VOLUME 119, NUMBER 30
JULY 30, 1997

JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Discrimination of 5′-GGGG-3′, 5′-GCGC-3′, and 5′-GGCC-3′ Sequences in the Minor Groove of DNA by Eight-Ring Hairpin Polyamides

Susanne E. Swalley, Eldon E. Baird, and Peter B. Dervan*

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

Received April 14, 1997[⊗]

Abstract: Eight-ring hairpin polyamides which differ only by the linear arrangement of pyrrole (Py) and imidazole (Im) amino acids were designed for recognition of six base pair DNA sequences containing four contiguous G,C base pairs. The respective DNA binding properties of three polyamides, ImImPyPy-γ-ImImPyPy-β-Dp, ImPyImPy-γ-ImPyImPy-β-Dp, and ImImImIm-γ-PyPyPyPy-β-Dp, were analyzed by footprinting and affinity cleavage on a DNA fragment containing the respective match sites 5′-TGGCCA-3′, 5′-TGCGCA-3′, and 5′-TGGGGA-3′. Quantitative footprint titrations demonstrate that ImImPyPy-γ-ImImPyPy-β-Dp binds the designed match site 5′-TGGCCA-3′ with an equilibrium association constant of $K_a = 1 \times 10^{10}$ M^{-1} and >250-fold specificity versus the mismatch sequences, 5′-TGCGCA-3′ and 5′-TGGGGA-3′. The polyamides ImPyImPy-γ-ImPyImPy-β-Dp and ImImImIm-γ-PyPyPyPy-β-Dp recognize their respective 5′-TGCGCA-3′ and 5′-TGGGGA-3′ match sites with reduced affinity relative to ImImPyPy-γ-ImImPyPy-β-Dp, but again with high specificity with regard to mismatch sites. These results expand the DNA sequence repertoire targeted by pyrrole-imidazole polyamides and identify sequence composition effects which will guide further second-generation polyamide design for DNA recognition.

Introduction

Small molecules that target specific 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 in the minor groove. 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] These "pairing rules" have been utilized to target a wide variety of sequences 5−13 base pairs in size.[2−8] Pyrrole-imidazole polyamides have

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

(1) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559.

(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.

(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) 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.

(4) (a) Mrksich, M; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572. (b) Geierstanger, G. H.; Dwyer, J. T.; Bathini, Y.; Lown, J. W.; Wemmer, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 4474. (c) Geierstanger, B. H.; Jacobsen, J-P; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Biochemistry* **1994**, *33*, 3055.

(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) Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6153. (d) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 8198. (e) Pilch, D. S.; Poklar, 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.

