Stereochemical Control of the DNA Binding Affinity, Sequence Specificity, and Orientation Preference of Chiral Hairpin Polyamides in the Minor Groove

David M. Herman, Eldon E. Baird, and Peter B. Dervan*

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

Received October 27, 1997

Abstract: Three-ring polyamides containing pyrrole (Py) and imidazole (Im) amino acids covalently coupled by γ-aminobutyric acid (γ) form six-ring hairpins that recognize five-base-pair sequences in the minor groove of DNA. Selective chiral substitution of the "γ-turn" enhances the properties of polyamide hairpins with regard to DNA affinity and sequence specificity. Polyamides of core sequence composition ImPyPy-γ-PyPyPy which differ by selective stereochemical substitution of the prochiral α-position in the γ-turn were prepared. The DNA binding properties of two enantiomeric polyamides were analyzed by footprinting and affinity cleavage on a DNA fragment containing two match sites (5′-TGTTA-3′ and 5′-ACATT-3′) and one 5′-TGTCA-3′ mismatch site. Quantitative footprint titrations demonstrate that replacement of γ-aminobutyric acid by (*R*)-2,4-diaminobutyric acid enhances DNA binding affinity for the 5′-TGTTA-3′ match site 13-fold ($K_a = 3.8 \times 10^9$ M^{-1}). The enhanced affinity is achieved without a compromise in sequence selectivity, which in fact increases and is found to be 100-fold higher relative to binding at a single base pair mismatch sequence, 5′-TGTCA-3′. An (*S*)-2,4-diaminobutyric acid linked hairpin binds with 170-fold reduced affinity relative to the *R*-enantiomer and only 5-fold sequence specificity versus a 5′-ACATT-3′ reversed orientation site. These effects are modulated by acetylation of the chiral amine substituents. This study identifies structural elements which should facilitate the design of new hairpin polyamides with improved DNA binding affinity, sequence specificity, and orientational selectivity.

Introduction

Small molecules that target specific predetermined DNA sequences have the potential to control gene expression. Polyamides containing *N*-methylpyrrole and *N*-methylimidazole amino acids are synthetic ligands that have an affinity and specificity for DNA comparable to naturally occurring DNA binding proteins.[1] DNA recognition depends on side-by-side amino acid pairings oriented N−C with respect to the 5′−3′ direction of the DNA helix in the minor groove.[2] Antiparallel pairing of imidazole (Im) opposite pyrrole (Py) recognizes a G·C base pair, while a Py/Im combination recognizes C·G.[2] A Py/Py pair is degenerate and recognizes either an A·T or T·A base pair.[2,3] An Im/Im pairing[4] is disfavored,[5] breaking a potential degeneracy for recognition.

In parallel with elucidation of the scope and limitations of the polyamide pairing rules, efforts have been made to prevent slipped binding motifs[6] as well as increase DNA binding affinity and sequence specificity by covalently linking polyamide subunits.[6−9] A hairpin polyamide motif with γ-aminobutyric acid (γ) serving as a turn-specific internal guide residue provides a synthetically accessible method for C−N linkage of polyamide subunits (Figure 1).[9] Head-to-tail linked polyamides bind

(1) (a) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559. (b) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 6953. (c) Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 7636.

(2) (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) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Science* **1994**, *266*, 646. (f) White, S.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 8756.

(3) (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) White, S.; Baird, E. E.; Dervan, P. B. *Biochemistry* **1996**, *35*, 12532. (d) Chen, X.; Ramakrishnan; Sundaralingham, M. *J. Mol. Biol.* **1997**, *267*, 1157.

(4) (a) Dwyer, T. J.; Geierstanger, B. H.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1992**, *114*, 5911. (b) Al-Said, N.; Lown, J. W. *Tetrahedron Lett.* **1994**, *35*, 7577. (c) Walker, W. L.; Landlaw, E. M., Dickerson, R. E.; Goodsell, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 5634. (d) Kopka, M. L.; Goodsell, D. S.; Han, G. W.; Chiu, T. K.; Lown, J. W.; Dickerson, R. E. *Structure* **1997**, *5*, 1033.

(5) (a) Geierstanger, B. H.; Dwyer, T. J.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 4474. (b) Singh, S. B.; Ajay; Wemmer, D. E.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7673. (c) White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569.

(6) (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. *Nat. Struct. Biol.* **1996**, *3*, 321. (c) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *Chem. Eur. J.* **1997**, *3*, 1608.

(7) For a review, see: Wemmer, D. E. and Dervan, P. B. *Curr. Opin. Struct. Biol.* **1997**, *7*, 355.

(8) (a) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 3663. (b) Dwyer, T. J.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 9900. (c) Chen, Y. H.; Lown, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 6995.

(9) (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) Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6153. (d) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1996**, *3*, 369. (e) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 8198. (f) Pilch, D. S.; Pokar, N. A.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8306. (i) de Claire, R. P. L.; Geierstanger B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7909.

S0002-7863(97)03722-0 CCC: $15.00
Published on Web 02/07/1998

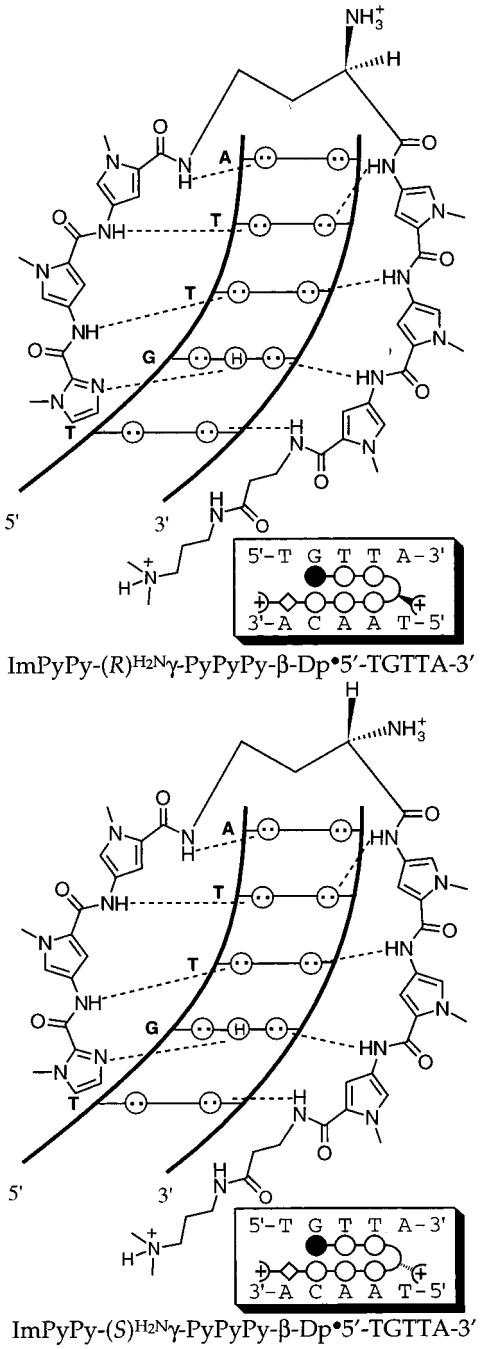


Figure 1. (top) Hydrogen bonding model of the 1:1 polyamide:DNA complex formed between the hairpin polyamide ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) with a 5′-TGTTA-3′ site. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. For schematic binding model, imidazole and pyrrole rings are represented as shaded and unshaded spheres, respectively, and the β-alanine residue is represented as an unshaded diamond. (bottom) Binding model for ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*) with a 5′-TGTTA-3′ site.

specifically to designated target sites with 100-fold enhanced affinity relative to unlinked subunits.[9] Eight-ring hairpin polyamides bearing a single positively charged tertiary amine group at the C-terminus have been shown to be cell permeable and to inhibit the transcription of specific genes in cell culture.[10] The relationship between biological regulation and the placement, frequency, and nature of charged moieties within the hairpin structure has yet to be determined. This provided impetus to elucidate the effects on hairpin polyamide DNA binding by selectively placed substituents within the γ-turn.

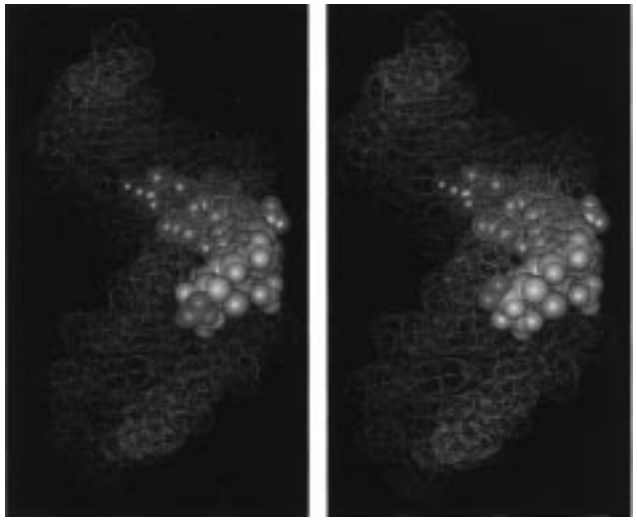


Figure 2. Computer-generated models of (left) ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp (gray) and (right) ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp (gray) bound in the minor groove of a double-stranded DNA van der Waals surface (yellow). Respective $-NH_3^+$ groups are shown in cyan. Models are derived from the NMR structure coordinates of ImPyPy-γ-PyPyPy-Dp·5′-TGTTA-3′ [9i] using InsightII software.

