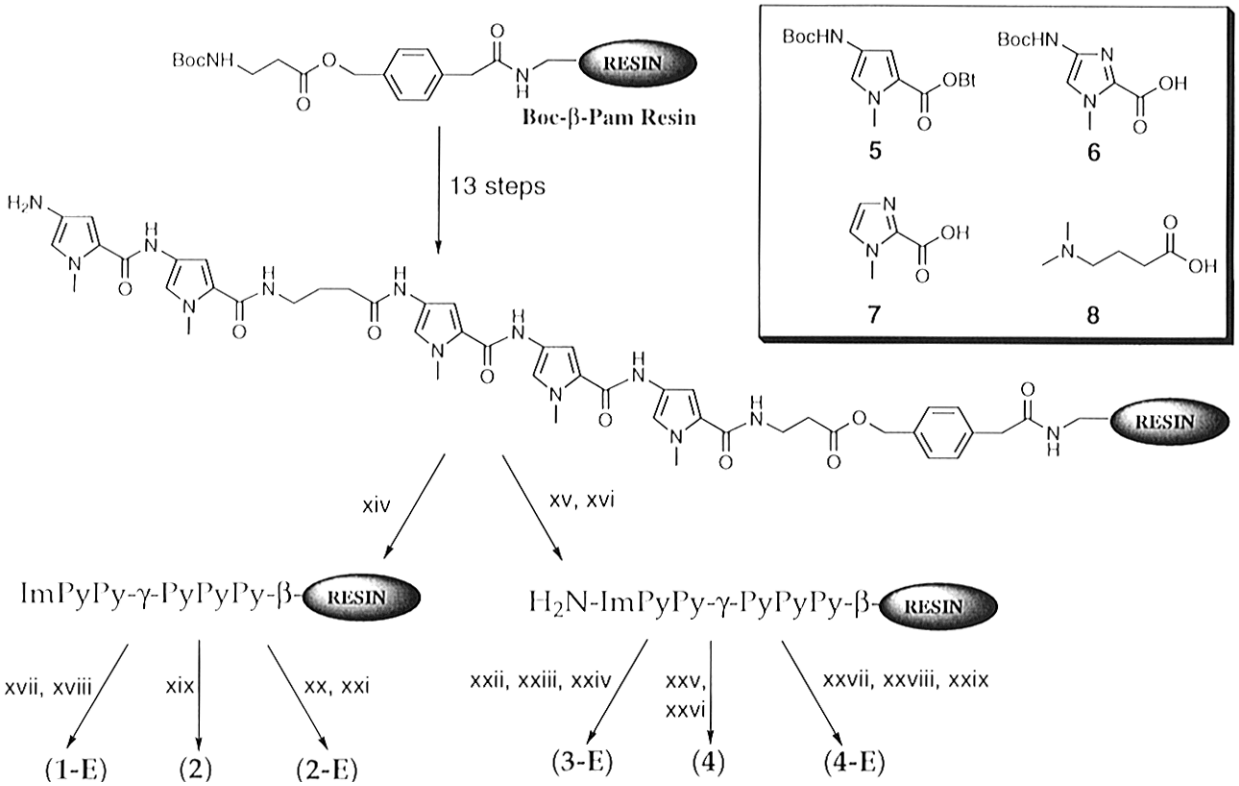

Figure 4. (box) Pyrrole and imidazole monomers used for synthesis of the polyamides described here: Boc-pyrrole-OBt ester **5**, Boc-imidazole-acid **6** and imidazole-2-carboxylic acid **7**.[1a] Solid phase synthetic scheme for ImPyPy-γ-PyPyPy-β-Dp-EDTA (**1-E**), ImPyPy-γ-PyPyPy-β-EtOH (**2**), ImPyPy-γ-PyPyPy-C7-EDTA (**2-E**), AcImPyPy-γ-PyPyPy-β-Dp-EDTA (**3-E**), Dp-ImPyPy-γ-PyPyPy-β-Me (**4**), and Dp-ImPyPy-γ-PyPyPy-β-C7-EDTA (**4-E**). Synthesis is initiated from commercially available Boc-β-alanine-Pam-resin (0.2 mmol/g): (i) 80% TFA/DCM, 0.4 M PhSH; (ii) BocPy-OBt, DIEA, DMF; (iii) 80% TFA/DCM, 0.4 M PhSH; (iv) BocPy-OBt, DIEA, DMF; (v) 80% TFA/DCM, 0.4 M PhSH; (vi) BocPy-OBt, DIEA, DMF; (vii) 80% TFA/DCM, 0.4 M PhSH; (viii) Boc-γ-aminobutyric acid (HBTU, DIEA), DMF; (ix) 80% TFA/DCM, 0.4 M PhSH; (x) BocPy-OBt, DIEA, DMF; (xi) 80% TFA/DCM, 0.4 M PhSH; (xii) BocPy-OBt, DIEA, DMF; (xiii) 80% TFA/DCM, 0.4 M PhSH; (xiv) imidazole-2-carboxylic acid (HBTU/DIEA); (xv) BocIm-OBt (DCC/HOBt), DIEA, DMF; (xvi) 80% TFA/DCM, 0.4 M PhSH; (xvii) 3,3′-diamino-*N*-methyldipropylamine, 55 °C; (xviii) EDTA-dianhydride, DMSO/NMP, DIEA, 55 °C; 0.1M NaOH; (xix) ethanol, 55 °C; (xx) 1,7-diaminoheptane, 55 °C; (xxi) EDTA-dianhydride, DMSO/NMP, DIEA, 55 °C; 0.1 M NaOH (xxii) acetic anhydride, DMF, DIEA; (xxiii) 3,3′-diamino-*N*-methyldipropylamine, 55 °C; (xxiv) EDTA-dianhydride, DMSO/NMP, DIEA, 55 °C; 0.1 M NaOH; (xxv) dimethylamino-γ (**8**), (HBTU/DIEA); (xxvi) $MeNH_2$ (saturated), DMF, 80 psi, 55 °C (xxvii) dimethylamino-γ (**8**), (HBTU/DIEA); (xxviii) 1,7-diaminoheptane, 55 °C; (xxix) EDTA-dianhydride, DMSO/NMP, DIEA, 55 °C; 0.1 M NaOH

indicating that the carboxy terminus of the polyamide is located at the 5′-side of each site. Furthermore, the relative location of the observed cleavage maxima is unchanged for recognition of 5′-TGTTA-3′ and 5′-TCTTA-3′, indicating similar placement of the polyamide C-termini at both sites. *These results indicate that all four polyamides may adapt two unique binding orientations.* Each binding orientation may represent a unique and distinguishable hairpin fold (Figure 2).[12]

Binding Affinity. Four polyamides of core sequence composition ImPyPy-γ-PyPyPy-β, but varying at the N- and C-terminus, bind both match sites 5′-TGTTA-3′ and 5′-TCTTA-3′ as a 1:1 complex (eq 2, $n = 1$) consistent with the hairpin motif. However, the relative discrimination between the two sites varies by 16-fold. Among the four ligands, polyamide (**1**) binds the 5′-TGTTA-3′ site with the highest affinity ($K_a = 1.4 \times 10^7\ M^{-1}$) and orientation specificity (16-fold). Replacement of the charged dimethylaminopropylamide tail group with an uncharged ethoxyamide group as in polyamide (**2**) results in a negligible decrease in affinity ($K_a = 1.1 \times 10^7\ M^{-1}$) but reduced specificity (9-fold). The decrease in orientational specificity indicates that the cationic tail group is necessary but not sufficient for optimal oriented polyamide binding.

In previous work we have observed that N-terminal acetylation reduces both the binding affinity and the sequence preference for target sites.[5b] The acetylated polyamide (**3**) binds the 5′-TGTTA-3′ site with reduced affinity and orientational specificity compared to (**1**). Replacement of the N-terminal acetyl group with a charged dimethylaminopropyl group results in a 4-fold increase in affinity at the 5′-TCTTA-3′ site but a complete loss of orientational specificity for polyamide (**4**). These results indicate that the N- or C-terminus position of charged substituents plays an important role in orientation preference.