S0002-7863(97)01184-0 CCC: $14.00

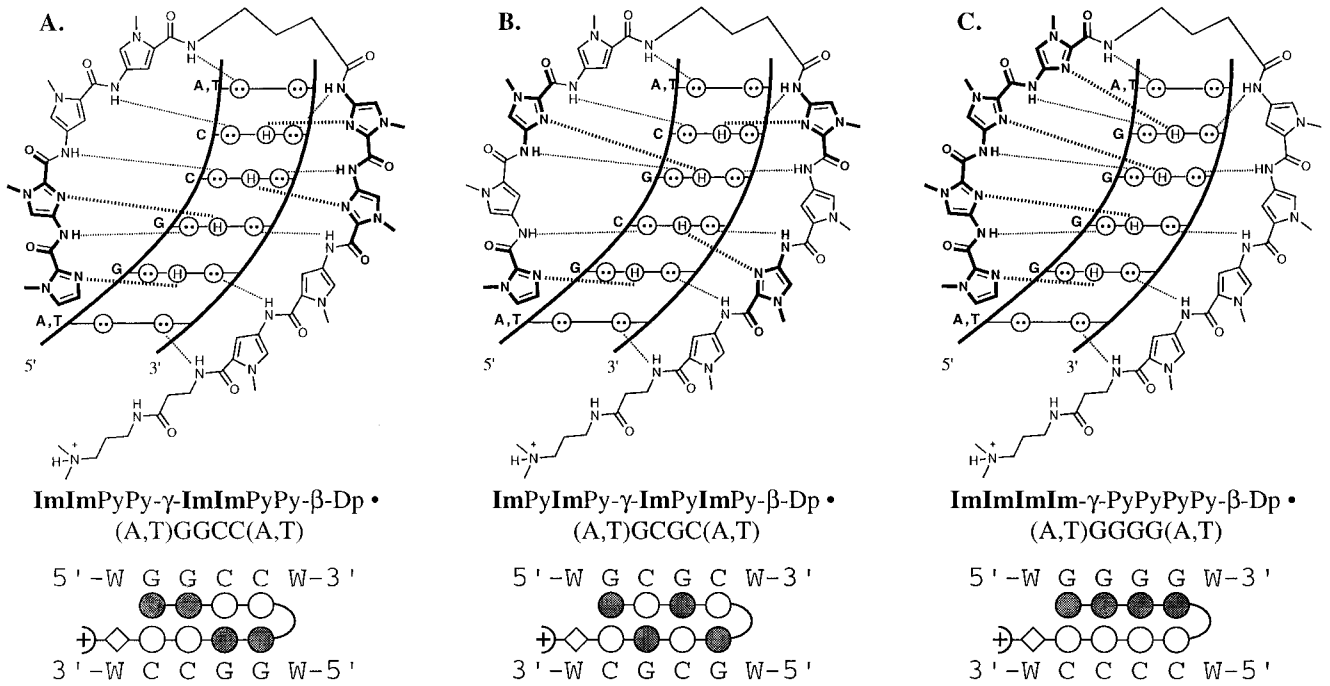


Figure 1. Binding models for the complexes formed between the DNA and the three isomeric eight-ring hairpin polyamides, with the imidazole amino acids in bold: (A) **Im**Im**Py**Py-γ-**Im**Im**Py**Py-β-Dp, (B) **Im**Py**Im**Py-γ-**Im**Py**Im**Py-β-Dp, and (C) **ImImImIm**-γ-PyPyPyPy-β-Dp. (Top) Hydrogen bonding models of 1:1 polyamide:DNA recognition. 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. (Bottom) Ball and stick models of hairpin-DNA interaction. Shaded and nonshaded circles denote imidazole and pyrrole carboxamides, respectively. Nonshaded diamonds represent the β-alanine residue. W represents either an A or T base.

recently been shown to be cell permeable and to inhibit the transcription of specific genes.[9,10] This provides impetus to explore the scope and limitations of this approach for DNA recognition. Discrimination of G·C from C·G by the Im/Py pair requires precise positioning for the key hydrogen bond between the imidazole N3 and the exocyclic amine of guanine.[5e] Given the structural requirement for binding the exocyclic amine groups of contiguous G,C base pairs, it remained to be determined if polyamides could be designed to recognize pure G,C sequences with very high affinity.

Two four-ring polyamides, ImPyImPy-Dp and ImPyPyPy-Dp (Dp = ((dimethylamino)propyl)amide), which differ by a single amino acid substitution, bind as 2:1 homodimers[6] to the respective match sites 5′-WGCGCW-3′ ($K_a \sim 1 \times 10^5$ M^{-1}) and 5′-WGWWCW-3′ ($K_a = 9 \times 10^6$ M^{-1}) with equilibrium association constants K_a that differ by a factor of ~90 (W = A·T or T·A).[6,8] One possible explanation for this difference is that the presence of four exocyclic amino groups in the minor groove presents a steric hindrance to deep polyamide binding, such that any polyamide designed for sequences containing four or more G·C base pairs will bind with low affinity.[6]

As a minimum first step to high affinity binding of G,C sequences, C−N covalently coupled polyamide subunits, i.e., "hairpin polyamides" may be used. To examine whether a core sequence of purely G,C base pairs could be recognized with high affinity and specificity, three eight-ring hairpin polyamides containing the maximum four imidazole amino acids, ImImPyPy-γ-ImImPyPy-β-Dp (**1**), ImPyImPy-γ-ImPyImPy-β-Dp (**2**), and ImImImIm-γ-PyPyPyPy-β-Dp (**3**) were synthesized by solid phase methods (Figures 1 and 2).[11] The corresponding EDTA analogs, ImImPyPy-γ-ImImPyPy-β-Dp-EDTA (**1-E**), ImPyImPy-γ-ImPyImPy-β-Dp-EDTA (**2-E**), and ImImImIm-γ-PyPyPyPy-β-Dp-EDTA (**3-E**) were also constructed to confirm the orientation of the hairpin motif. Specific hydrogen bonds are expected to form between each imidazole N3 and one of the four individual guanine 2-amino groups on the floor of the minor groove (Figure 1). Covalent head-to-tail linkage of polyamide subunits by a γ-aminobutyric acid (γ) linker to form "hairpin" polyamides has been shown to increase affinity by >100-fold relative to the unlinked subunits.[5] The C-terminal β-alanine (β) residue increases both the affinity and specificity of the hairpin motif and facilitates solid phase synthesis.[5b] We report here the affinities, binding orientations, and relative selectivities of the three polyamides as determined by three separate techniques: MPE·Fe(II) footprinting,[12] affinity cleaving,[13] and DNase I footprinting.[14] Information about binding site size and location is provided by MPE·Fe(II) footprinting,

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

(7) (a) Trauger, J. W.; Baird, E. E.; Mrksich, M. Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6160. (b) Trauger, J. W.; J. W.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1996**, *3*, 369. (c) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer D. E. *Nature Struct. Biol.* **1996**, *3*, 321.

(8) Kelly, J. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 6981.