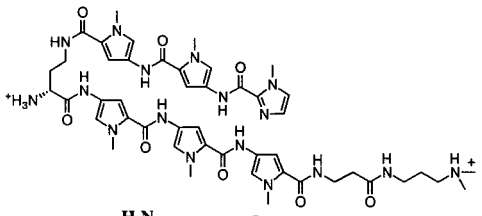
Analysis of the NMR structure[9i] of a hairpin polyamide of sequence composition ImPyPy-γ-PyPyPy complexed with a 5′-TGTTA-3′ target site suggested that substitutions at the α-position of the γ-aminobutyric acid residue could be accommodated within the hairpin−DNA complex. Modeling indicated that replacing the α-H of γ with an amino group that confers an *R*-configuration at the α-carbon could be accommodated within the floor and walls of the minor groove (Figures 1 and 2a). Formally this is accomplished by substitution of the γ-residue with (*R*)-2,4-diaminobutyric acid ((*R*)$^{H_2N}\gamma$). On the other hand, the (S)-2,4,-diaminobutyric acid ((*S*)$^{H_2N}\gamma$) linked hairpin is predicted to clash with the walls of the minor groove of the DNA helix (Figures 1 and 2b). We describe here the synthesis of a new class of chiral hairpin polyamides and their characterization with regard to DNA binding affinity and sequence specificity.

Substitution of the prochiral γ-turn with either enantiomer of 2,4-diaminobutyric acid provides the dicationic six-ring enantiomeric polyamides (+)-ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) and (−)-ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*), which were synthesized by solid-phase methods. As a control, the monocationic polyamide (+)-ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-EtOH (**2**-*R*), which lacks a charge at the C-terminus, was prepared. To further study steric effects, the α-acetamido polyamides (+)-ImPyPy-(*R*)$^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*R*) and (−)-ImPyPy-(*S*)$^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*S*) were also studied (Figure 3).[11] The EDTA analogues ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*R*·Fe(II)), ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*S*·Fe(II)), ImPyPy-(*R*)EDTA·Fe(II)γ-PyPyPy-β-Dp (**5**-*R*·Fe(II)), and ImPyPy-(*S*)EDTA·Fe(II)γ-PyPyPy-β-Dp (**5**-*S*·Fe(II)) were constructed to confirm the binding orientation of the modified hairpins at each DNA binding site (Figure 3).

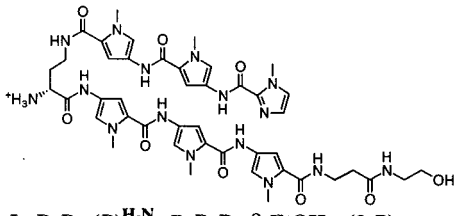
We report here the DNA binding affinity, orientation, and sequence selectivity of five six-ring hairpin polyamides **1**-*R*, **1**-*S*, **2**-*R*, **3**-*R*, and **3**-*S* for five-base-pair (5-bp) 5′-TGTTA-3′,

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

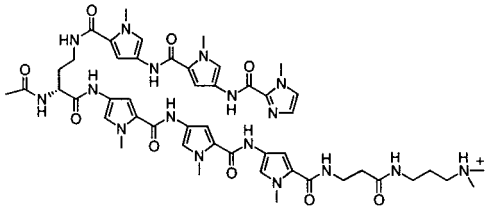
(11) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.



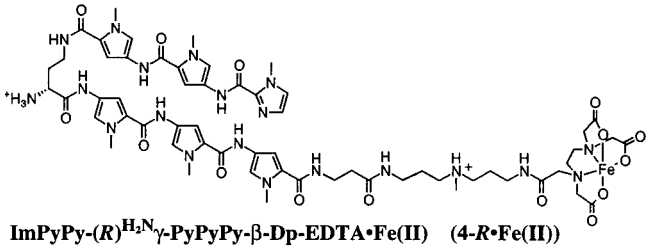
ImPyPy-(*R*)^{H_2N}γ-PyPyPy-β-Dp (1-*R*)



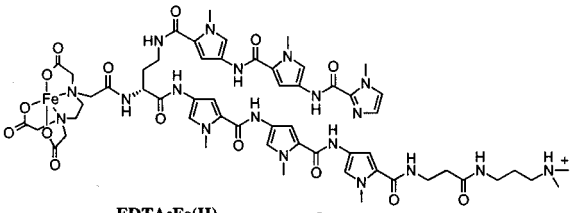
ImPyPy-(*R*)^{H_2N}γ-PyPyPy-β-EtOH (2-*R*)



ImPyPy-(*R*)^{Ac}γ-PyPyPy-β-Dp (3-*R*)



ImPyPy-(*R*)^{H_2N}γ-PyPyPy-β-Dp-EDTA•Fe(II) (4-*R*•Fe(II))



ImPyPy-(*R*)$^{EDTA•Fe(II)}$γ-PyPyPy-β-Dp (5-*R*•Fe(II))

Figure 3. Structures of the six-ring hairpin polyamides ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-Dp (**1**-*R*), ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-EtOH (**2**-*R*), ImPyPy-(*R*)^{Ac}γ-PyPyPy-*β*-Dp (**3**-*R*), ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-Dp-EDTA•Fe(II) (**4**-*R*•Fe(II)), and ImPyPy-(*R*)$^{EDTA•Fe(II)}$γ-PyPyPy-*β*-Dp (**5**-*R*•Fe(II)). Structures of the corresponding (*S*)^{H_2N}γ linked hairpins are not shown.

5′-ACATT-3′, and 5′-TGTCA-3′ sequences. Three separate techniques are used to characterize the DNA binding properties of the designed polyamides: MPE•Fe(II) footprinting,[12] affinity cleaving,[13] and DNase I footprinting.[14] Information about binding site size is gained from MPE•Fe(II) footprinting, while affinity cleavage studies determine the binding orientation and stoichiometry of the 1:1 hairpin:DNA complex. Quantitative DNase I footprint titrations allow determination of equilibrium association constants (K_a) of the polyamides for respective match and mismatch binding sites.

Results

Synthesis. Two polyamide-resins, ImPyPy-(*R*)^{Fmoc}γ-PyPyPy-*β*-Pam-resin and ImPyPy-(S)^{Fmoc}γ-PyPyPy-*β*-Pam-resin, were synthesized in 14 steps from Boc-*β*-alanine-Pam-resin (1 g of resin, 0.2 mmol/g of substitution) using previously described Boc-chemistry machine-assisted protocols (Figure 4).[11] (*R*)- and (*S*)-2,4-diaminobutyric acid residues were introduced as orthogonal protected *N*-α-Fmoc-*N*-γ-Boc derivatives (HBTU, DIEA). Fmoc-protected polyamide-resins ImPyPy-(R)^{Fmoc}γ-PyPyPy-*β*-Pam-resin and ImPyPy-(S)^{Fmoc}γ-PyPyPy-*β*-Pam-resin were treated with 1:4 DMF:piperidine (22 °C, 30 min) to provide ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-Pam-resin and ImPyPy-(*S*)^{H_2N}γ-PyPyPy-*β*-Pam-resin, respectively. A single-step aminolysis of the resin ester linkage was used to cleave the polyamide from the solid support. A sample of resin (240 mg) was treated with either (dimethylamino)propylamine (55 °C, 18 h) to provide **1**-*R*, **1**-*S*, **3**-*R*, and **3**-*S* or ethanolamine (55 °C, 18 h) to provide **2**-*R*. Resin cleavage products were purified by reverse phase HPLC to provide ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-Dp (**1**-*R*), ImPyPy-(*S*)^{H_2N}γ-PyPyPy-*β*-Dp (**1**-*S*), and ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-EtOH (**2**-*R*). The stereochemical purity of **1**-*R* was determined to be >98% by Mosher amide analysis.[15] **1**-*R*,*R* and **1**-*R*,*S* Mosher amides were prepared by reaction of **1**-*R* with HOBt activated esters generated in situ from (*R*)-α-methoxy-α-(trifluoromethyl)phenylacetic acid and (*S*)-α-methoxy-α-(trifluoromethyl)phenylacetic acid. For synthesis of analogues modified with EDTA at the carboxy terminus, the amine-resin was treated with Boc-anhydride (DMF, DIEA, 55 °C, 30 min) to provide ImPyPy-(*R*)^{Boc}γ-PyPyPy-*β*-Pam-resin and ImPyPy-(*S*)^{Boc}γ-PyPyPy-*β*-Pam-resin (Figure 4). A sample of Boc-resin was then cleaved with 3,3′-diamino-*N*-methyldipropylamine (55 °C, 18 h) and purified by reversed phase HPLC to provide either ImPyPy-(*R*)^{Boc}γ-PyPyPy-*β*-Dp-NH$_2$ (**1**-*R*-Boc-NH$_2$) or ImPyPy-(*S*)^{Boc}γ-PyPyPy-*β*-Dp-NH$_2$ (**1**-*S*-Boc-NH$_2$), which afford free primary amine groups at the C-terminus suitable for postsynthetic modification. The polyamide-amines **1**-*R*-Boc-NH$_2$ and **1**-*S*-Boc-NH$_2$ were treated with an excess of EDTA-dianhydride (DMSO/NMP, DIEA, 55 °C, 15 min), and the remaining anhydride was hydrolyzed (0.1 M NaOH, 55 °C, 10 min). The Boc-protected EDTA-modified polyamides ImPyPy-(*R*)^{Boc}γ-PyPyPy-*β*-Dp-EDTA (**4**-*R*-Boc) and ImPyPy-(*S*)^{Boc}γ-PyPyPy-*β*-Dp-EDTA (**4**-*S*-Boc) were isolated by HPLC. Individual Boc-EDTA-polyamides were deprotected with neat TFA (22 °C, 1 h) to provide the respective C-terminal EDTA derivatives, ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-Dp-EDTA (**4**-*R*) and ImPyPy-(*S*)^{H_2N}γ-PyPyPy-*β*-Dp-EDTA (**4**-*S*). For the synthesis of acetamide-turn or EDTA-turn derivatives, a sample of the α-amino polyamide ImPyPy-(*R*)^{H_2N}γ-PyPyPy-*β*-Dp (**1**-*R*) or ImPyPy-(*S*)^{H_2N}γ-PyPyPy-*β*-Dp (**1**-*S*) was treated with an excess of either acetic anhydride or EDTA-dianhydride (DMSO/NMP, DIEA 55 °C, 30 min) and the remaining anhydride was hydrolyzed (0.1 M NaOH, 55 °C, 10 min). The polyamides ImPyPy-(*R*)^{Ac}γ-PyPyPy-*β*-Dp (**3**-*R*), ImPyPy-(*S*)^{Ac}γ-PyPyPy-*β*-Dp (**3**-*S*), ImPyPy-(*R*)^{EDTA}γ−PyPyPy-*β*-Dp (**5**-*R*) and ImPyPy-(*S*)^{EDTA}γ−PyPyPy-*β*-Dp (**5**-*S*) were then isolated by reverse phase HPLC. The six-ring hairpin polyamides described here are soluble in aqueous solution at concentrations of ≤10 mM.