Subunit Dipoles. Dipole−dipole interactions could potentially account in part for the observed DNA-binding orientation preference (Figure 11). Calculated dipoles[13] for the individual monomer units AcIm, AcPy, and Im as well as the three-ring subunits ImPyPy and AcImPyPy reveal that only monomers and subunits containing an unsubstituted Im residue could have a dipole antiparallel to the walls of the helix. Subunit dipoles could stabilize side-by-side antiparallel placement of polyamide subunits in the minor groove. In addition, dipole interactions between the DNA backbone and the polyamide could lead to 5′ to 3′ aligned with N to C oriented binding *if there exists a net dipole on each antiparallel strand of the DNA double helix*.[14]

Induced Chirality. Pyrrole−imidazole polyamides are achiral molecules which form chiral complexes with DNA. In principle there exists two nonsuperimposable hairpin folds which

(12) Mismatched pairing of Im/Py opposite C·G cannot be ruled out for recognition of 5′-TCTTA-3′ sites; however, mismatches of this class are energetically unfavorable and have been found to reduce DNA binding by at least 3 kcal/mol.

(13) Dipoles were calculated using MacSpartan v 1.0, see: Mecozzi, S.; West, A. P.; Dougherty, D. A. *J. Am. Chem. Soc.* **1996**, *93*, 10566−10571.

(14) Norden, B. Personal communication.

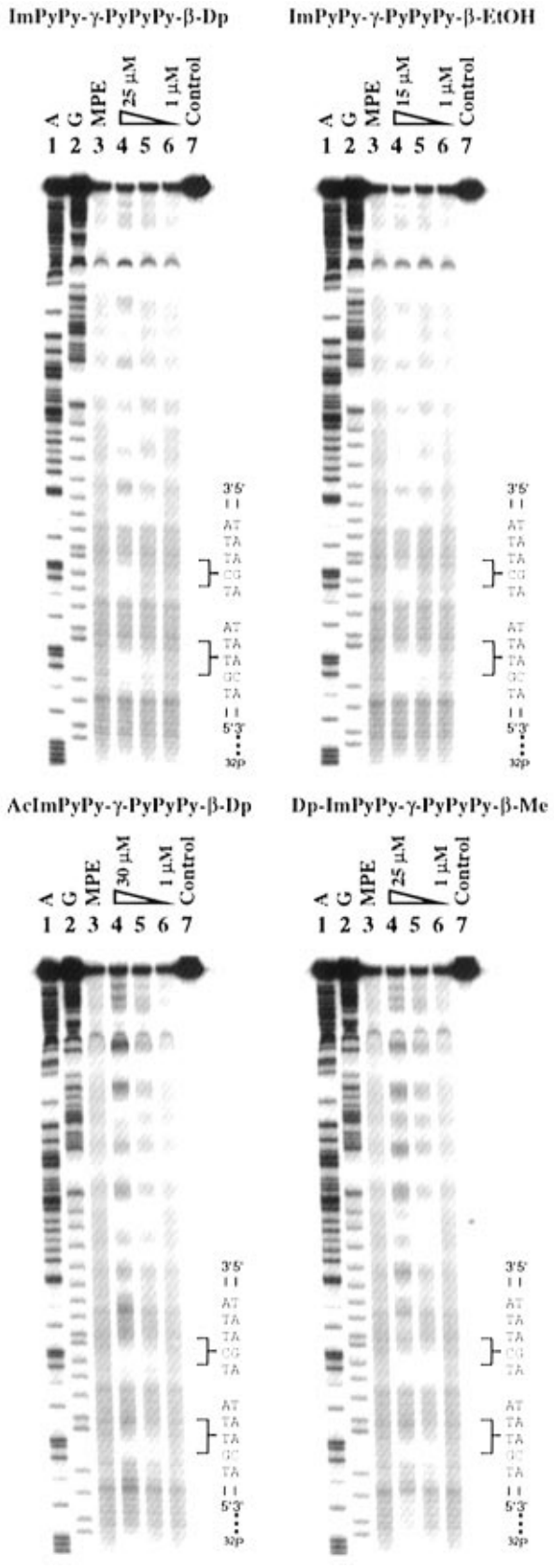

Figure 5. MPE·Fe(II) footprinting experiments on the 3′-^{32}P-labeled 266-bp*Eco* RI/*Pvu* II restriction fragment from plasmid pDEH2. The 5′-TGTTA-3′ and 5′-TCTTA-3 sites are shown on the right side of the autoradiog.[21] Lane 1, A reaction; lane 2, G reaction; lane 3 MPE·Fe(II) standard; lane 7, intact DNA; lanes 4−6: 25 μM, 5 μM, 1 μM ImPyPy-γ-PyPyPy-β-Dp (**1**), 15 μM, 5 μM, 1 μM ImPyPy-γ-PyPyPy-β-EtOH (**2**), 30 μM, 10 μM, 1 μM AcImPyPy-γ-PyPyPy-β-Dp (**3**), or 25 μM, 10 μM, 1 μM Dp-ImPyPy-γ-PyPyPy-β-Me (**4**). All lanes contain 15 kcpm 3′-radiolabeled DNA, 25 mM Tris-acetate buffer (pH 7.0), 10 mM NaCl, and 100 μM/base pair calf thymus DNA.

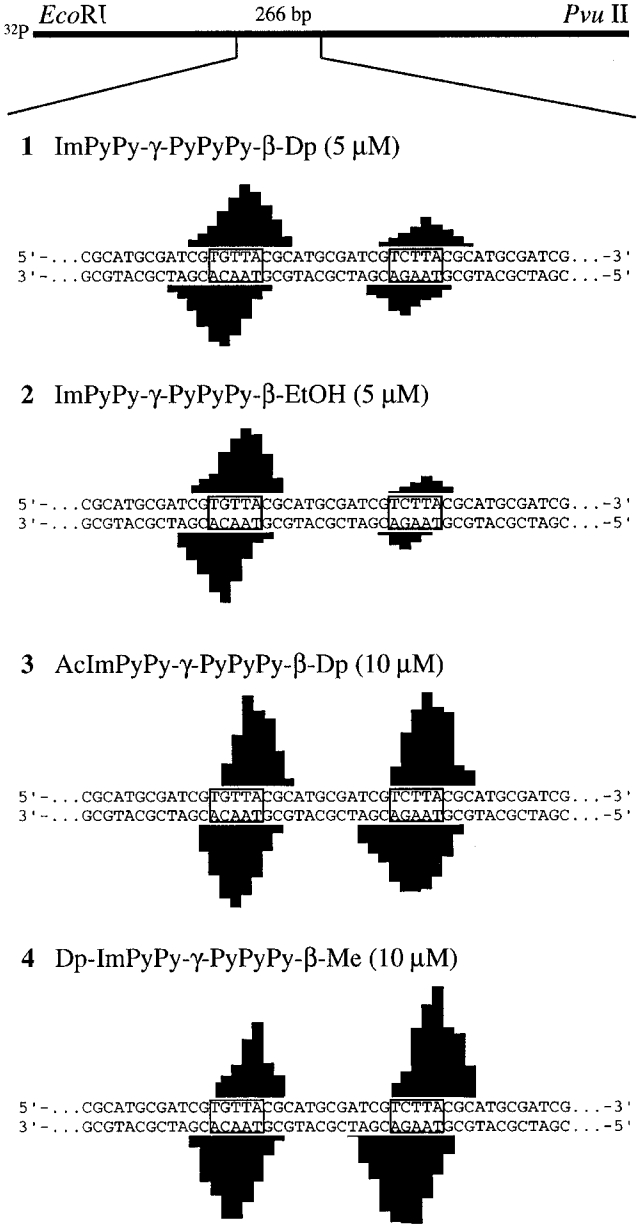

Figure 6. (Top) Illustration of the 266 bp restriction fragment with the position of the sequence indicated. (Bottom) MPE·Fe(II) protection patterns of ImPyPy-γ-PyPyPy-β-Dp (**1**) or ImPyPy-γ-PyPyPy-β-EtOH (**2**) for 5 μM concentration and AcImPyPy-γ-PyPyPy-β-Dp (**3**) or Dp-ImPyPy-γ-PyPyPy-β-Me (**4**) for 10 μM concentration. Bar heights are proportional to the relative protection from cleavage at each band.

are related by mirror plane symmetry (Figure 2). Fold 1 is responsible for the preferred 5′ to 3′ N to C orientation for pyrrole−imidazole polyamide DNA recognition. Fold 2 corresponds to the 3′ to 5′ N to C recognition observed here. In the absence of DNA each fold should be energetically equivalent. Upon binding to DNA, adjacent pyrrole and imidazole monomers twist to accommodate the right handed B-form DNA helix. Twisting of polyamide monomers results in an induced asymmetry upon binding. A hairpin polyamide−DNA complex composed of an asymmetrically folded polyamide may be expected to display differential energetics for oriented binding.