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

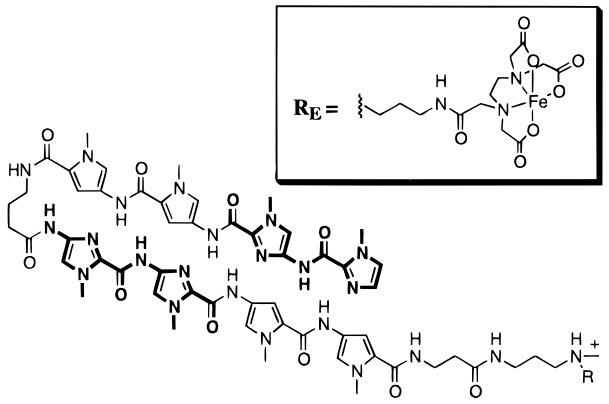
(10) For other approaches to small molecule transcription antagonists, see: (a) Maher, J. L.; Dervan, P. B.; Wold, B. *Biochemistry* **1992**, *31*, 71. (b) Duvalentin, G.; Thuong, N. T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 504. (c) Ho, S. N.; Boyer, S. H.; Schreiber, S. L.; Danishefsky, S. J.; Crabtree, G. R. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 9203. (d) Liu, C.; Smith, B. M.; Ajito, K.; Komatsu, H.; Gomezpaloma, L.; Li, T. H.; Theodorakis, E. A.; Nicolaou, K. C.; Vogt, P. K. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93* 940.

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

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

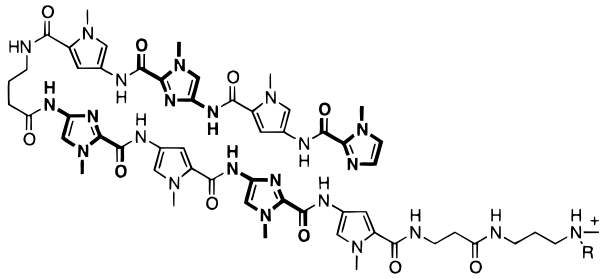
(13) (a) Schultz, P. G.; Taylor, J. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1982**, *104*, 6861. (b) Schultz, P. G.; Dervan, P. B. *J. Biomol. Struct. Dyn.* **1984**, *1*, 1133. (c) Taylor, J. S.; Schultz, P. B.; Dervan, P. B. *Terahedron* **1984**, *40*, 457.

(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.



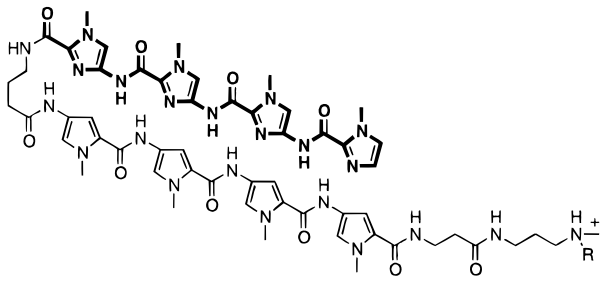
(1) R = CH_3, **ImIm**PyPy-γ-**ImIm**PyPy-β-Dp

(1-E•Fe(II)) R = R_E, **ImIm**PyPy-γ-I**mIm**PyPy-β-Dp-EDTA•Fe(II)



(2) R = CH_3, **Im**Py**Im**Py-γ-**Im**Py**Im**Py-β-Dp

(2-E•Fe(II)) R = R_E, **Im**Py**Im**Py-γ-**Im**Py**Im**Py-β-Dp-EDTA•Fe(II)



(3) R = CH_3, **ImImImIm**-γ-PyPyPyPy-β-Dp

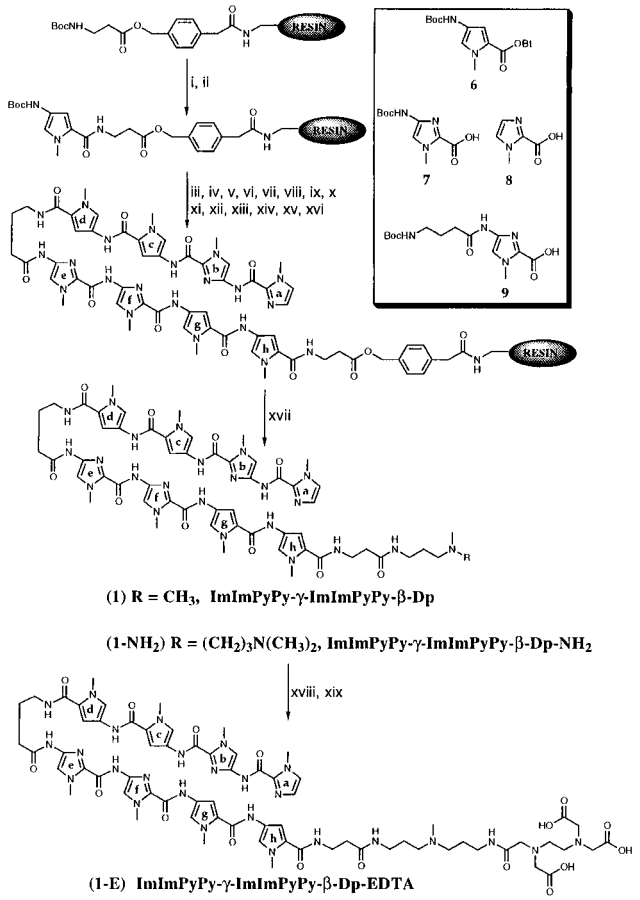
(3-E•Fe(II)) R = R_E, **ImImImIm**-γ-PyPyPyPy-β-Dp-EDTA•Fe(II)

Figure 2. Structures of the eight-ring hairpin polyamides **1**−**3** and the corresponding Fe(II)·EDTA affinity cleaving derivatives. The imidazole amino acids are in bold print.

while binding orientations are demonstrated by affinity cleaving. Quantitative DNase I footprint titrations allow the determination of equilibrium association constants (K_a) of the polyamides for their respective match sites.