(12) (a) Van Dyke, M. W.; Hertzberg, R. P.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 5470. (b) Van Dyke, M. W.; Dervan, P. B. *Science* **1984**, *225*, 1122.

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

(14) (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.

(15) (a) Dale, J. A.; Mosher, H. S. *J. Am. Chem. Soc.* **1973**, *95*, 512. (b) Yamaguchi, S. In *Asymmetric Synthesis*: *Analytical Methods*; Morrison, J. D., Ed.; Academic Press: New York, 1983; Vol. 1, pp 125−152.

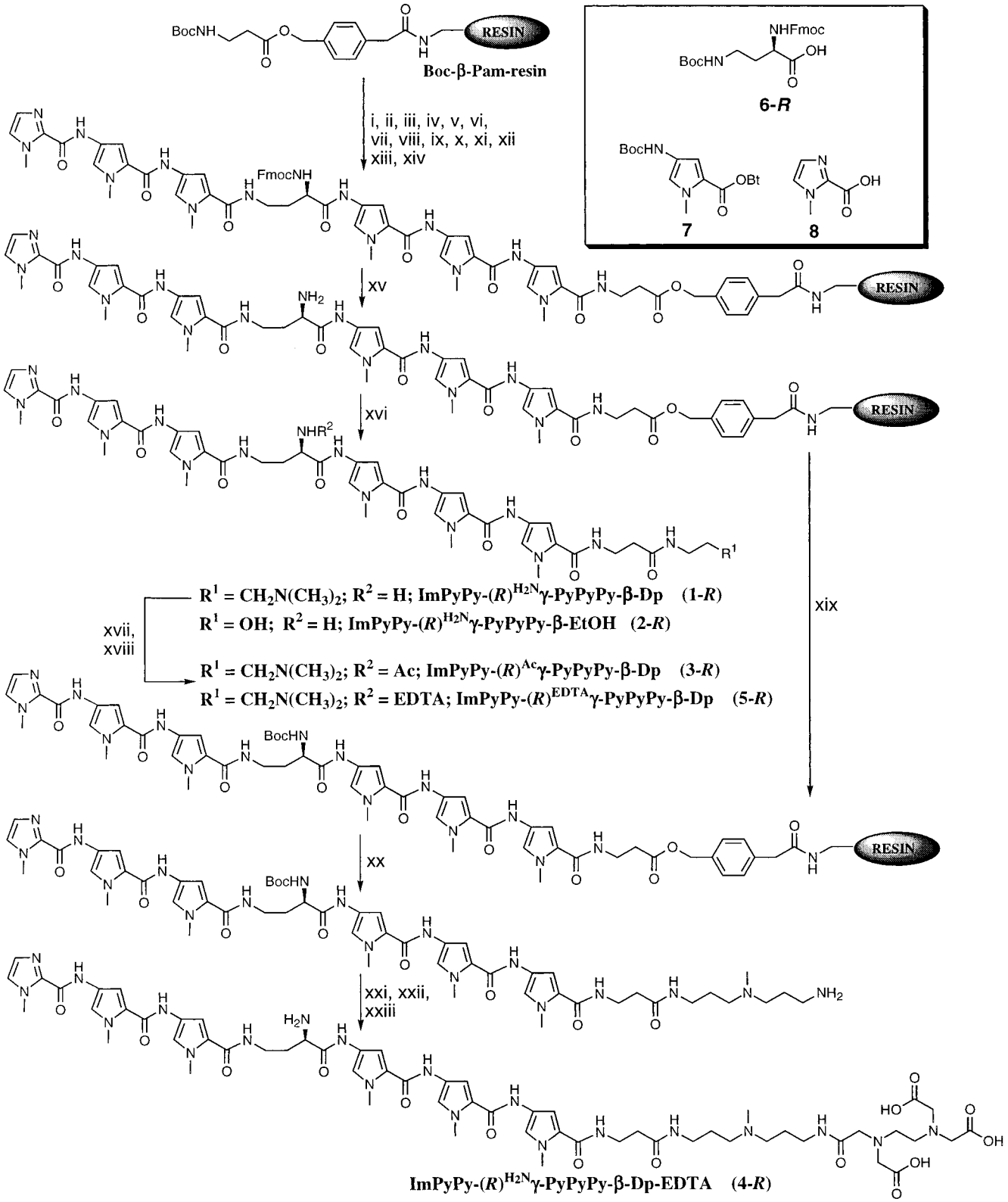


Figure 4. Solid-phase synthetic scheme exemplified for ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*), ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-EtOH (**2**-*R*), ImPyPy-$(R)^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*R*), ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA (**4**-*R*), and ImPyPy-$(R)^{EDTA}\gamma-$PyPyPy-β-Dp (**5**-*R*): (i) 80% TFA/DCM, 0.4 M PhSH; (ii) Boc-Py-OBt, DIEA, DMF; (iii) 80% TFA/DCM, 0.4 M PhSH; (iv) Boc-Py-OBt, DIEA, DMF; (v) 80% TFA/DCM, 0.4 M PhSH; (vi) Boc-Py-OBt,DIEA, DMF; (vii) 80% TFA/DCM, 0.4 M PhSH; (viii) Fmoc-α-Boc-γ-diaminobutyric 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) imidazole-2-carboxylic acid (HBTU/DIEA); (xv) 80% Piperdine:DMF (25 °C, 30 min) (xvi) (*N*,*N*-dimethylamino)propylamine (55 °C, 18 h) for **1**-*R*; ethanolamine (55 °C, 18 h) for **2**-*R*; (xvii) Ac_2O (for **3**-*R*), EDTA-dianhydride (for **5**-*R*), DIEA, DMF (55 °C, 30 min); (xviii) 0.1 NaOH (55 °C, 10 min); (xix) Boc_2O, DIEA, DMF; (xx) 3,3′-diamino-*N*-methyldipropylamine (55 °C, 18 h); (xxi) EDTA-dianhydride, DMSO, NMP, DIEA (55 °C, 30 min); (xxii) 0.1 M NaOH, (55 °C, 10 min) (xxiii) TFA: (inset) Pyrrole, imidazole, and diaminobutyric acid monomers for solid-phase synthesis: (*R*)-Fmoc-α-Boc-γ-diaminobutyric acid (**6**-*R*), Boc-pyrrole-OBt ester[11] (Boc-Py-OBt) (**7**), and imididazole-2-carboxylic acid[2a] (Im-OH) (**8**).

Figure 5. (top) Illustration of the 135-bp restriction fragment with the position of the sequence indicated. Bar heights are proportional to the relative protection from cleavage at each band. Binding sites determined by MPE·Fe(II) footprinting and quantitated by DNase I footprint titrations are boxed. (bottom) MPE·Fe(II) protection patterns of 1.25 μM ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*), 1.25 μM ImPyPy-(*R*)$^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*R*), 1.25 μM ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*), and 2.5 μM ImPyPy-(*S*)$^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*S*).

Binding Site Size and Location by MPE·Fe(II) Footprinting. MPE·Fe(II) footprinting[12] on 3′- and 5′-^{32}P end-labeled 135-bp restriction fragments reveals that the polyamides, each at 1 μM concentration, bind to the 5′-TGTTA-3′ match site (25 mM Tris-acetate, 10 mM NaCl, 100 μM/base pair of calf thymus DNA, pH 7.0 and 22 °C) (Figure 5). Compounds **1**-*R* and **3**-*R*, each at 1.25 μM, protect both the cognate 5′-TGTTA-3′ site and the single base pair mismatch sequence 5′-TGTCA-3′. Remarkably, binding sequence preferences vary for the polyamides depending on the stereochemistry of the amine substituent. At 1.25 μM and 2.5 μM concentration, respectively, polyamides **1**-*S* and **3**-*S* bind a 5′-ACATT-3′ reverse orientation match site in addition to the target match site 5′-TGTTA-3′ (Figure 5). The sizes of the asymmetrically 3′-shifted footprint cleavage protection patterns for the polyamides are consistent with 5-bp binding sites (Figure 5).