Implications for the Design of Minor Groove Binding Polyamides. The results reported here establish that a hairpin polyamide binds with at least a 16-fold orientation preference with the N-termini of the polyamide subunits located at the 5′-end of the targeted DNA strand. Although the cationic dimethylaminopropyl end group is not required for oriented hairpin polyamide recognition in the minor groove of DNA,

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

polyamide (N→C)		5′-TGTTA-3′	5′-TCTTA-3′	orientation[c]
ImPyPy-γ-PyPyPy-β-Dp	**1**	1.4×10^7 (0.4)	8.8×10^5 (1.0)	16
ImPyPy-γ-PyPyPy-β-EtOH	**2**	1.1×10^7 (0.2)	1.2×10^6 (1.0)	9
AcImPyPy-γ-PyPyPy-β-Dp	**3**	7.6×10^6 (1.1)	1.9×10^6 (1.2)	4
Dp-ImPyPy-γ-PyPyPy-β-Me	**4**	7.2×10^6 (1.4)	9.1×10^6 (0.1)	0.8

[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 $CaCl_2$. [c] Orientation preference calculated from K_a(5′-TGTTA-3′)/K_a(5′-TCTTA-3′).

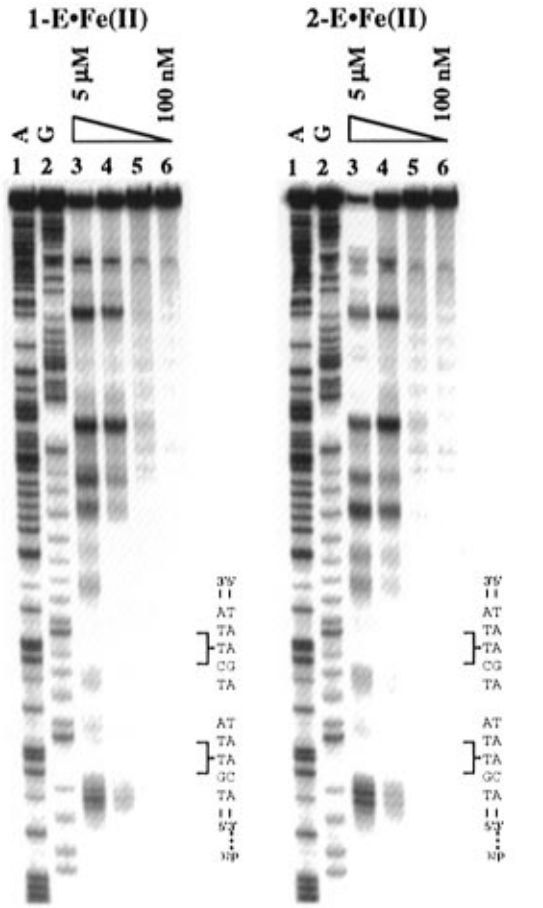

Figure 7. Storage phosphor autoradiogram of 8% denaturing polyacrylamide gels used to separate the fragments[21] generated by affinity cleaving experiments performed with (a) ImPyPy-γ-PyPyPy-β-Dp-EDTA·Fe(II) and (b) AcImPyPy-γ-PyPyPy-β-Dp-EDTA·Fe(II): lanes 1 and 2, A and G sequencing lanes; lanes 3−6, digestion products obtained in the presence of **2**: 5 μM, 2.5 μM, 1 μM, and 100 nM polyamide. The targeted binding sites are indicated on the right side of the autoradiograms. All reactions contain 20 kcpm 3′-^{32}P restriction fragment, 25 mM tris-actetate, 20 mM NaCl, 100 μM bp calf thymus DNA, pH 7.

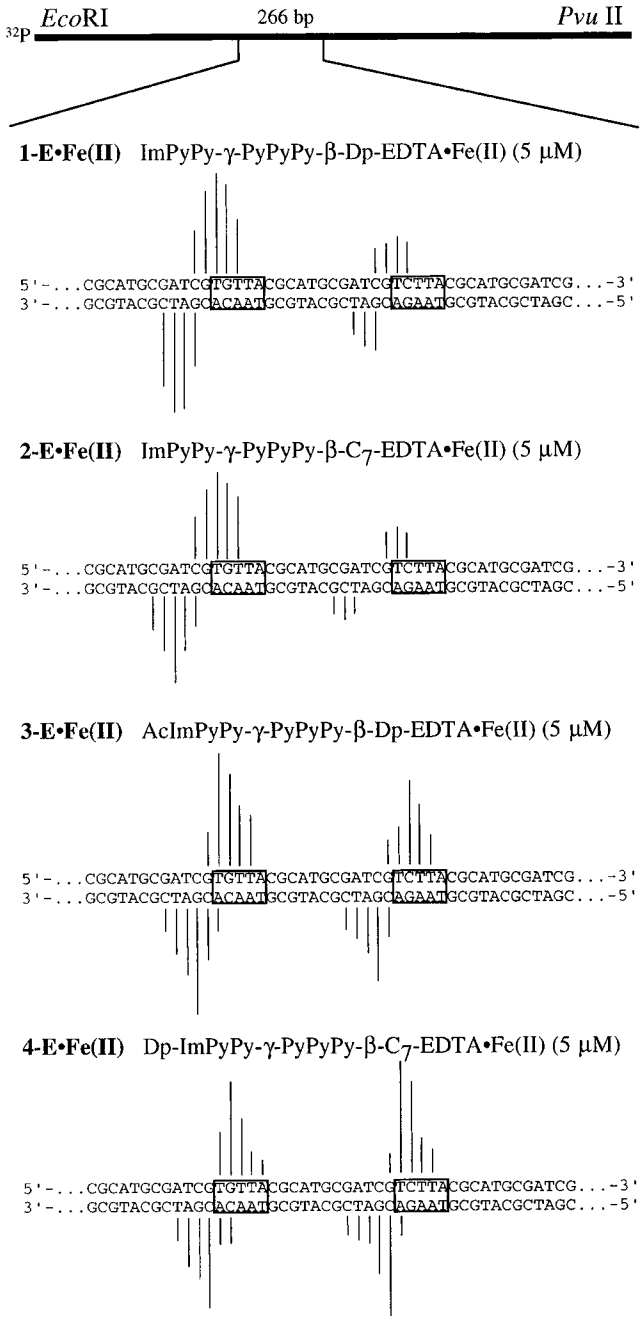

Figure 8. Results from affinity cleavage with ImPyPy-γ-PyPyPy-β-Dp-EDTA (**1-E**), ImPyPy-γ-PyPyPy-β-C7-EDTA (**2-E**) AcImPyPy-γ-PyPyPy-β-Dp-EDTA (**3-E**), and Dp-ImPyPy-γ-PyPyPy-β-C7-EDTA (**4-E**) at 5 μM concentration. (Top) Illustration of the 266 bp restriction fragment with the position of the sequence indicated. Only the two designated target sites are boxed. Arrow heights are proportional to the relative cleavage intensities at each base pair.

the N- or C-terminus position of the charge appears important. Pyrrole−imidazole polyamide DNA-binding orientation preference defines a new design feature which must be considered for application of the pairing rules for DNA targeting. For optimal orientation and hence sequence-specificity, a positive charge at the C-terminus and no substitution of an N-terminal imidazole is preferred.