Results

Synthesis of Polyamides. The polyamides ImImPyPy-γ-ImImPyPy-β-Dp (**1**), ImPyImPy-γ-ImPyImPy-β-Dp (**2**), and ImImImIm-γ-PyPyPyPy-β-Dp (**3**) were synthesized in a stepwise manner from Boc-β-alanine-Pam resin (1 g resin/0.2 mmol/g substitution) using Boc-chemistry machine-assisted protocols in 16, 12, and 18 steps, respectively (Figure 3).[11] The polyamides ImImPyPy-β-Dp (**4**) and ImPyImPy-β-Dp (**5**) were synthesized by manual solid phase methods in 8 and 6 steps, respectively. The γ-Im and PyIm subunits were introduced to polyamides **1**, **2**, and **5** as dimer-blocks in order to avoid the slow coupling of γ or Py to Im. A sample of resin (240 mg) was then cleaved by a single-step aminolysis reaction with ((dimethylamino)propyl)amine (55 °C, 18 h) and subsequently purified by reverse phase HPLC chromatography to provide



(1) R = CH_3, ImImPyPy-γ-ImImPyPy-β-Dp

(1-NH_2) R = $(CH_2)_3N(CH_3)_2$, ImImPyPy-γ-ImImPyPy-β-Dp-NH_2

(1-E) ImImPyPy-γ-ImImPyPy-β-Dp-EDTA

Figure 3. Solid phase synthetic scheme exemplified for ImImPyPy-γ-ImImPyPy-β-Dp, ImImPyPy-γ-ImImPyPy-β-Dp-NH_2, and ImImPyPy-γ-ImImPyPy-β-Dp-EDTA starting from commercially available Boc-β-Pam-resin: (i) 80% TFA/DCM, 0.4 M PhSH; (ii) Boc-Py-OBt, DIEA, DMF; (iii) 80% TFA/DCM, 0.4 M PhSH; (iv) Boc-Py-OBt, DIEA, DMF; (v) 80% TFA/DCM, 0.4 M PhSH; (vi) Boc-Im-OBt (DCC, HOBt), DIEA, DMF; (vii) 80% TFA/DCM, 0.4 M PhSH; (viii) Boc-γ-Im-OBt (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) Boc-Im-OBt (DCC, HOBt), DIEA, DMF; (xv) 80% TFA/DCM, 0.4 M PhSH; (xvi) Im-OBt (HBTU, DIEA), DMF; (xvii) *N*,*N*-dimethylaminopropylamine (55 °C, 18 h) for **1**; 3,3′-diamino-*N*-methyldipropylamine (55 °C, 18h) for **1-NH_2**; (xviii) EDTA-dianhydride, DMSO, NMP, DIEA (55 °C, 1 h); (xix) 0.1M NaOH, (55 °C, 30 min). (Inset) Pyrrole and imidazole monomers for solid phase synthesis: Boc-Pyrrole-OBt ester (Boc-Py-OBt) **6**, Boc-imidazole acid (Boc-Im-OH) **7**, imididazole-2-carboxylic acid (Im-OH) **8**,[2a] Boc-γ-Imidazole acid (Boc-γ-Im-OH) **9**.[11]

ImImPyPy-γ-ImImPyPy-β-Dp (**1**) (26 mg), ImPyImPy-γ-ImPyImPy-β-Dp (**2**) (19 mg), ImImImIm-γ-PyPyPyPy-β-Dp (12 mg) (**3**), ImImPyPy-β-Dp (**4**) (14 mg), and ImPyImPy-β-Dp (**5**) (5 mg).

Reaction of a sample of resin (240 mg) with the symmetric triamine 3,3′-diamino-*N*-methyldipropylamine (55 °C, 18 h) and subsequent purification by reverse phase HPLC chromatography provides polyamides modified with the trifluoroacetate salt of a free aliphatic primary amine group suitable for post-synthetic modification. The polyamide amines ImImPyPy-γ-ImImPyPy-β-Dp-NH_2 (**1-NH_2**), ImPyImPy-γ-ImPyImPy-β-Dp-NH_2 (**2-NH_2**), and ImImImIm-γ-PyPyPyPy-β-Dp-NH_2 (**3-NH_2**) were then treated with an excess of the dianhydride of EDTA (DMSO/NMP, DIEA, 55 °C), and the remaining anhydride was hydrolyzed (0.1 M NaOH, 55 °C). The respective EDTA modified polyamides, ImImPyPy-γ-ImImPyPy-β-Dp-EDTA (**1-E**), ImPyImPy-γ-ImPyImPy-β-Dp-EDTA (**2-E**), and ImIm-

```
5'-AATTCGAGCTCGGTACCCGGGGATCCTATGTCAGTCATGGGGATGACTGTCAGTCATGGCCATGACTGTCAGTCATGCGCATGACTGTCAGTCTTAAGCAGCT
3'-TTAAGCTCGAGCCATGGGCCCCTAGGATACAGTCAGTACCCCTACTGACAGTCAGTACCGGTACTGACAGTCAGTACGCGTACTGACAGTCAGAATTCGTCGA

TGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTA
ACCGCATTAGTACCAGTATCGACAAAGGACACACTTTAACAATAGGCGAGTGTTAAGGTGTGTTGTATGCTCGGCCTTCGTATTTCACATTTCGGACCCCACGGAT

ATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAG-3'
TACTCACTCGATTGAGTGTAATTAACGCAACGCGAGTGACGGGCGAAAGGTCAGCCCTTTGGACAGCACGGTC-5'
```