Binding Orientation by Affinity Cleaving. Affinity cleavage experiments[13] using hairpin polyamides modified with EDTA·Fe(II) at either the C-terminus or on the γ-turn were used to determine polyamide binding orientation and stoichiometry. Affinity cleavage experiments were performed on the same 3′- and 5′-^{32}P end-labeled 135-bp restriction fragment (25

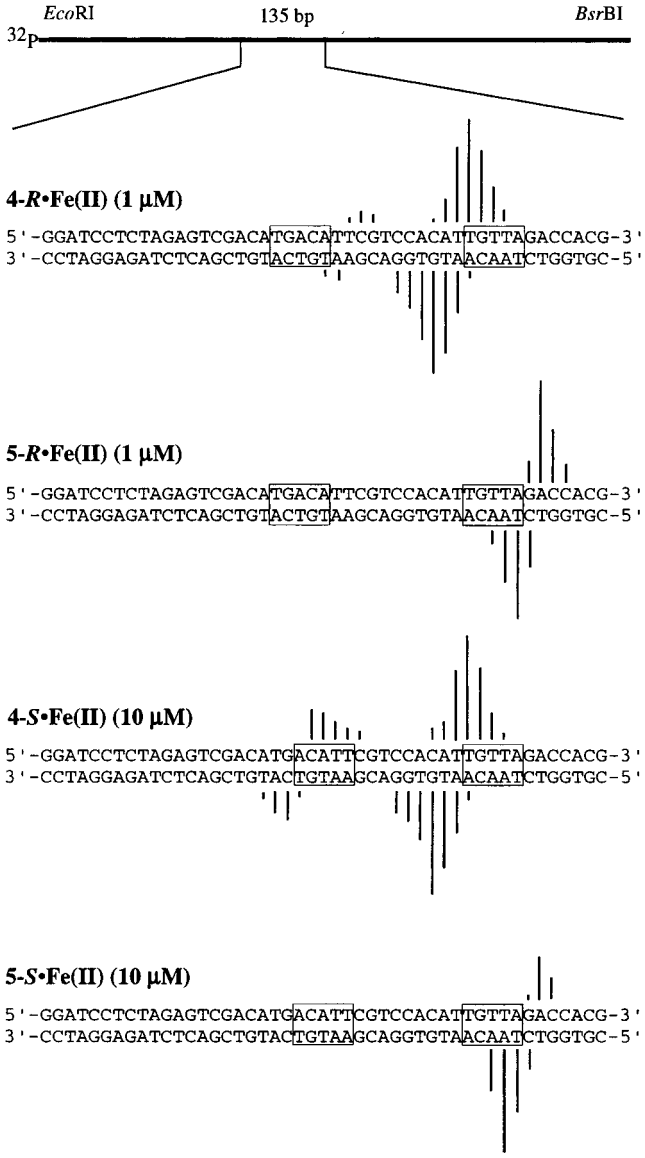


Figure 6. (top) Illustration of the 135-bp restriction fragment with the position of the sequence indicated. Line heights are proportional to the relative cleavage at each band. Binding sites determined by MPE·Fe(II) footprinting and quantitated by DNase I footprint titrations are boxed. (bottom) Affinity cleavage patterns for ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*R*·Fe(II)) and ImPyPy-(R)$^{EDTA\cdot Fe(II)}\gamma$-PyPyPy-$\beta$-Dp (**5**-*R*·Fe(II)) at 1 μM concentration and ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*S*·Fe(II)) and ImPyPy-(*S*)$^{EDTA\cdot Fe(II)}\gamma$-PyPyPy-$\beta$-Dp (**5**-*S*·Fe(II)) at 10 μM concentration.

mM Tris-acetate, 10 mM NaCl, 100 μM/base pair of calf thymus DNA, pH 7.0 and 22 °C). The observed cleavage patterns for ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*R*·Fe(II)), ImPyPy-(*R*)$^{EDTA\cdot Fe(II)}\gamma$-PyPyPy-$\beta$-Dp (**5**-*R*·Fe(II)), ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*S*·Fe(II)), ImPyPy-(*S*)$^{EDTA\cdot Fe(II)}\gamma$-PyPyPy-$\beta$-Dp (**5**-*S*·Fe(II)) (Figures 6 and 7) are in all cases 3′-shifted, consistent with minor groove occupancy. In the presence of 3.3 μM of **4**-*R*·Fe(II) and 10 μM **4**-*S*·Fe(II), which have an EDTA·Fe(II) moiety at the C-terminus, a single cleavage locus proximal to the 5′-side of the 5′-TGTTA-3′ match sequence is revealed (Figure 6). In the presence of 3.3 μM **5**-*R*·Fe(II) and 10 μM **5**-*S*·Fe(II), which have an EDTA·Fe(II) moiety appended to the γ-turn, a single cleavage locus is revealed proximal to the 3′-side of the 5′-TGTTA-3′ match sequence (Figure 6). Cleavage loci are more concise for the γ-turn EDTA·Fe(II) placement relative to carboxy-terminal

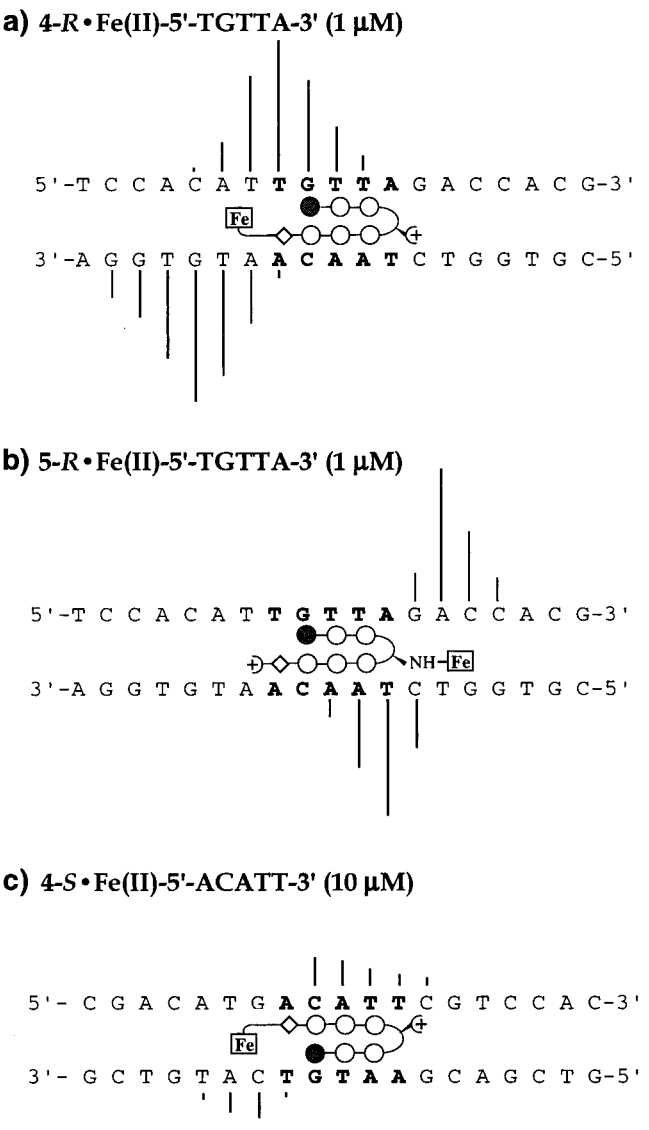


Figure 7. Affinity cleavage patterns and ball-and-stick models of the six-ring EDTA·Fe(II) analogues. Bar heights are proportional to the relative cleavage intensities at each base pair. Shaded and nonshaded circles denote imidazole and pyrrole carboxamides, respectively. Nonshaded diamonds represent β-alanine. The boxed Fe denotes the EDTA·Fe(II) cleavage moiety: (a) ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*R*·Fe(II)·5′-TGTTA-3′) at 1 μM concentration; (b) 1 μM ImPyPy-$(R)^{EDTA\cdot Fe(II)}\gamma$-PyPyPy-$\beta$-Dp (**5**-*R*·Fe(II)·5′-TGTTA-3′); (c) 10 μM ImPyPy-$(S)^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA·Fe(II) (**4**-*S*·Fe(II)·5′-ACATT-3′).

placement, consistent with the shorter tether (Figure 7). Cleavage loci are observed at both the 5′- and 3′-sides of the 5′-TGTCA-3′ single base pair mismatch site in the presence of 10 μM of **4**-*R*·Fe(II) (Figure 6). The cleavage pattern observed at the 3′-side of the site is approximately 3-fold more intense than cleavage at the 5′-side. For polyamide **4**-*S*·Fe(II) at 10 μM concentration, a single cleavage locus is revealed proximal to the 5′-side of the 5′-ACATT-3′ reverse orientation match site (Figure 7).

Energetics by Quantitative DNase I Footprinting. Quantitative DNase I footprint titrations[14] (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 constant (K_a) of each six-ring hairpin polyamide for the three resolved sites (Figures 8 and 9). The 5′-TGTTA-3′ site is bound by polyamides in the following order: ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) ($K_a = 3.8 \times 10^9$ M^{-1}) ≈ ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-EtOH (**2**-*R*) ($K_a = 3.3 \times 10^9$ M^{-1}) > ImPyPy-$(R)^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*R*) ($K_a = 3.0 \times 10^8$ M^{-1}) ≈ ImPyPy-γ-PyPyPy-β-Dp ($K_a = 2.9 \times 10^8$ M^{-1}) > ImPyPy-$(S)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*) ($K_a = 2.2 \times 10^7$ M^{-1}) > ImPyPy-$(S)^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*S*) ($K_a < 5.0 \times 10^6$ M^{-1}). Equilibrium association constants for recognition of the 5′-TGACA-3′ single base pair mismatch site are as follows: ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) ($K_a = 3.5 \times 10^7$ M^{-1}) ≈ ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-EtOH (**2**-*R*) ($K_a = 3.1 \times 10^7$ M^{-1}) > ImPyPy-$(R)^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*R*) ($K_a < 5 \times 10^6$ M^{-1}) ≈ ImPyPy-γ-PyPyPy-β-Dp ($K_a = 4.8 \times 10^6$ M^{-1}). The polyamides ImPyPy-$(S)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*) and ImPyPy-$(S)^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*S*) recognize the 5′-ACATT-3′ reverse orientation sequence with $K_a = 4.6 \times 10^6$ M^{-1} and $K_a < 5 \times 10^6$ M^{-1}, respectively. It should be noted that a detailed comparison of the relative mismatch binding energetics cannot be made since the 5′-TGACA-3′ and 5′-ACATT-3′ binding sites overlap. The relative affinity of 5′-TGTTA-3′ match site binding varies from 100-fold to 5-fold depending on the stereochemistry of the γ-turn substitutions (Table 1).