Experimental Section

Dicyclohexylcarbodiimide (DCC), hydroxybenzotriazole (HOBt), 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 0.2 mmol/g Boc-β-alanine(4-carboxamidomethyl)-benzylestercopoly(styrene-divinylbenzene) resin (Boc-β-Pam-Resin), and Boc-γ-aminobutyric acid were purchased from Peptides International. *N*,*N*-Diisopropylethylamine (DIEA), *N*,*N*-dimethylformamide (DMF), *N*-methylpyrrolidone (NMP), DMSO/NMP, acetic anhydride (Ac_2O), and 0.0002 M potassium cyanide/pyridine were purchased from Applied Biosystems. Dichloromethane (DCM) and triethylamine (TEA) was reagent grade from EM, thiophenol (PhSH) and dimethylaminopropylamine were from Aldrich, trifluoroacetic acid (TFA) was from

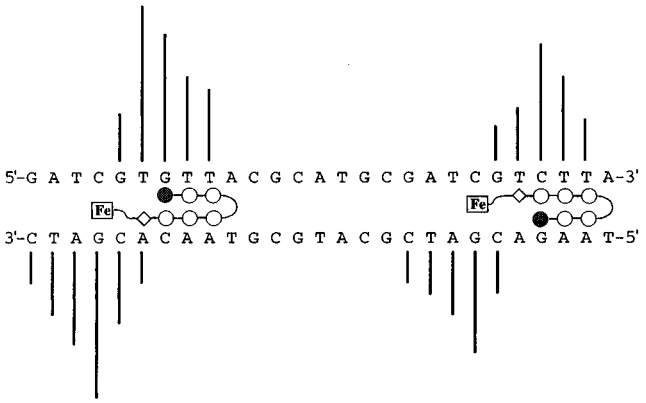

Figure 9. Affinity cleavage patterns and ball and stick models of the six-ring EDTA·Fe(II) analog **1-E·Fe(II)** bound to the match sites, 5′-TGTTA-3′ and 5′-TCTTA-3′. Bar heights are proportional to the relative cleavage intensities at each base pair. Shaded and nonshaded circles denote imidazole and pyrrole carboxamides, respectively. Nonshaded diamonds represent the β-alanine residue. The boxed **Fe** denotes the EDTA·Fe(II) cleavage moiety. Cleavage patterns show a single binding orientation at each match site.

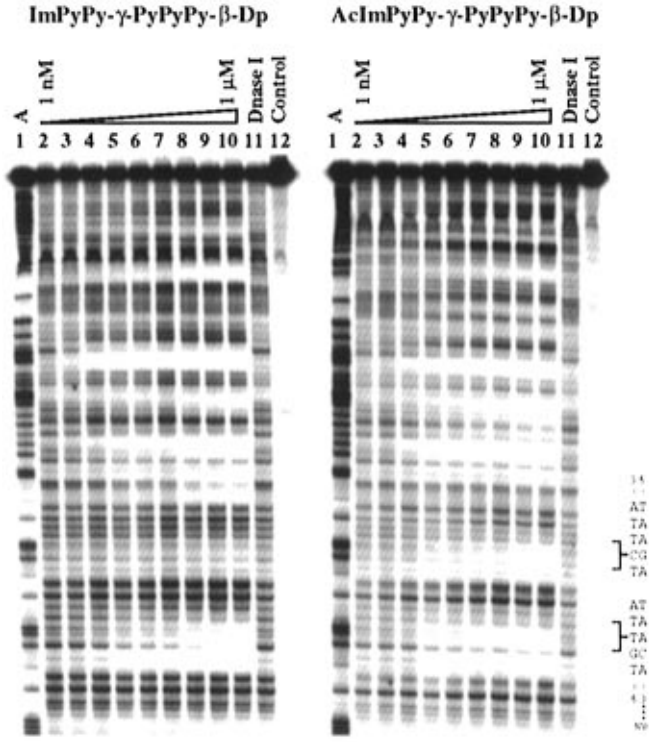

Figure 10. Quantitative DNase I footprint titration experiment with ImPyPy-γ-PyPyPy-β-Dp (**1**) and AcImPyPy-γ-PyPyPy-β-Dp on the 3′-^{32}P-labeled 266 bp *Eco* RI/*Pvu* II restriction fragment from plasmid pDEH2. The binding sites 5′-TGTTA-3′ and 5′-TCTTA-3′ are shown on the right side of the storage phosphor autoradiogram.[21] All reactions contain 20 kcpm restriction fragment, 10 mM tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0. Lane 1, A reaction; lane 2, G reaction; lanes 3 and 21, DNase I standard; lanes 4−20 contain 1.0 nM, 10 nM, 65 nM, 100 nM, 150 nM, 250 nM, 500 nM, 750 nM, and 1 μM polymide, respectively.

Halocarbon, phenol was from Fisher, and ninhydrin was from Pierce. All reagents were used without further purification.

Quik-Sep polypropylene disposable filters were purchased from Isolab Inc. and were used for filtration of DCU. Disposable polypropylene filters were also used for washing resin for ninhydrin and picric acid tests and for filtering predissolved amino acids into reaction vessels. A shaker for manual solid phase synthesis was obtained from St. John Associates, Inc. Screw-cap glass peptide synthesis reaction vessels (5 and 20 mL) with a #2 sintered glass frit were made as

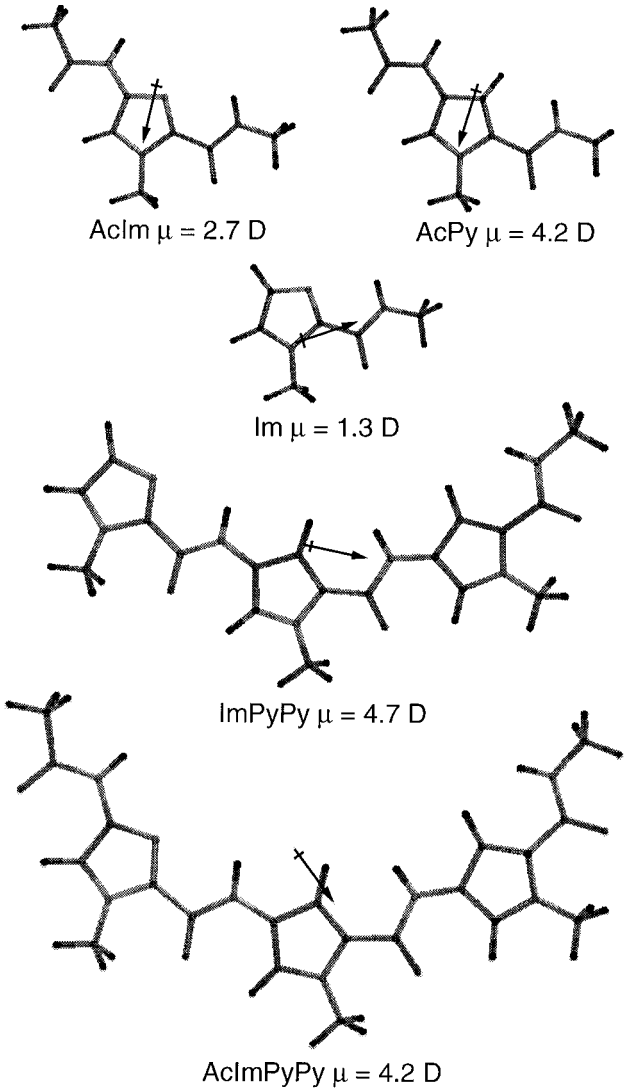

Figure 11. Dipoles calculated for individual polyamide monomers and subunits. All C-termini are the methylamide derivatives, all N-termini are either unmodified, Im and ImPyPy or acetamide derivatives, AcIm, AcPy, and AcImPyPy. Arrows depict direction of dipole moments.

described by Kent.[15] ^{1}H NMR spectra were recorded on a General Electric-QE NMR spectrometer at 300 MHz in DMSO-d_6, with chemical shifts reported in parts per million relative to residual solvent. UV spectra were measured in water on a Hewlett-Packard Model 8452A diode array spectrophotometer. Matrix-assisted, laser desorption/ionization time of flight mass spectrometry (MALDI-TOF) was performed 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 a Beckman Gold system using a RAINEN C_{18}, Microsorb MV, 5 μm, 300 × 4.6 mm reversed phase column in 0.1% (wt/v) TFA with acetonitrile as eluent and a flow rate of 1.0 mL/min, gradient elution 1.25% acetonitrile/min. Preparatory reverse phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 × 100 mm, 100 μm C18 column equipped with a guard, 0.1% (wt/v) TFA, 0.25% acetonitrile/min. Water (18 MΩ) was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered.