Figure 4. Sequence of the 282 base pair *Eco*RI/*Pvu*II restriction fragment from plasmid pSES9hp. Only the 6-bp sites that are boxed and in bold print (5′-TGGGGA-3′, 5′-TGGCCA-3′, and 5′-TGCGCA-3′) were analyzed by quantitative DNase I footprint titrations. The 5′-TGGCGT-3′ and 5′-TGGTCA-3′ sites that are in bold print but unboxed are single base pair mismatches recognized by ImImPyPy-γ-ImImPyPy-β-Dp (**1**). The other two bolded sites, 5′-TGGGGT-3′ and 5′-TGCGCT-3′, are further match sequences for polyamides **3** and **2**, respectively.

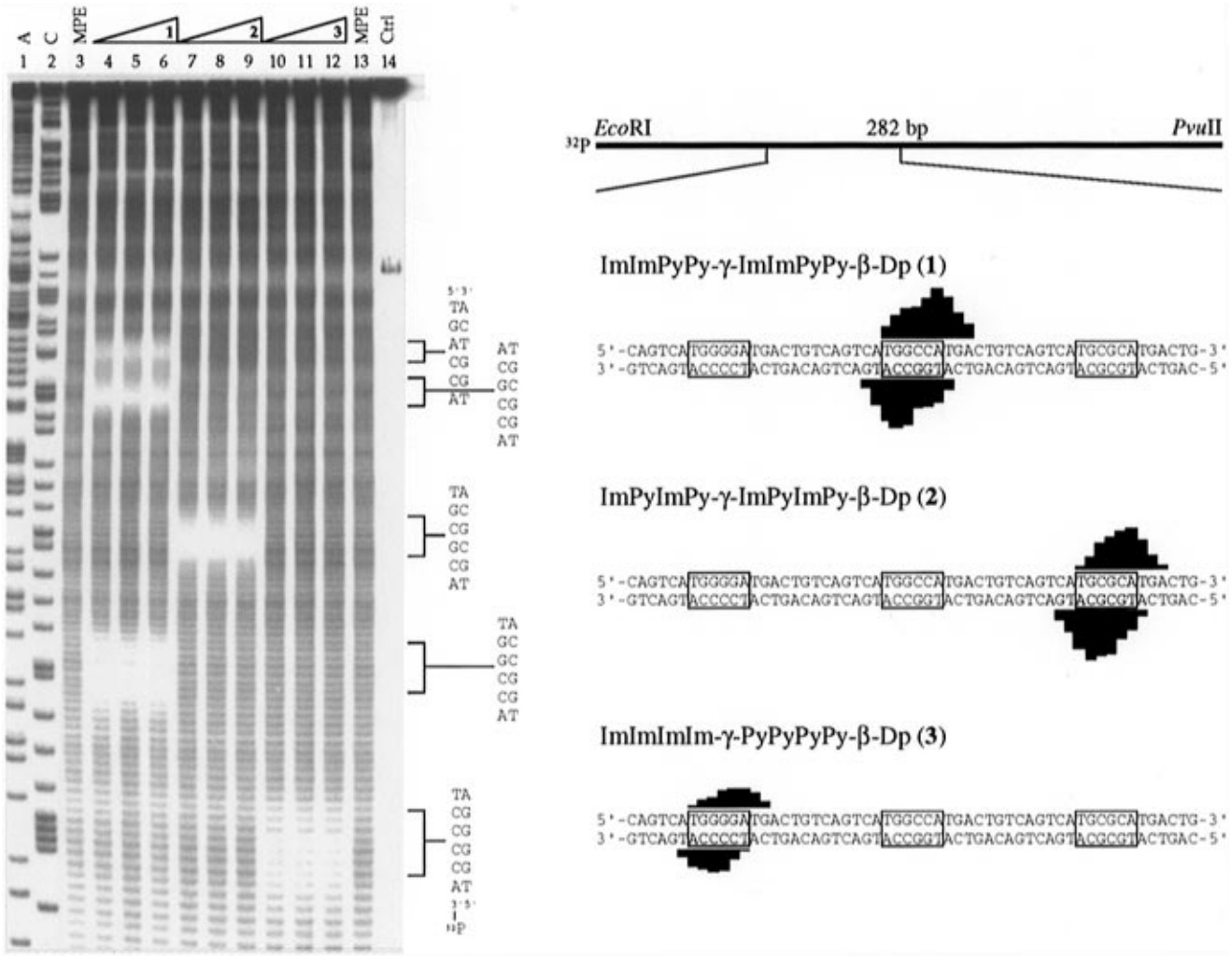


Figure 5. (Left) MPE·Fe(II) footprinting experiment[12] on the 3′-^{32}P-labeled 282 bp *Eco*RI/*Pvu*II restriction fragment from plasmid pSES9hp. The targeted 5′-TGGCCA-3′, 5′-TGCGCA-3, and 5′-TGGGGA-3′ sites are shown on the right side of the autoradiogram, as are the 5′-TGGTCA3′ and 5′-TGGCGT-3′ sites. Lane 1, A reaction; lane 2, C reaction; lanes 3 and 13, MPE·Fe(II) standard; lanes 4−6, 2, 5, and 10 μM ImImPyPy-γ-ImImPyPy-β-Dp (**1**); lanes 7−9, 2, 5, and 10 μM ImPyImPy-γ-ImPyImPy-β-Dp (**2**); lanes 10−12, 2, 5, and 10 μM ImImImIm-γ-PyPyPyPy-β-Dp (**3**); lane 14, intact DNA. 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. (Right) MPE·Fe(II) protection patterns for ImImPyPy-γ-ImImPyPy-β-Dp, ImPyImPy-γ-ImPyImPy-β-Dp, and ImImImIm-γ-PyPyPyPy-β-Dp at 10 μM concentration. (Top) Illustration of the 282 bp restriction fragment with the position of the sequence indicated. Bar heights are proportional to the relative protection from cleavage at each band. Boxes represent equilibrium binding sites determined by the published model, and only sites that were quantitated by DNase I footprint titrations are boxed.