Discussion

Binding Site Size and Orientation. MPE·Fe(II) footprinting reveals that the polyamides bind with highest affinity to the 5′-TGTTA-3′ match site, the 5′-TGACA-3′ single base pair mismatch site for polyamides **1**-*R* and **3**-*R*, and the 5′-ACATT-3′ reverse orientation match site for polyamides **1**-*S* and **3**-*S* (Figure 5). Affinity cleaving experiments using polyamides with EDTA·Fe(II) placed at either the carboxy terminus or the γ-turn confirm that polyamides derived from both (*R*)- and (*S*)-2,4-diaminobutyric acid bind to the 5′-TGTTA-3′ target site with a single orientation (Figure 6). The observation of a single cleavage locus is consistent only with an oriented 1:1 polyamide: DNA complex in the minor groove and rules out any dimeric overlapped or extended binding motifs. The hairpin binding model is further supported by the location of the cleavage locus at either the 5′- or 3′-side of the 5′-TGTTA-3′ target site corresponding to EDTA·Fe(II) placement at the polyamide carboxy-terminus or the γ-turn, respectively (Figures 6 and 7). Polyamide subunits linked by $(R)^{H_2N}\gamma$ bind the symmetric single base pair mismatch sequence 5′-TGACA-3′ in two distinct orientations. Polyamides linked with $(S)^{H_2N}\gamma$ bind to a 5′-ACATT-3′ reverse orientation match sequence as revealed by a unique cleavage locus at the 5′-side of the site.

Binding Affinity. All six polyamides bind to the 5′-TGTTA-3′ target site with binding isotherms (eq 2, $n = 1$) consistent with binding as an intramolecular hairpin (Figure 9). However, the relative match site binding affinity varies by nearly 1000-fold depending on the stereochemistry of the γ-turn and the nature of the substituents. Among the six polyamides, ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) binds to the targeted 5′-TGTTA-3′ site with the highest affinity. ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp binds with an equilibrium association constant ($K_a = 3 \times 10^9$ M^{-1}), a factor of 10 greater than that of the parent polyamide, ImPyPy-γ-PyPyPy-β-Dp, ($K_a = 3 \times 10^8$ M^{-1}) (Table 1). Replacement of the C-terminal (dimethylamino)-propylamide group of **1**-*R* with an ethoxyamide group as in ImPyPy-γ-PyPyPy-β-EtOH (**2**-*R*) results in no decrease in binding affinity ($K_a = 3 \times 10^9$ M^{-1}). Acetylation of the γ-turn amino group as in ImPyPy-$(R)^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*R*) reduces binding affinity 10-fold ($K_a = 3 \times 10^8$ M^{-1}) relative to that of **1**-*R*.

The observation that polyamides which differ only by replacement of the (dimethylamino)propylamide group **1**-*R* with

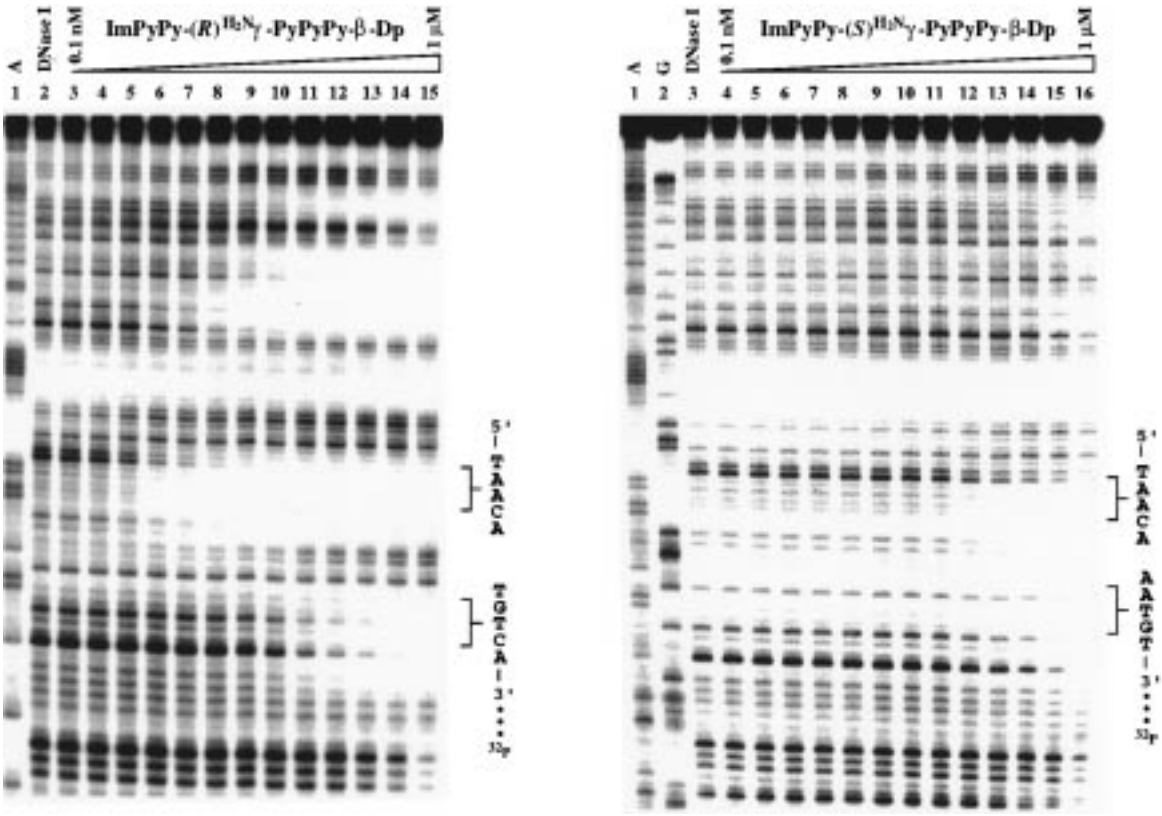


Figure 8. Quantitative DNase I footprint titration experiment with (left) ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) and (right) ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*) on the 3′-end labeled 135-bp restriction fragment: lane 1, A reaction; lane 2, DNase I standard; lanes 3−15, 0.1 nM, 0.2 nM, 0.5 nM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1.0 μM polyamide. The 5′-TGTTA-3′ and 5′-TGACA-3′ sites for **1**-*R* and the 5′-TGTTA-3′ and 5′-ACATT-3′ sites for **1**-*S* that were analyzed are shown on the right side of the autoradiogram. All reactions contain 20 kcpm restriction fragment, 10 mM Tris·HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

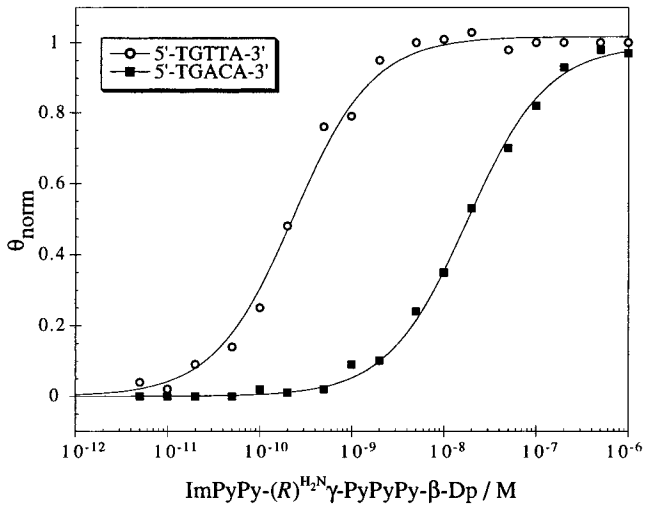


Figure 9. Data from quantitative DNase I footprint titration experiments for ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) binding to the 5′-TGTTA-3′ and 5′-TGTCA-3′ sites. θ_{norm} points were obtained using storage phosphor autoradiography and processed as described in the Experimental Section. The data for the binding of **1**-*R* to the 5′-TGTTA-3′ match site is indicated by open circles and binding to the 5′-TGACA-3′ mismatch site by closed squares. The solid curves are the best-fit Langmuir binding titration isotherms obtained from a nonlinear least-squares algorithm using eq 2 ($n = 1$).

an ethoxyamide group **2**-*R* bind with similar affinity indicates that interactions between the cationic (dimethylamino)propyl tail group with anionic phosphate residues or the negative electrostatic potential in the floor of the minor groove[16] may not contribute substantially to the energetics of hairpin−DNA binding. Furthermore, these results indicate that the observed binding enhancement of **1**-*R*, in relation to ImPyPy-γ-PyPyPy-β-Dp, is not simply the difference between a monocationic and dicationic ligand binding to the polycationic DNA helix.[16] The modest increased binding affinity of polyamide **1**-*R* may result from electrostatic interactions between the precisely placed amine group and the floor of the minor groove. Alternately, the increased affinity could indicate a reduction in the degrees of freedom accessible to the free hairpin in solution resulting from a steric effect or an electrostatic interaction between the positively charged amine group and the negative potential of the α-carbonyl group.

Polyamides linked with (*S*)$^{H_2N}\gamma$, ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*), and ImPyPy-(*S*)$^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*S*) bind to the 5′-TGTTA-3′ match site with 100-fold ($K_a = 2 \times 10^7$ M^{-1}) and 1000-fold ($K_a < 5 \times 10^6$ M^{-1}) reduced affinity relative to the (*R*)$^{H_2N}\gamma$ linked polyamide **1**-*R* (Table 1). These results demonstrate that the DNA binding affinity of chiral hairpin polyamides can be predictably regulated as a function of the stereochemistry of the turn residue.

Sequence Specificity. Polyamides with a variety of substitutions at the γ-turn bind preferentially to the 5′-TGTTA-3′ match site, while overall specificity versus binding at reverse orientation and mismatch sites is modified. Replacing the α-proton in the γ residue of ImPyPy-γ-PyPyPy-β-Dp with an amino group that confers a chiral α-hydrogen (*R*) configuration provides the most specific polyamide ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-

(16) (a) Zimmer, C.; Wahnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31. (b) Pullman, B. *Adv. Drug. Res.* **1990**, *18*, 1. (c) Breslauer, K. J.; Ferrante, R.; Marky, L. A.; Dervan, P. B.; Youngquist, R. S. In *Structure and Expression*: *DNA and Its Drug Complexes*; Sarma, R. H., Sarma, M. H. Eds; Academic Press: New York, 1988; pp 273−289.