ImPyPy-γ-PyPyPy-β-EtOH (2). ImPyPy-γ-PyPyPy-β-Pam-resin was synthesized by machine-assisted solid phase methods. A sample of resin (240 mg, 0.18 mmol/g[16]) was placed in a glass scintillation vial and treated with neat ethanolamine (2 mL). The reaction mixture was placed in an oven and periodically agitated (55 °C, 24 h). Upon completion of polyamide cleavage, the reaction mixture was filtered to remove resin, 0.1% (wt/v) TFA added (6 mL), and the resulting solution purified by reversed phase HPLC chromatography. ImPyPy-

(15) Kent, S. B. H. *Ann. Rev. Biochem.* **1988**, *57*, 957.

γ-PyPyPy-β-EtOH is recovered upon lyophyllization as a white powder (6.8 mg, 16% recovery). UV λ_{max} 246, 306 (50 000); ^{1}H NMR (DMSO-d_6) δ 10.47 (s, 1 H), 9.90 (s, 1 H), 9.88 (s, 1 H), 9.87 (s, 1 H), 9.82 (s, 1 H), 8.03 (t, 1 H, J = 5.2 Hz), 7.95 (t, 1 H, J = 6.0 Hz), 7.85 (t, 1 H, J = 5.2 Hz), 7.38 (s, 1 H), 7.25 (d, 1 H, J = 1.4 Hz), 7.20 (d, 1 H, J = 1.3 Hz), 7.16 (d, 1 H, J = 1.4 Hz), 7.14 (m, 2 H), 7.11 (d, 1 H, J = 1.5 Hz), 7.05 (s, 1 H), 7.00 (d, 1 H, J = 1.3 Hz), 6.87 (d, 1 H, J = 1.4 Hz), 6.84 (d, 1 H, J = 1.4 Hz), 6.78 (d, 1 H, J = 1.4 Hz), 3.96 (s, 3 H), 3.81 (s, 3 H), 3.80 (s, 3 H), 3.79 (s, 3 H) 3.77 (s, 3 H), 3.76 (s, 3 H), 3.34 (m, 4 H), 3.17 (q, 2 H, J = 5.7 Hz), 3.05 (q, 2 H, J = 5.9 Hz), 2.27 (m, 4 H), 1.72 (quintet, 2 H, J = 6.9 Hz). MALDI-TOF-MS, 936.3 (937.0 calc. for M + H).

Dp-ImPyPy-γ-PyPyPy-β-Me (4). Dp-ImPyPy-γ-PyPyPy-β-Pam-resin was synthesized by machine-assisted solid phase methods. A sample of resin (240 mg, 0.18 mmol/g[16]) was placed in a glass scintillation vial and treated with a saturated solution of methylamine in DMF (20 mL). The reaction mixture was placed in sealed Parr apparatus and periodically agitated (55 °C, 24 h, 80 psi). Upon completion of cleavage, the reaction mixture was cooled to room temperature and filtered to remove resin. Excess DMF was removed *in vacuo*, 0.1% (wt/v) TFA added (6 mL), and the resulting solution purified by reversed phase HPLC chromatography. Upon lyophyllization Dp-ImPyPy-γ-PyPyPy-β-Me is recovered as a white powder (20.6 mg, 47% recovery). UV λ_{max}, 248, 312 (50 000); ^{1}H NMR (DMSO-d_6) δ 10.34 (s, 1 H), 9.53 (s, 1 H) 9.90 (s, 3 H), 9.85 (s, 1 H), 9.5 (br s, 1 H), 8.03 (m, 2 H), 7.81 (q, 1 H, J = 4.2 Hz), 7.42 (s, 1 H), 7.24 (d, 1 H, J = 1.6 Hz), 7.22 (d, 1 H, J = 1.5 Hz), 7.17 (d, 1 H, J = 1.5 Hz), 7.17 (m, 2 H), 7.12 (d, 1 H, J = 1.5 Hz), 7.00 (d, 1 H, J = 1.6 Hz), 6.88 (d, 1 H, J = 1.6 Hz), 6.85 (d, 1 H, J = 1.6 Hz), 6.80 (d, 1 H, J = 1.5 Hz), 3.93 (s, 3 H), 3.83 (s, 3 H), 3.81 (m, 6 H), 3.77 (m, 6 H), 3.34 (q, 2 H, J = 6.1 Hz), 3.18 (q, 2 H, J = 5.4 Hz), 3.03 (q, 2 H, J = 5.2 Hz), 2.76 (d, 6 H, J = 4.5 Hz), 2.54 (d, 3 H, J = 4.3 Hz), 2.37 (t, 2 H, J = 6.3 Hz), 2.26 (m, 4 H), 1.87 (quintet, 2 H, J = 6.6 Hz), 1.76 (quintet, 2 H, J = 6.8 Hz). MALDI-TOF-MS, 1034.4 (1035.2 calc. for M + H).

ImPyPy-γ-PyPyPy-β-Dp-NH$_2$ (1-NH$_2$). A sample of ImPyPy-γ-PyPyPy-β-Pam-resin (350 mg, 0.18 mmol/g[16]) was placed in a glass scintillation vial and treated with neat 3,3′-diamino-*N*-methyldipropylamine (2 mL). The reaction mixture was placed in an oven and periodically agitated (55 °C, 24 h). Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with 0.1% (wt/v) TFA to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImPyPy-γ-PyPyPy-β-Dp-NH$_2$ (28 mg, 40% recovery) as a white powder. ^{1}H NMR (DMSO-d_6) δ 10.47 (s, 1 H), 9.91 (s, 1 H), 9.90 (s, 1 H), 9.88 (s, 1 H), 9.84 (s, 1 H), 9.2 (br s, 1 H), 8.0 (m, 3 H), 7.8 (br s, 3 H), 7.33 (s, 1 H), 7.26 (d, 1 H, J = 1.2 Hz), 7.20 (d, 1 H, J = 1.4 Hz), 7.14 (m, 4 H), 7.03 (m, 2 H), 6.88 (d, 1 H, J = 1.4 Hz), 6.85 (m, 2 H), 3.97 (s, 3 H), 3.82 (m, 9 H), 3.78 (m, 6 H), 3.33 (q, 2 H, J = 5.7 Hz), 3.2- 3.0 (m, 8 H), 2.81 (q, 2 H, J = 5.8 Hz), 2.71 (d, 3 H, J = 4.4 Hz), 2.39 (t, 2 H, J = 5.8 Hz), 2.22 (t, 2 H, J = 6.1 Hz), 1.85 (quintet, 2 H, J = 6.2 Hz), 1.78 (m, 4 H). MALDI-TOF-MS, 1022.1 (1021.2 calc. for M + H).