ImIm-γ-PyPyPyPy-β-Dp-EDTA (**3-E**), were then isolated by reversed phase HPLC. Eight-ring hairpin polyamides are stable for storage at room temperature as the trifluoroacetate salt, are soluble in aqueous solution at concentrations ≤1 mM, and can be synthesized by solid phase methods in sufficient quantity (20−200 mg) and purity for biological applications.[9,11]

Identification of Binding Site Size and Location by MPE·Fe(II) Footprinting. MPE·Fe(II) footprinting (25 mM Tris-acetate, 10 mM NaCl, 100 μM/base pair calf thymus DNA, pH 7.0 and 22 °C)[12] on the 3′- and 5′-^{32}P end-labeled 282 base pair *Eco*RI/*Pvu*II restriction fragment from the cloned plasmid pSES9hp (Figure 4) reveals that polyamides **1**−**3** are protecting their targeted match sequences, 5′-TGGCCA-3′, 5′-TGCGCA-3′, and 5′-TGGGGA-3′, respectively (Figure 5). Polyamide **1** also recognizes two singly-mismatched sequences, 5′-TGGCGT-3′ and 5′-TGGTCA-3′. The footprinting patterns for the eight-ring hairpin polyamides are consistent with binding sites that are six base pairs in length.

Identification of Binding Orientation by Affinity Cleaving. Affinity cleavage assays (25 mM Tris-acetate, 10 mM NaCl, 100 μM/base pair calf thymus DNA, pH 7.0 and 22 °C)[13] were performed in order to identify the binding orientations of the