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

polyamide	match site 5′-TGTTA−3′	reverse site 5′-ACATT−3′	mismatch site 5′-TGACA−3′	specificity[c]
ImPyPy-γ-PyPyPy-β-Dp[9b]	2.9×10^8	nd	4.8×10^6	60
ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp	3.8×10^9 (0.2)	nd	3.5×10^7 (1.0)	100
ImPyPy-$(S)^{H_2N}\gamma$-PyPyPy-β-Dp	2.2×10^7 (0.8)	4.6×10^6 (2.0)[d]	nd	5
ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-EtOH	3.3×10^9 (0.9)	nd	3.1×10^7 (0.4)	100
ImPyPy-$(R)^{Ac}\gamma$-PyPyPy-β-Dp	3.0×10^8 (1.3)	nd	$<5.0 \times 10^6$	≥ 60
ImPyPy-$(S)^{Ac}\gamma$-PyPyPy-β-Dp	$<5.0 \times 10^6$	$<5.0 \times 10^6$	nd	nd

[a] The reported association constants are the average values obtained from three DNase I footprint titration experiments. The standard deviation for each data set is indicated 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 $MgCl_2$, and 5 mM $CaCl_2$. [b] The five base-pair binding sites are in capital letters. [c] Specificity is calculated by K_a(match site)/K_a(mismatch site). [d] Mismatch site is 5′-(ACATT)-3′ for ImPyPy-$(S)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*S*) and ImPyPy-$(S)^{Ac}\gamma$-PyPyPy-β-Dp (**3**-*S*) as determined by MPE·Fe(II) footprinting and affinity cleaving.

β-Dp (**1**-*R*). The ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp·5′-TGTTA-3′ complex forms with 100-fold preference relative to the ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp·5′-TGTCA-3′ mismatch complex. Substitution of the charged (dimethylamino)propyl tail group with an ethoxyamide group as in **2**-*R* does not alter binding specificity. The modest increase in specificity against single base mismatch sequences for polyamides **1**-*R* and **2**-*R* (100-fold) relative to that for the parent unsubstituted hairpin polyamide (60-fold) implicates chiral hairpin polyamides as an optimized class of small molecules for recognition of the DNA minor groove.

Binding Orientation. In principle, a polyamide:DNA complex can form at two different DNA sequences depending on the alignment of the polyamide (N−C) with the walls of the minor groove (5′-3′).[2f] 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 individual recognized DNA strands would bind to "forward match" 5′-WGWWW-3′ sequences (W = A or T). Placement of the polyamide N-terminus at the 3′-side of each recognized strand would result in targeting "reverse match" 5′-WCWWW-3′ sequences. For hairpin polyamides, there is an energetic preference for "forward" alignment of each polyamide subunit (N−C) with respect to the backbone (5′−3′) of the DNA double helix.[2f]

In addition to decreasing the affinity for the 5′-TGTTA-3′ match site, replacing the α-proton of γ-turn in ImPyPy-γ-PyPyPy-β-Dp with $(S)^{H_2N}\gamma$ changes the mismatch sequence preference from the 5′-TGACA-3′ site bound by the $(R)^{H_2N}\gamma$-linked polyamides to a 5′-ACATT-3′ reverse match site. Binding to the reverse site may result from the presence of the steric bulk of the amino or acetamido groups in the floor of the minor groove, preventing the deep polyamide binding required for specific DNA recognition. However, an analysis of hairpin folding requirements for "forward" and "reverse" binding reveals an additional model.

In principle, there exist two non-superimposable hairpin folds which are related by mirror plane symmetry (Figure 10). One hairpin fold is responsible for the preferred 5′ to 3′ N to C orientation, while the other fold corresponds to the 3′ to 5′ N to C reverse orientation binding. For an achiral hairpin polyamide in the absence of DNA, each non-superimposable fold should be energetically equivalent. However, an asymmetrically folded hairpin polyamide with a chiral substituent could potentially display differential energetics for oriented binding (Figure 10). In the forward folded hairpin (5′ to 3′ N to C), $(R)^{H_2N}\gamma$ directs the amine functionality away from the DNA helix, while the *S*-enantiomer is predicted to direct the amine into the floor of the minor groove. For the "reverse" fold hairpin, $(S)^{H_2N}\gamma$ directs the amine functionality away from the floor of the DNA helix, while the amine of the *R*-enantiomer

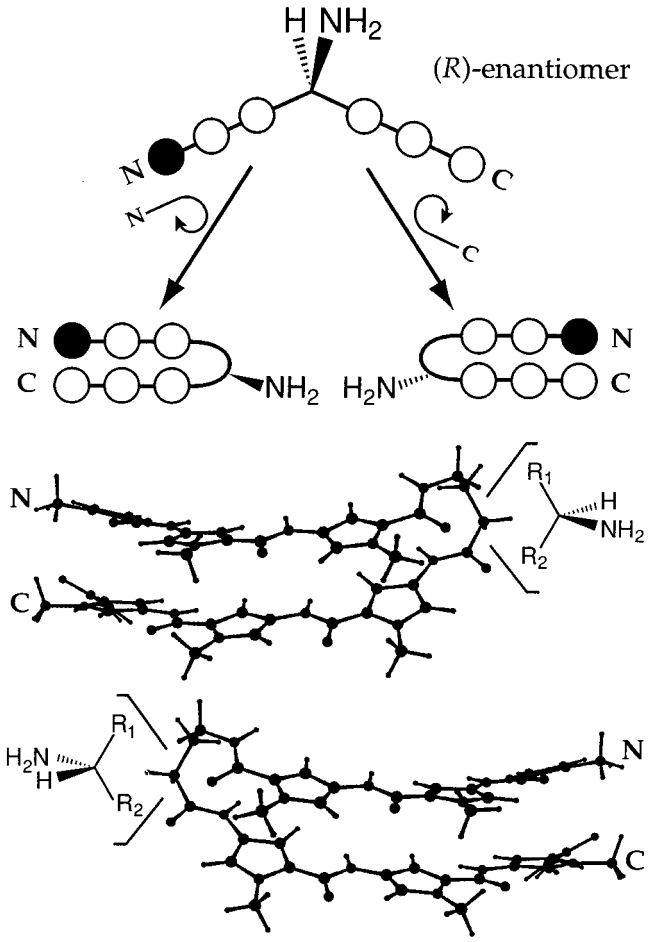


Figure 10. Model for chiral hairpin folding. Filled and unfilled circles represent Py and Im residues, respectively, α-amino, and α−H are highlighted and shown in the $(R)^{H_2N}\gamma$. Folding pathways leading to hairpin structures suitable for (left) polyamide (N−C) recognition of DNA "forward orientation" (5′−3′) and (right) polyamide (C−N) recognition of DNA "reverse orientation" (5′−3′). The corresponding stereochemistry of the α-position of the γ-turn is highlighted for each fold. "Forward" hairpin structure model for ImPyPy-γ-PyPyPy was generated from NMR structure coordinates[9i] using Chem3D software. "Reverse" hairpin structure (bottom) was generated by performing a mirror transformation of the "forward" hairpin.

is predicted to clash with the floor of the helix. The modest enhanced specificity of chiral polyamides **1**-*R* and **2**-*R* relative to the unsubstituted parent hairpin may result from stabilization of the forward binding mode and/or destabilization of the reverse binding hairpin fold.

Implications for the Design of Minor Groove Binding Molecules. The results presented here reveal properties of chiral structure elements that will guide future polyamide design: (i)

Amine substituents on the $(R)^{H_2N}\gamma$-turn amino acid enhance DNA binding affinity and specificity relative to the unsubstituted parent hairpin, providing for an optimized class of hairpin polyamides. (ii) Acetamido substituents at the $(R)^{H_2N}\gamma$-turn do not compromise affinity or specificity relative to the parent hairpin, providing for a convenient synthetic attachment point at the "capped" end of the molecule. (iii) $(S)^{H_2N}\gamma$ linked hairpins bind with enhanced affinity to reverse orientation sights relative to the parent hairpin, while $(R)^{H_2N}\gamma$ linked hairpin binds with enhanced specificity relative to the parent hairpin. These observations indicate that γ-turn substituents may regulate hairpin polyamide binding orientational preference. The results presented here set the stage for preparation of a variety of new chiral hairpin polyamide structures for specific recognition in the DNA minor groove.

Experimental Section

Materials. DNA restriction fragment labeling protocols, MPE·Fe(II) footprinting, affinity cleaving, DNase I footprinting, determination of equilibrium association constants, and quantitation by storage phosphor autoradiography were as previously described.[1b,c] (*R*)-α-Methoxy-α-(trifluoromethyl)phenylacetic acid ((*R*)MPTA) and (*S*)-α-methoxy-α-(trifluoromethyl)phenylacetic acid ((*S*)MPTA) were from Aldrich. Screw-cap glass peptide synthesis reaction vessels (5 and 20 mL) with a no. 2 sintered glass frit were made as described by Kent.[17] ^{1}H NMR spectra were recorded on a General Electric-QE NMR spectrometer at 300 MHz 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. Optical rotations were recorded on a JASCO Dip 1000 digital polarimeter. 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. 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. All DNA manipulations were performed according to standard protocols.[18]

ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (1-*R*). ImPyPy-$(R)^{Fmoc}\gamma$-PyPyPy-β-Pam-resin was synthesized in a stepwise fashion by machine-assisted solid-phase methods.[11] (*R*)-2-Fmoc-4-Boc-diaminobutyric acid (0.7 mmol) was incorporated as previously described for Boc-γ-aminobutyric acid. ImPyPy-$(R)^{Fmoc}\gamma$-PyPyPy-β-Pam-resin was placed in a glass 20-mL peptide synthesis vessel and treated with DMF (2 mL), followed by piperidine (8 mL) and agitated (22 °C, 30 min). ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Pam-resin was isolated by filtration and washed sequentially with an excess of DMF, DCM, MeOH, and ethyl ether and the amine-resin dried in vacuo. A sample of ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Pam-resin (240 mg, 0.18 mmol/g[19]) was treated with neat (dimethylamino)-propylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL), and the resulting solution purified by reversed phase HPLC. ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp is recovered upon lyophilization of the appropriate fractions as a white powder (32 mg, 66% recovery): $[\alpha]^{20}_D$ +14.6 (*c* 0.05, H_2O); UV (H_2O) λ_{max} 246, 310 (50 000); ^{1}H NMR (DMSO-d_6) δ 10.56 (s, 1 H), 10.47 (s, 1 H), 9.97 (s, 1 H), 9.94 (s, 1 H), 9.88 (s, 1 H), 9.4 (br s, 1 H), 8.28 (s, 3 H), 8.22 (m, 1 H), 8.03 (m, 2 H), 7.38 (s, 1 H), 7.25 (d, 1 H, J = 1.6 Hz), 7.22 (d, 1 H, J = 1.5 Hz), 7.19 (d, 1 H, J = 1.5 Hz), 7.16 (d, 1 H, J = 1.6 Hz), 7.14 (d, 1 H, J = 1.8 Hz), 7.12 (d, 1 H, J = 1.7 Hz), 7.03 (m, 2 H), 6.95 (d, 1 H, J = 1.6 Hz), 6.91 (d, 1 H, J = 1.6 Hz), 6.85 (d, 1 H, J = 1.6 Hz), 3.96 (s, 3 H), 3.83 (s, 3 H), 3.81 (m, 6 H), 3.79 (s, 3 H), 3.76 (s, 3 H), 3.33 (q, 2 H, J = 6.3 Hz), 3.25 (q, 2 H, J = 5.7 Hz), 3.05 (q, 2 H, J = 5.9 Hz), 2.96 (q, 2 H, J = 5.3 Hz), 2.71 (d, 6 H, J = 4.9 Hz), 2.32 (t, 2 H, J = 7.1 Hz), 1.95 (q, 2 H, J = 5.9 Hz), 1.70 (quintet, 2 H, J = 7.3 Hz); MALDI-TOF-MS (monoisotopic) [M + H] 992.5 (992.5 calcd for $C_{47}H_{62}N_{17}O_8$).

ImPyPy-$(R)(R)^{MTPA}\gamma$−PyPyPy-β-Dp (1-*R*,*R*). (*R*)-α-methoxy-α-(trifluoromethyl)phenylacetic acid (117 mg, 0.5 mmol) and HOBt (70 mg, 0.5 mmol) were dissolved in DMF (1 mL), DCC (100 mg, 0.5 mmol) was added, and the solution was agitated for 30 min at 22 °C. A sample of the activated ester solution (100 μL, 0.05 mmol) was added to ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (**1**-*R*) (10 mg, 0.01 mmol), DIEA (50 μL) added, and the solution agitated for 3 h (22 °C). DMF (1 mL) followed by 0.1% (wt/v) TFA (6 mL) was then added to the reaction mixture and the resulting solution purified by reversed phase HPLC (1% acetonitrile/min) under conditions which were determined to separate the diastereomers. ImPyPy-$(R)(R)^{MTPA}\gamma$−PyPyPy-β-Dp is recovered as a white powder upon lyophilization of the appropriate fractions (6 mg, 53% recovery): ^{1}H NMR (DMSO-d_6) δ 10.50 (s, 1 H), 10.14 (s, 1 H), 9.92 (s, 2 H), 9.88 (s, 1 H), 9.2 (br s, 1 H), 8.43 (d, 1 H, J = 7.0 Hz), 8.02 (m, 3 H), 7.92 (m, 1 H), 7.47 (m, 2 H), 7.41 (m, 2 H), 7.36 (s, 1 H), 7.24 (m, 1 H), 7.19 (m, 1 H), 7.15 (m, 1 H), 7.12 (m, 3 H), 7.01 (m, 2 H), 6.90 (m, 3 H), 6.83 (m, 1 H), 4.46 (q, 1 H, J = 5.5 Hz), 3.94 (s, 3 H), 3.79 (m, 9 H), 3.75 (m, 6 H), 3.32 (m, 4 H), 3.05 (m, 2 H), 2.94 (m, 2 H), 2.68 (d, 6H, J = 4.0 Hz), 2.28 (t, 2 H, J = 6.3 Hz), 1.93 (q, 2 H, J = 6.1 Hz), 1.66 (quintet, 2 H, J = 6.0 Hz), 1.18 (s, 3 H); MALDI-TOF-MS (monoisotopic) 1208.5 [M + H] (1208.5 calcd for $C_{57}H_{69}F_3N_{17}O_{10}$).

ImPyPy-$(R)(S)^{MTPA}\gamma$-PyPyPy-β-Dp 1-*R*,*S*. ImPyPy-$(R)(S)^{MTPA}\gamma$-PyPyPy-β-Dp was prepared from (*S*)-α-methoxy-α-(trifluoromethyl)-phenylacetic acid as described for **1**-*R*,*R* (5 mg, 45% recovery): ^{1}H NMR (DMSO-d_6) δ 10.47 (s, 1 H), 10.08 (s, 1 H), 9.92 (s, 2 H), 9.88 (s, 1 H), 9.2 (br s, 1 H), 8.43 (d, 1 H, J = 6.9 Hz), 8.02 (m, 3 H), 7.46 (m, 2 H), 7.40 (m, 2 H), 7.36 (s, 1 H), 7.23 (m, 1 H), 7.19 (m, 1 H), 7.14 (m, 1 H), 7.12 (m, 3 H), 7.01 (m, 2 H), 6.87 (m, 3 H), 6.83 (m, 1 H), 4.44 (q, 1 H, J = 6.5 Hz), 3.94 (s, 3 H), 3.79 (m, 9 H), 3.75 (m, 6 H), 3.28 (m, 4 H), 3.06 (m, 4 H), 2.94 (m, 2 H), 2.69 (d, 6H, J = 4.5 Hz), 2.28 (t, 2 H, J = 6.5 Hz), 1.93 (q, 2 H, J = 6.1 Hz), 1.66 (quintet, 2 H, J = 6.0 Hz), 1.18 (s, 3 H); MALDI-TOF-MS (monoisotopic) [M + H] 1209.0 (1208.5 calcd for $C_{57}H_{69}F_3N_{17}O_{10}$).

ImPyPy-$(S)^{H_2N}\gamma$-PyPyPy-β-Dp (1-*S*). ImPyPy-$(S)^{H_2N}\gamma$-PyPyPy-β-Dp was prepared as described for **1**-*R* (23 mg, 49% recovery): $[\alpha]^{20}_D$ −14.2 (*c* 0.04, H_2O); ^{1}H NMR (DMSO-d_6) identical to that for **1**-*R*; MALDI-TOF-MS (monoisotopic) [M + H] 992.5 (992.5 calcd for $C_{47}H_{62}N_{17}O_8$).

ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-EtOH (2-*R*). A sample of ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Pam-resin (240 mg, 0.18 mmol/g[19]) was treated with neat ethanolamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL), and the resulting solution purified by reversed phase HPLC to provide ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-EtOH as a white powder upon lyophilization of the appropriate fractions (21 mg, 46% recovery): $[\alpha]^{20}_D$ +18.6 (*c* 0.04, H_2O); UV (H2O) λ_{max} 246, 310 (50 000); ^{1}H NMR (DMSO-d_6) δ 10.55 (s, 1 H), 10.48 (s, 1 H), 9.97 (s, 1 H), 9.94 (s, 1 H), 9.89 (s, 1 H), 8.24 (m, 4 H), 8.00 (t, 1 H, J = 4.1 Hz), 7.89 (t, 1 H, J = 5.8 Hz), 7.38 (s, 1 H), 7.25 (d, 1 H, J = 1.6 Hz), 7.22 (d, 1 H, J = 1.6 Hz), 7.21 (d, 1 H, J = 1.5 Hz), 7.16 (m, 2 H), 7.14 (d, 1 H, J = 1.6 Hz), 7.03 (d, 1 H, J = 1.7 Hz), 6.99 (d, 1 H, J = 1.4 Hz), 6.95 (d, 1 H, J = 1.6 Hz), 6.91 (d, 1 H, J = 1.5 Hz), 6.78 (d, 1 H, J = 1.5 Hz), 5.33 (m, 1 H), 3.95 (s, 3 H), 3.83 (s, 3 H), 3.81 (m, 6 H), 3.79 (s, 3 H), 3.76 (s, 3 H), 3.37 (q, 2 H, J = 6.2 Hz), 3.07 (q, 2 H, J = 5.9 Hz), 2.29 (t, 2 H, J = 7.1 Hz), 1.93 (q, 2 H, J = 5.8 Hz), 1.20 (m, 4 H); MALDI-TOF-MS (monoisotopic) [M + H] 951.4 (951.4 calcd for $C_{44}H_{55}N_{16}O_9$).

ImPyPy-$(R)^{Ac}\gamma$-PyPyPy-β-Dp (3-*R*). A sample of ImPyPy-$(R)^{H_2N}\gamma$-PyPyPy-β-Dp (4 mg) in DMSO (1 mL) was treated with a solution of acetic anhydride (1 mL) and DIEA (1 mL) in DMF (1 mL) and heated (55 °C) with periodic agitation for 30 min. Residual acetic anhydride was hydrolyzed (0.1 M NaOH, 1 mL, 55 °C, 10 min), 0.1% (wt/v) TFA was added (6 mL), and the resulting solution was purified by reversed phase HPLC to provide ImPyPy-$(R)^{Ac}\gamma$-PyPyPy-β-Dp as a white powder upon lyophilization of the appropriate fractions (2 mg,

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

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

(19) 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.