ImPyPy-γ-PyPyPy-β-C7-NH$_2$ (2-NH$_2$). A sample of ImPyPy-γ-PyPyPy-β-Pam-resin (350 mg, 0.18 mmol/g[16]) was placed in a glass scintillation vial and treated with neat 1,7-diaminoheptane (2 mL). The reaction mixture was placed in an oven and periodically agitated (55 °C, 24 h). Resin was removed by filtration through a disposable propylene filter, and the resulting solution dissolved with 0.1% (wt/v) TFA to a total volume of 8 mL and purified directly by preparatory reversed phase HPLC to provide ImPyPy-γ-PyPyPy-β-Dp-NH$_2$ (28 mg, 39% recovery) as a white powder. ^{1}H NMR (DMSO-d_6) δ 10.62 (s, 1 H), 9.95 (s, 1 H), 9.92 (s, 1 H), 9.91 (s, 1 H), 9.87 (s, 1 H), 8.09 (t, 1 H, J = 5.3 Hz), 8.05 (t, 1 H, J = 5.6 Hz), 7.90 (t, 1 H, J = 5.7 Hz), 7.6 (br s, 3 H), 7.46 (s, 1 H), 7.29 (d, 1 H, J = 1.7 Hz), 7.22 (d, 1 H, J = 1.5 Hz), 7.16 (m, 4 H), 7.13 (d, 1 H, J = 1.3 Hz), 7.02 (d, 1 H, J = 1.6 Hz), 6.88 (d, 1 H, J = 1.6 Hz), 6.86 (d, 1 H, J = 1.6 Hz), 6.82 (d, 1 H, J = 1.6 Hz), 3.98 (s, 3 H), 3.83 (s, 3 H), 3.82 (s, 3 H), 3.81 (s, 3 H), 3.79 (s, 3 H), 3.78 (s, 3 H), 3.34 (q, 2 H, J = 5.9 Hz), 3.18 (q, 2 H, J = 6.2 Hz), 3.03 (q, 2 H, J = 5.9 Hz), 2.73 (m, 2 H), 2.27 (m, 4 H), 1.77 (t, 2 H, J = 5.8 Hz), 1.46 (m, 2 H), 1.35 (m, 2 H), 1.23 (m, 6 H). MALDI-TOF-MS, 1006.6 (1006.2 calc. for M + H).

AcImPyPy-γ-PyPyPy-β-Dp-NH$_2$ (3-NH$_2$). AcImPyPy-γ-PyPyPy-β-resin was synthesized by machine assisted methods. A sample of resin (350 mg, 0.18 mmol/g[17]) was placed in a 20 mL glass scintillation vial and treated with 2 mL of 3,3′-diamino-*N*-methyldipropylamine (55 °C, 18 h). Resin was removed by filtration through a propylene filter, and the resulting solution diluted with 0.1% (wt/v) TFA to a total volume of 8 mL and purified directly by reversed phase HPLC to provide AcImPyPy-γ-PyPyPy-β-Dp-NH$_2$ (29 mg, 43% recovery) as a white powder. ^{1}H NMR (DMSO-d_6) δ 10.26 (s, 1 H), 10.17 (s, 1 H), 9.92 (m, 2 H), 9.90 (s, 1 H), 9.87 (s, 1 H), 9.5 (br s, 1 H), 8.12 (t, 1 H, J = 5.6 Hz), 8.09 (m, 2 H), 7.9 (br s, 3 H), 7.41 (s, 1 H), 7.25 (d, 1 H, J = 1.5 Hz), 7.21 (d, 1 H, J = 1.3 Hz), 7.16 (m, 3 H), 7.11 (d, 1 H, J = 1.6 Hz), 7.03 (d, 1 H, J = 1.3 Hz), 6.89 (d, 1 H, J = 1.6 Hz), 6.85 (m, 2 H), 3.93 (s, 3 H), 3.83 (s, 3 H), 3.82 (s, 3 H), 3.81 (s, 3 H), 3.79 (s, 3 H), 3.78 (s, 3 H), 3.37 (q, 2 H, J = 5.9 Hz), 3.2- 3.0 (m, 8 H), 2.84 (q, 2 H, J = 5.7 Hz), 2.71 (d, 3 H, J = 4.3 Hz), 2.33 (t, 2 H, J = 6.6 Hz), 2.26 (t, 2 H, J = 6.8 Hz), 2.00 (s, 3 H), 1.88 (quintet, 2 H, J = 6.8 Hz), 1.77 (m, 4 H). MALDI-TOF-MS, 1078.0 (1078.2 calc. for M + H).

Dp-ImPyPy-γ-PyPyPy-β-C7-NH$_2$ (4-NH$_2$). A sample of Dp-ImPyPy-γ-PyPyPy-β-resin (350 mg, 0.16 mmol/g[17]) was placed in a 20 mL glass scintillation vial and treated with 2 mL of 1,7-diaminoheptane (55 °C, 18 h). Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with 0.1% (wt/v) TFA to a total volume of 8 mL and purified directly by reversed phase HPLC to provide Dp-ImPyPy-γ-PyPyPy-β-C7-NH$_2$ (37 mg, 55% recovery) as a white powder. ^{1}H NMR (DMSO-d_6) δ 10.40 (s, 1 H), 9.98 (s, 1 H), 9.92 (d, 2 H), 9.91 (s, 1 H), 9.87 (s, 1 H), 9.6 (br s, 1 H), 8.09 (m, 1 H), 8.05 (t, 1 H, J = 5.1 Hz), 7.88 (t, 1 H, J = 5.4 Hz), 7.6 (br s, 3 H), 7.43 (s, 1 H), 7.25 (d, 1 H, J = 1.6 Hz), 7.21 (d, 1 H, J = 1.5 Hz), 7.17 (m, 3 H), 7.13 (d, 1 H, J = 1.6 Hz), 7.02 (d, 1 H, J = 1.6 Hz), 6.89 (d, 1 H, J = 1.7 Hz), 6.86 (d, 1 H, J = 1.6 Hz), 6.83 (d, 1 H, J = 1.6 Hz), 3.93 (s, 3 H), 3.83 (s, 3 H), 3.82 (s, 3 H), 3.81 (s, 3 H), 3.80 (s, 3 H), 3.78 (s, 3 H), 3.32 (q, 2 H, J = 5.7 Hz), 3.21 (q, 2 H, J = 6.2 Hz), 3.05 (m, 4 H), 2.76 (d, 6 H, J = 3.9 Hz), 2.72 (m, 2 H), 2.39 (t, 2 H, J = 6.1 Hz), 2.28 (m, 4 H), 1.91 (quintet, 2 H, J = 6.6 Hz), 1.77 (quintet, 2 H, J = 6.4 Hz), 1.47 (m, 2 H), 1.36 (m, 2 H), 1.24 (m, 6 H). MALDI-TOF-MS, 1133.8 (1134.3 calc. for M + H).

ImPyPy-γ-PyPyPy-β-Dp-EDTA (1-E). EDTA-dianhydride (50 mg) was dissolved in 1 mL of DMSO/NMP solution and 1 mL of DIEA by heating at 55 °C for 5 min. The dianhydride solution was added to ImPyPy-γ-PyPyPy-β-Dp-NH$_2$ (**1-NH$_2$**) (8.0 mg, 7 μmol) in 750 μL of DMSO. The mixture was heated at 55 °C for 25 min, treated with 3 mL of 0.1 M NaOH, and heated at 55 °C for 10 min. Aqueous TFA (0.1%) was added to adjust the total volume to 8 mL, and the solution purified directly by reversed phase HPLC chromatography to provide **1-E** as a white powder (4 mg, 40% recovery). ^{1}H NMR (DMSO-d_6) δ 10.45 (s, 1 H), 9.89 (s, 1 H), 9.88 (s, 1 H), 9.86 (s, 1 H), 9.82 (s, 1 H), 9.2 (br s, 1 H), 8.39 (t, 2 H, J = 6.0 Hz), 8.06 (m, 3 H), 7.37 (s, 1 H), 7.25 (d, 1 H, J = 1.5 Hz), 7.17 (d, 1 H, J = 1.3 Hz), 7.13 (m, 4 H), 7.02 (m, 2 H), 6.87 (d, 1 H, J = 1.2 Hz), 6.85 (m, 2 H), 3.96 (s, 3 H), 3.81 (s, 3 H), 3.80 (s, 3 H), 3.79 (s, 3H), 3.77 (s, 3 H), 3.76 (s, 3 H), 3.68 (m, 4 H), 3.33 (q, 2 H, J = 5.3 Hz), 3.3−3.0 (m, 16 H), 2.85 (m, 2 H), 2.69 (d, 3 H, J = 4.3 Hz), 2.33 (t, 2 H, J = 5.5 Hz), 2.24 (t, 2 H, J = 6.6 Hz), 1.73 (m, 6 H). MALDI-TOF-MS, 1295.4 (1295.3 calc. for M + H).