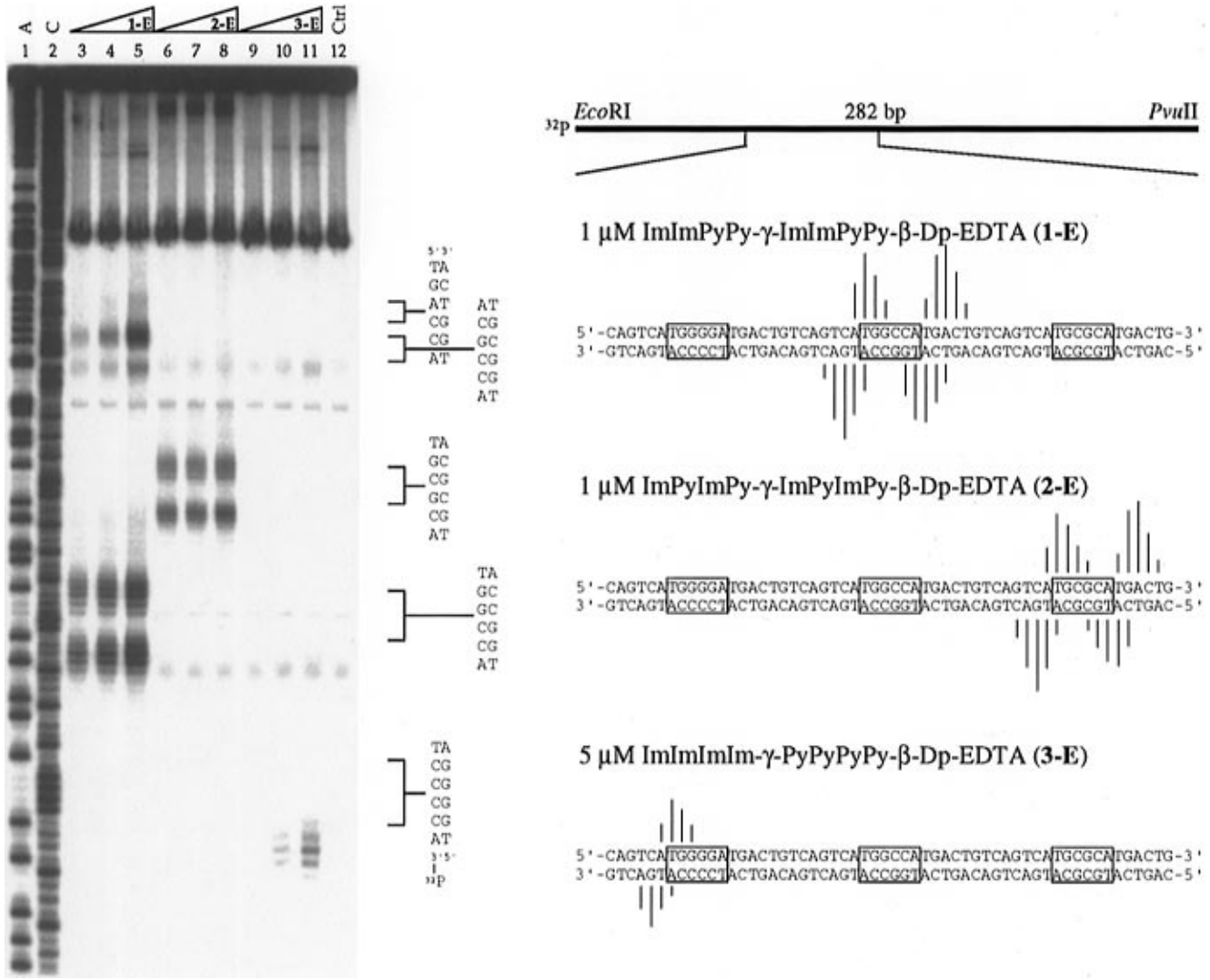


Figure 6. (Left) Affinity cleaving experiment[13] on the 3′-^{32}P-labeled 282 bp *Eco*RI/*Pvu*II restriction fragment from plasmid pSES9hp. The targeted 5′-TGGCCA-3′, 5′-TGCGCA-3, and 5′-TGGGGA-3′ sites are shown on the right side of the autoradiogram, as are the 5′-TGGTCA3′ and 5′-TGGCGT-3′ sites. Lane 1, A reaction; lane 2, C reaction; lanes 3−5, 1, 2, and 5 μM ImImPyPy-γ-ImImPyPy-β-Dp-EDTA (**1-E**); lanes 6−8, 1, 2, and 5 μM ImPyImPy-γ-ImPyImPy-β-Dp-EDTA (**2-E**); lanes 9−11, 1, 2, and 5 μM ImImImIm-γ-PyPyPyPy-β-Dp-EDTA (**3-E**); lane 12, intact DNA. 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. (Right) Affinity cleavage patterns for ImImPyPy-γ-ImImPyPy-β-Dp-EDTA and ImPyImPy-γ-ImPyImPy-β-Dp-EDTA at 1 μM concentration and ImImImIm-γ-PyPyPyPy-β-Dp-EDTA at 5 μM concentration. (Top) Illustration of the 282 bp restriction fragment with the position of the sequence indicated. Bar heights are proportional to the relative cleavage at each band. Boxes represent equilibrium binding sites determined by the published model, and only sites that were quantitated by DNase I footprint titrations are boxed.

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

polyamide	5′-TGGCCA-3′	5′-TGCGCA-3′	5′-TGGGGA-3′
ImImPyPy-γ-ImImPyPy-β-Dp (**1**)	9.7×10^9 (1.3)	$<2 \times 10^7$	$<2 \times 10^7$
ImPyImPy-γ-ImPyImPy-β-Dp (**2**)	$<10^7$	3.7×10^7 (0.9)	$<10^7$
ImImImIm-γ-PyPyPyPy-β-Dp (**3**)	$<5 \times 10^6$	$<5 \times 10^6$	2.8×10^7 (0.2)
ImImPyPy-β-Dp (**4**)	2.4×10^7 (0.6)	$<10^5$	$<10^5$
ImPyImPy-β-Dp (**5**)	$<2 \times 10^5$	$<2 \times 10^5$	$<2 \times 10^5$

[a] Values reported are the mean values measured from at least three DNase I 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 $CaCl_2$.