50% recovery): $[\alpha]_{20}^{D}$ +20.5 (*c* 0.06, H_2O); UV (H_2O) λ_{max} 242, 304 (50 000); ^{1}H NMR (DMSO-d_6) δ 10.49 (s, 1 H), 10.06 (s, 1 H), 9.94 (m, 2 H), 9.00 (s, 1 H), 9.4 (br s, 1 H), 8.21 (d, 1 H, J = 7.8 Hz), 8.06 (m, 2 H), 8.00 (t, 1 H, J = 6.2 Hz), 7.39 (s, 1 H), 7.27 (d, 1 H, J = 1.7 Hz), 7.21 (d, 1 H, J = 1.6 Hz), 7.18 (m, 2 H), 7.14 (m, 2 H), 7.03 (m, 2 H), 6.90 (d, 1 H, J = 1.6 Hz), 6.86 (m, 2 H), 4.43 (q, 1 H, J = 7.5 Hz), 3.96 (s, 3 H), 3.82 (m, 9 H), 3.73 (m, 6 H), 3.37 (q, 2 H, J = 5.8 Hz), 3.11 (q, 2 H, J = 6.9 Hz), 2.98 (q, 2 H, J = 5.4 Hz), 2.79 (q, 2 H, J = 5.3 Hz), 2.71 (d, 6 H, J = 4.7 Hz), 2.33 (t, 2 H, J = 6.2 Hz), 1.97 (s, 3 H), 1.70 (quintet, 2 H, J = 6.0 Hz) MALDI-TOF-MS (av) 1035.1 (1035.2 calcd for M + H).

ImPyPy-(*S*)$^{Ac}\gamma$-PyPyPy-β-Dp (3-*S*). ImPyPy-(*S*)$^{Ac}\gamma$-PyPyPy-β-Dp was prepared as described for **3**-*R* (2 mg, 50% recovery): $[\alpha]_{20}^{D}$ −16.4 (*c* 0.07, H_2O); ^{1}H NMR (DMSO-d_6) is identical to that for **3**-*R*; MALDI-TOF-MS (monoisotopic) [M + H] 1034.6 (1034.5 calcd for $C_{49}H_{64}N_{17}O_9$).

ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Dp-NH$_2$ (4-*R*-Boc-NH$_2$). A sample of ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Pam-resin (300 mg, 0.18 mmol/g[19]) was treated a solution of Boc-anhydride (500 mg) and DIEA (1 mL) in DMF (4 mL) and heated (55 °C) with periodic agitation for 30 min. ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Pam-resin was isolated by filtration and washed sequentially with an excess of DMF, DCM, MeOH, and ethyl ether, and the dried in vacuo. A sample of ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Pam-resin (240 mg, 0.18 mmol/g[19]) was treated with neat 3,3′-diamino-*N*-methyldipropylamine (2 mL) and heated (55 °C) with periodic agitation for 16 h. The reaction mixture was then filtered to remove resin, 0.1% (wt/v) TFA added (6 mL), and the resulting solution purified by reversed phase HPLC to provide ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Dp-NH$_2$ as a white powder upon lyophilization of the appropriate fractions (18 mg, 36% recovery): $[\alpha]^{20}_{D}$ −30 (*c* 0.05, H_2O); UV (H_2O) λ_{max} 240, 306 (50 000); ^{1}H NMR (DMSO-d_6) δ 10.59 (s, 1 H), 10.16 (s, 1 H), 10.04 (m, 2 H), 10.00 (s, 1 H), 9.4 (br s, 1 H), 8.31 (d, 1 H, J = 7.8 Hz), 8.16 (m, 2 H), 8.10 (t, 1 H, J = 6.2 Hz), 7.89 (t, 1 H, J = 5.8 Hz), 7.49 (s, 1 H), 7.37 (d, 1 H, J = 1.7 Hz), 7.22 (d, 1 H, J = 1.6 Hz), 7.21 (d, 1 H, J = 1.5 Hz), 7.16 (m, 2 H), 7.14 (d, 1 H, J = 1.6 Hz), 7.03 (d, 1 H, J = 1.7 Hz), 6.99 (d, 1 H, J = 1.4 Hz), 6.95 (d, 1 H, J = 1.6 Hz), 6.91 (d, 1 H, J = 1.5 Hz), 6.78 (d, 1 H, J = 1.5 Hz), 5.33 (m, 1 H), 3.95 (s, 3 H), 3.83 (s, 3 H), 3.81 (m, 6 H), 3.79 (s, 3 H), 3.76 (s, 3 H), 3.37 (q, 2 H, J = 6.2 Hz), 3.07 (q, 2 H, J = 5.9 Hz), 2.29 (t, 2 H, J = 7.1 Hz), 1.93 (q, 2 H, J = 5.8 Hz), 1.20 (m, 4 H); MALDI-TOF-MS (monoisotopic) [M + H] 1135.3 (1135.6 calcd for $C_{54}H_{75}N_{18}O_{10}$).

ImPyPy-(*S*)$^{Boc}\gamma$-PyPyPy-β-Dp-NH$_2$ (4-*S*-Boc-NH$_2$). ImPyPy-(*S*)$^{Boc}\gamma$-PyPyPy-β-Dp-NH$_2$ was prepared as described for **4**-*R*. (16 mg, 32% recovery): $[\alpha]^{20}_{D}$ −30 (*c* 0.05, H_2O); ^{1}H NMR (DMSO-d_6) is identical to **4**-*R*-Boc-NH$_2$; MALDI-TOF-MS (monoisotopic) [M + H] 1135.4 (1135.6 calcd for $C_{54}H_{75}N_{18}O_{10}$).

ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Dp-EDTA (4-*R*-Boc). Excess EDTA-dianhydride (50 mg) was dissolved in DMSO/NMP (1 mL) and DIEA (1 mL) by heating at 55 °C for 5 min. The dianhydride solution was added to ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Dp-NH$_2$ (10.4 mg, 10 μmol) dissolved in DMSO (750 μL). The mixture was heated (55 °C, 25 min) and the remaining EDTA-anhydride hydrolyzed (0.1 M NaOH, 3 mL, 55 °C, 10 min). Aqueous TFA (0.1wt %/v) was added to adjust the total volume to 8 mL and the solution purified directly by reversed phase HPLC to provide ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Dp-EDTA (**4**-*R*-Boc) as a white powder upon lyophilization of the appropriate fractions (4 mg, 40% recovery): MALDI-TOF-MS (monoisotopic) [M + H] 1409.6 (1409.7 calcd for $C_{64}H_{89}N_{20}O_{17}$).

ImPyPy-(*S*)$^{Boc}\gamma$-PyPyPy-β-Dp-EDTA (4-*S*-Boc). ImPyPy-(*S*)$^{Boc}\gamma$-PyPyPy-β-Dp-NH$_2$ (12.0 mg, 12 μmol) was converted to **4**-*S*-Boc as described for **4**-*R*-Boc (4 mg, 33% recovery): MALDI-TOF-MS (monoisotopic) [M + H] 1409.7 (1409.7 calcd for $C_{64}H_{89}N_{20}O_{17}$).

ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA (4-*R*). A sample of ImPyPy-(*R*)$^{Boc}\gamma$-PyPyPy-β-Dp-EDTA (2.1 mg) in DMSO (750 μL) was placed in a 50-mL flask and treated with TFA (15 mL, 22 °C, 2 h). Excess TFA was removed in vacuo, water added (6 mL), and the resulting solution purified by reversed phase HPLC to provide ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA as a white powder upon lyophilization of the appropriate fractions (1.3 mg, 50% recovery): MALDI-TOF-MS (monoisotopic) [M + H] 1309.5 (1309.6 calcd for $C_{59}H_{81}N_{20}O_{15}$).

ImPyPy-(*S*)$^{H_2N}\gamma$-PyPyPy-β-Dp-EDTA (4-*S*). ImPyPy-(*S*)$^{Boc}\gamma$-PyPyPy-β-Dp-EDTA (3.0 mg) was converted to **4**-*S* as described for **4**-*R* (1 mg, 33% recovery): MALDI-TOF-MS (monoisotopic) [M + H] 1309.5 (1309.6 calcd for $C_{59}H_{81}N_{20}O_{15}$).

ImPyPy-(*R*)$^{EDTA}\gamma$-PyPyPy-β-Dp (5-*R*). Excess EDTA-dianhydride (50 mg) was dissolved in DMSO/NMP (1 mL) and DIEA (1 mL) by heating at 55 °C for 5 min. The dianhydride solution was added to ImPyPy-(*R*)$^{H_2N}\gamma$-PyPyPy-β-Dp (1.0 mg, 1 μmol) dissolved in DMSO (750 μL). The mixture was heated (55 °C, 25 min), and the remaining EDTA-anhydride was hydrolyzed (0.1 M NaOH, 3 mL, 55 °C, 10 min). Aqueous TFA (0.1 wt/v) was added to adjust the total volume to 8 mL and the solution purified directly by reversed phase HPLC to provide **5**-*R* as a white powder upon lyophilization of the appropriate fractions (0.6 mg, 60% recovery): MALDI-TOF-MS (monoisotopic) [M + H] 1266.4 (1266.6 calcd for $C_{57}H_{76}N_{19}O_{15}$).

ImPyPy-(*S*)$^{EDTA}\gamma$-PyPyPy-β-Dp (5-*S*). ImPyPy-(*S*)$^{EDTA}\gamma$−PyPyPy-β-Dp was prepared from **1**-*S* as described for **5**-*R* (6.8 mg, 16% recovery): MALDI-TOF-MS (monoisotopic) [M + H] 1266.5 (1266.6 calcd for $C_{57}H_{76}N_{19}O_{15}$).

Preparation of 3′- and 5′-End-Labeled Restriction Fragments. The plasmid pMM5 was linearized with EcoRI and BsrBI and then either 3′- or 5′-^{32}P end-labeled 135-bp fragment isolated as described. Chemical sequencing reactions were performed according to published methods.[20]

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 D.M.H., and the Howard Hughes Medical Institute for a predoctoral fellowship to E.E.B. We thank G.M. Hathaway for MALDI-TOF mass spectrometry.

Supporting Information Available: Text giving materials, DNA restriction fragment labeling protocols, MPE·Fe(II) and DNase I footprinting procedures, affinity cleaving procedures, and methods for determination of equilibrium association constants for all compounds, Figure S1 (autoradiograms for MPE·Fe(II) footprinting experiments) and Figure S2 (autoradiograms for affinity cleaving experiments) (10 pages). See any current masthead page for ordering and Web access instructions.

JA9737228

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