ImPyPy-γ-PyPyPy-β-C7-EDTA (2-E). Compound **2-E** was prepared from compound **2-NH$_2$** (8.0 mg, 7 μmol) as described for **1-E** (2 mg, 20% recovery). ^{1}H NMR (DMSO-d_6) δ 10.51 (s, 1 H), 9.94 (s, 1 H), 9.92 (s, 1 H), 9.91 (s, 1 H), 9.86 (s, 1 H), 8.33 (t, 1 H, J = 5.5 Hz), 8.09 (t, 1 H, J = 5.8 Hz), 8.04 (t, 1 H, J = 5.0 Hz), 7.87 (t, 1 H, J = 5.2 Hz), 7.40 (s, 1 H), 7.27 (d, 1 H, J = 1.6 Hz), 7.21 (d, 1 H, J = 1.5 Hz), 7.17 (m, 3 H), 7.14 (d, 1 H, J = 1.7 Hz), 7.05 (s, 1 H), 7.02 (d, 1 H, J = 1.5 Hz), 6.89 (d, 1 H, J = 1.6 Hz), 6.86 (d, 1 H, J = 1.5 Hz), 6.82 (d, 1 H, J = 1.5 Hz), 3.97 (s, 3 H), 3.83 (s, 3 H), 3.82 (s, 3 H), 3.81 (s, 3 H), 3.79 (s, 3 H), 3.78 (s, 3 H), 3.67 (m, 4 H), 3.33 (q, 2 H, J = 5.6 Hz), 3.22 (m, 4 H), 3.10 (m, 4 H), 3.04 (m, 4 H), 2.92

(16) Resin substitution has been corrected for the weight of the polyamide chain. The change in resin substitution can be calculated as L_{new}(mmol/g) = $L_{old}/(1 + L_{old}(W_{new} - W_{old}) \times 10^{-3})$, where L is the loading (mmol of amine per 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. Peptide Protein Res.* **1991**, *37*, 513.

(m, 4 H), 1.77 (t, 2 H, J = 6.4 Hz), 1.35 (m, 4 H), 1.22 (m, 6 H). MALDI-TOF-MS, 1280.4 (1280.3 calc. for M + H).

AcImPyPy-γ-PyPyPy-β-Dp-EDTA (3-E). Compound **3-E** was prepared from compound **3-NH$_2$** (8.0 mg, 7 μmol) as described for **1-E** (3 mg, 30%). ^{1}H NMR (DMSO-d_6) δ 10.25 (s, 1 H), 10.00 (s, 1 H), 9.91 (m, 2 H), 9.90 (s, 1 H), 9.85 (s, 1 H), 8.42 (t, 1 H, J = 5.6 Hz), 8.08 (m, 3 H), 7.41 (s, 1 H), 7.26 (d, 1 H, J = 1.5 Hz), 7.20 (d, 1 H, J = 1.5 Hz), 7.16 (m, 3 H), 7.14 (d, 1 H, J = 1.5 Hz), 7.04 (d, 1 H, J = 1.2 Hz), 6.88 (d, 1 H, J = 1.6 Hz), 6.86 (m, 2 H), 3.93 (s, 3 H), 3.83 (s, 3 H), 3.82 (s, 3 H), 3.81 (s, 3 H), 3.79 (s, 3 H), 3.78 (s, 3 H), 3.71 (m, 4 H), 3.58 (q, 2 H, J = 5.2 Hz), 3.4- 3.0 (m, 18 H), 2.71 (d, 3 H, J = 4.3 Hz), 2.33 (t, 2 H, J = 5.4 Hz), 2.26 (t, 2 H, J = 5.7 Hz), 2.00 (s, 3 H), 1.74 (m, 6 H). MALDI-TOF-MS, 1352.8 (1352.4 calc. for M + H).

Dp-ImPyPy-γ-PyPyPy-β-C7-EDTA (4-E). Compound **4-E** was prepared from compound **4-NH$_2$** (8.0 mg, 7 μmol) as described for **1-E** (4 mg, 40% recovery). ^{1}H NMR (DMSO-d_6) δ 10.36 (s, 1 H), 9.92 (s, 1 H), 9.88 (m, 2 H), 9.87 (s, 1 H), 9.82 (s, 1 H), 9.2 (br s, 1 H), 8.03 (m, 3 H), 7.98 (m, 1 H), 7.41 (s, 1 H), 7.22 (d, 1 H, J = 1.5 Hz), 7.20 (d, 1 H, J = 1.6 Hz), 7.14 (m, 3 H), 7.10 (d, 1 H, J = 1.6 Hz), 7.00 (d, 1 H, J = 1.5 Hz), 6.87 (d, 1 H, J = 1.3 Hz), 6.83 (d, 1 H, J = 1.4 Hz), 6.78 (d, 1 H, J = 1.6 Hz), 3.92 (s, 3 H), 3.82 (s, 3 H), 3.80 (s, 3 H), 3.79 (s, 3 H), 3.77 (s, 3 H), 3.76 (s, 3 H), 3.41 (m, 4 H), 3.3−3.0 (m, 16 H), 2.74 (d, 6 H, J = 4.3 Hz), 2.71 (m, 2 H), 2.31 (t, 2 H, J = 5.4 Hz), 2.22 (t, 2 H, J = 5.6 Hz), 1.77 (m, 2 H), 1.39 (m, 2 H), 1.28 (m, 4 H), 1.20 (m, 6 H). MALDI-TOF-MS, 1407.7 (1408.6 calc. for M + H).

DNA Reagents and Materials. Sonicated, deproteinized calf thymus DNA, from Pharmacia, was dissolved in filter sterilized water to a final concentration of 1 mM in base-pairs and stored at 4 °C. Glycogen was purchased from Boehringer-Manheim as a 20 mg/mL aqueous solution. Nucleotide triphosphates were purchased from Pharmacia and used as supplied. Nucleoside triphosphates labeled with ^{32}P (≥3000 C$_i$/mmol) were obtained from Dupont-New England Nuclear. Cerenkov radioactivity was measured with a Beckman LS 2801 scintillation counter. All enzymes were purchased from Boehringer-Manheim and used according to the supplier's recommended protocol in the activity buffer provided. Plasmid pUC19 was obtained from Worthington Biochemical. A solution of 0.5 M EDTA, pH 8.0, was purchased from Ultrapure. Phosphoramidites were from Glen Research. The pH of buffer solutions was recorded using a digital pH/mV meter (Model no. 611, Orion Research) and a ROSS semimicro combination pH electrode. General manipulation of duplex DNA and oligonucleotides were performed according to established procedures.[17,18]

Construction of Plasmid DNA. Oligodeoxynucleotides were synthesized by standard automated solid support chemistry using an Applied Biosystems model 380B DNA synthesizer and *O*-cyanoethyl-*N*,*N*-diisopropyl phosphoramidites. Plasmid pDEH2 was prepared by hybridization of a complementary set of synthetic oligonucleotides: (1) 5′-GATCCGCATGC GATCGTGTTACGCATGCGATCGTCTTACG-CATGCGATCG-3′ and (2) 5′-AGCT CGATCGCATGCGTAAGAC-GATCGCATGCGTAACACGATCGCATGCG-3′. The complementary oligonucleotides were annealed and then ligated to the large pUC19 *Bam* HI/*Hind* III restriction fragment using T4 DNA ligase. The ligated plasmid was then used to transform Epicurean Coli XL-1 Blue Supercompetent cells. Colonies were selected for α-complementation on 25 mL Luria-Bertani medium agar plates containing 50 mg/mL ampicillin and treated with XGAL and IPTG solutions. Large scale plasmid purification was performed using Qiagen purification kits. The presence of the desired insert was determined by dideoxy sequencing using a USB Sequenase v. 2.0 kit. Plasmid DNA concentration was determined at 260 nm using the relation 1 OD unit = 50 μg/mL duplex DNA.

Preparation of 3′- and 5′-End-Labeled Restriction Fragments. The plasmid pDEH2 was linearized with *Eco* RI and then treated with Sequenase v. 2.0, deoxyadenosine 5′-[α-^{32}P]triphosphate, and thymidine 5′-[α-^{32}P]triphosphate. The 3′-end labeled fragment was then digested with *Pvu* II and loaded onto a 7% nondenaturing polyacrylamide gel. The desired 266 base-pair band was visualized by autoradiography and isolated. For MPE and affinity cleaving reactions, pDEH2 also was 5′-^{32}P labeled. First, the plasmid pDEH2 was digested with *Eco* RI and then dephosphorylated with calf intestine alkaline phosphatase. The digested plasmid then was 5′-^{32}P labeled with T4 polynucleotide kinase and deoxyadenosine 5′-[γ-^{32}P]triphosphate, digested with *Pvu* II, and loaded onto a 7% nondenaturing polyacrylamide gel. The desired 266 base-pair band was visualized by autoradiography and isolated. Chemical sequencing adenine-specific reactions were performed on all labeled fragments.[19]

MPE·Fe(II) Footprint Titrations.[8] All reactions were performed in a total volume of 40 μL. A polyamide stock solution (ImPyPy-γ-PyPyPy-β-Dp, ImPyPy-γ-PyPyPy-β-EtOH, AcImPyPy-γ-PyPyPy-β-Dp, or Dp-ImPyPy-γ-PyPyPy-β-Me) or H_2O (for reference lanes) was added to an assay buffer containing either 3′-^{32}P labeled or 5′-^{32}P labeled restriction fragment (20 000 cpm), affording final solution conditions of 25 mM tris-acetate, 10 mM NaCl, 100 μM/bp calf thymus DNA, pH 7, and either (i) polyamide stock solution or (ii) no polyamide (for reference lanes). Solutions were incubated at 22 °C for 24 h. A fresh 50 μM MPE·Fe(II) solution was made from 100 μL of a 100 μM MPE solution and 100 μM ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2\cdot 6H_2O$) solution. Then, 4 μL of a 50 μM MPE·Fe(II) solution was added, and the solution was allowed to equilibrate for 10 min at 22 °C. Cleavage was initiated by the addition of 4 μL of a 50 mM dithiothreitol solution and allowed to proceed for 15 min at 22 °C. Reactions were stopped by ethanol precipitation, resuspended in 1 × TBE/80% formamide loading buffer, denatured by heating at 85 °C for 15 min, and placed on ice. Reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross-linking, 7 M urea) in 1 × TBE at 2000 V for 2.5 h. Gels were dried on a slab dryer and then exposed to a storage phophor screen at 22 °C. The data were analyzed by performing volume integrations of the cleavage bands and reference bands using the ImageQuant v. 3.3 software. Background-corrected volume integration of rectangles encompassing cleavage bands was normalized to a maximum value of 0.95.

Affinity Cleaving Titrations. All affinity cleavage reactions[7] were performed in a total volume of 40 μL. A polyamide stock solution (ImPyPy-γ-PyPyPy-β-Dp-EDTA, ImPyPy-γ-PyPyPy-β-C7-EDTA, AcImPyPy-γ-PyPyPy-β-Dp-EDTA, or Dp-ImPyPy-γ-PyPyPy-β-C7-EDTA) or H_2O (for reference lanes) was added to an assay buffer containing either 3′-^{32}P labeled or 5′-^{32}P labeled restriction fragment (20 000 cpm), affording final solution conditions of 25 mM tris-acetate, 20 mM NaCl, 100 μM calf thymus DNA, pH 7, and either (i) 100 nM-5 μM polyamide or (ii) no polyamide (for reference lanes). Solutions were incubated at 22 °C for 24 h. Then, 4 μL of a 100 μM ferrous ammonium sulfate solution was added, and the solution was allowed to equilibrate for 20 min at 22 °C. Affinity cleaving reactions were initiated by the addition of 4 μL of a 100 mM dithiothreitol solution and reacted for 30 min at 22 °C. Reactions were stopped by the addition of 10 μL of a solution containing 1.5 M NaOAc (pH 5.5), 0.28 mg/mL glycogen, and 14 μM base pairs calf thymus DNA, and ethanol precipitated. Reactions were resuspended in 1 × TBE/80% formamide loading buffer, denatured by heating at 85 °C for 15 min, and placed on ice. Reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross-linking, 7 M urea) in 1 × TBE at 2000 V for 2.5 h. Gels were dried on a slab dryer and then exposed to a storage phophor screen at 22 °C. The data were analyzed by performing volume integrations of the cleavage bands and reference bands using the ImageQuant v. 3.3 software. Background-corrected volume integration of rectangles encompassing cleavage bands was normalized to a maximum value of 3.95.

DNase I Footprinting. All reactions[9] were carried out in a volume of 40 μL. We note that no carrier DNA was used in these reactions. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were 10 mM tris·HCl buffer (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, and 20 kcpm 3′-radiolabeled DNA. The solutions were allowed to equilibrate for 24 h at 22 °C. Cleavage was initiated by the addition of 4 μL of a DNase I stock solution (diluted with 1 mM DTT to give a stock concentration of 0.1 u/mL) and allowed to proceed for 7 min

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

(18) Gait, M. J. *Oligonucleotide Synthesis: A Practical Approach*; IRL Press: Oxford, 1984.

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

at 22 °C. The reactions were stopped by the addition of 1.25 M sodium chloride solution containing 100 mM EDTA, 0.2 mg/mL glycogen, and 28 μM bp calf thymus DNA, and then ethanol precipitated. The cleavage products were resuspended in 100 mM tris-borate−EDTA/80% formamide loading buffer, denatured at 85 °C for 15 min, placed on ice, and immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V for 2.5 h. The gels were dried under vacuum at 80 °C and then quantitated using storage phosphor technology.[20]

The data were analyzed by performing volume integrations of the 5′-TGTTA-3′ and 5′-TCTTA-3′ sites and a reference site. The apparent DNA target site saturation, θ_{app}, was calculated for each concentration of polyamide using the following equation

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

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 fit to a Langmuir binding isotherm (eq 2, $n = 1$) by minimizing the difference between θ_{app} and θ_{fit}, using the modified Hill equation

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

where $[L]_{tot}$ corresponds to the total polyamide concentration, K_a corresponds to the monomeric association constant, and θ_{min} and θ_{max} represent the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. Data were fit using a nonlinear least-squares fitting procedure of KaleidaGraph software (version 2.1, Abelbeck software) with K_a, $\theta_{max,}$ and θ_{min} as the adjustable parameters. All acceptable fits had a correlation coefficient of $R > 0.97$. 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. Photostimuable storage phosphor imaging plates (Kodak Storage Phosphor Screen SO230 obtained from Molecular Dynamics) were pressed flat against dried gel samples and exposed in the dark at 22 °C for 12−24 h. A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens.[20] The data were analyzed by performing volume integration of the target site and reference blocks using the ImageQuant v. 3.3 software.

Acknowledgment. We are grateful to the National Institutes of Health (GM-2768) for research support, the National Science Foundation for a predoctoral fellowship to S.W., and the Howard Hughes Medical Institute for a predoctoral fellowship to E.E.B. We thank G. M. Hathaway for MALDI-TOF mass spectrometry.

JA971569B

(20) Johnston, R. F.; Picket, S. C.; Barker, D. L. *Electrophoresis* **1990**, *11*, 355.

(21) We note that residual salt is most likely responsible for the discontinuity present at the top of gels and does not interfere with analysis of the target sites.