EDTA analogues of the three hairpin polyamides: ImImPyPy-γ-ImImPyPy-β-Dp-EDTA (**1-E**), ImPyImPy-γ-ImPyImPy-β-Dp-EDTA (**2-E**), and ImImImIm-γ-PyPyPyPy-β-Dp-EDTA (**3-E**) (Figure 6). Polyamides **1-E** and **2-E** recognize their respective palindromic match sequences, 5′-TGGCCA-3′ and 5′-TGCGCA-3′, in two equivalent orientations, consistent with hairpin formation. In contrast, polyamide **3-E** recognizes a nonpalindromic sequence, 5′-TGGGGA-3′, in a single orientation with cleavage visible only on the 5′-side of the site, as predicted by the hairpin model.

Determination of Binding Affinities by Quantitative DNase I Footprinting. Quantitative DNase I footprint titrations (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$, pH 7.0 and 22 °C)[14] were performed to determine the equilibrium association constant (K_a) of each polyamide to its respective match sequence (Table 1). Also, equilibrium association

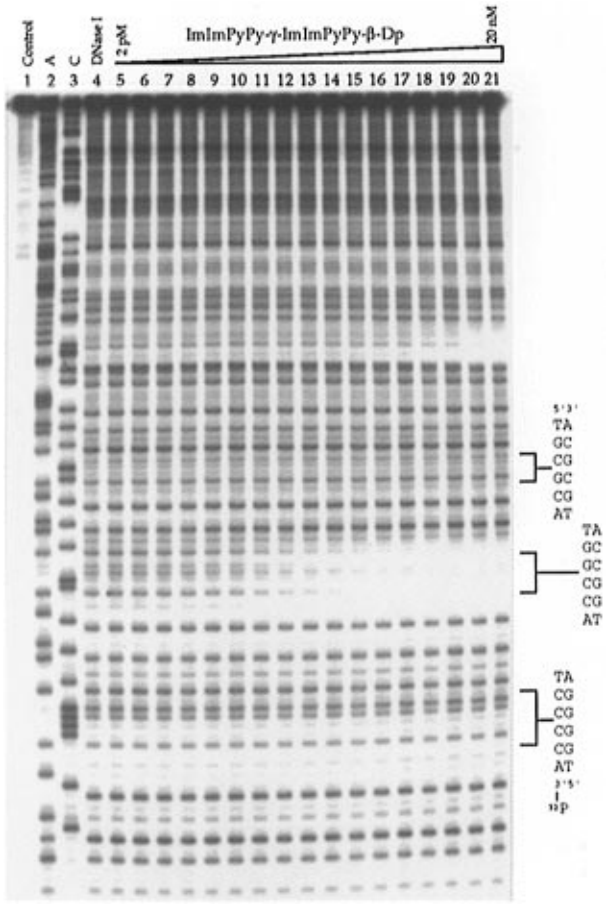


Figure 7. Quantitative DNase I footprint titration experiment[14] with ImImPyPy-γ-ImImPyPy-β-Dp on the *Eco*RI/*Pvu*II restriction fragment from plasmid pSES9hp: lane 1, intact DNA; lane 2, A reaction; lane 3, C reaction; lane 4, DNase I standard; lanes 5−21, 2 pM, 5 pM, 10 pM, 20 pM, 40 pM, 65 pM, 100 pM, 150 pM, 250 pM, 400 pM, 650 pM, 1 nM, 1.5 nM, 2.5 nM, 5 nM, 10 nM, and 20 nM ImImPyPy-γ-ImImPyPy-β-Dp. The 5′-TGGCCA-3′, 5′-TGCGCA-3′, and 5′-TGGGGA-3′ sites 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 $CaCl_2$.

constants were determined for ImImPyPy-β-Dp (**4**), the unlinked homodimer of polyamide **1** as well as for ImPyImPy-β-Dp (**5**), the unlinked homodimer of polyamide **2**. This allows a direct comparison between the linked and unlinked polyamides for identical DNA sequences. ImImPyPy-γ-ImImPyPy-β-Dp recognizes its match site 5′-caTGGCCAtg-3′ with an equilibrium association constant of $K_a = 9.7 \times 10^9$ M^{-1} (Figures 7 and 8), an increase of 400-fold over the unlinked ImImPyPy-β-Dp ($K_a = 2.4 \times 10^7$ M^{-1}). The single base pair mismatch sites identified by MPE·Fe(II) footprinting , 5′-TGGCGT-3′ and 5′-TGGTCA-3′, are positioned too close together to allow accurate quantitation of each individual site. Neither site is bound by ImImPyPy-γ-ImImPyPy-β-Dp with an equilibrium association constant greater than $\sim 4 \times 10^8$ M^{-1}, indicating at least 25-fold specificity toward these two single base pair mismatch sequences.

In contrast to polyamide **1**, both ImPyImPy-γ-ImPyImPy-β-Dp and ImImImIm-γ-PyPyPyPy-β-Dp bind their respective match sites with much lower affinity: the former recognizes the site 5′-caTGCGCAtg-3′ with a $K_a = 3.7 \times 10^7$ M^{-1} (an increase of ≥190-fold over the unlinked ImPyImPy-β-Dp), while the latter binds the site 5′-caTGGGGAtg-3′ with a $K_a = 2.8 \times 10^7$ M^{-1}. The hairpin polyamides **1** and **2** are bound with approximately two orders of magnitude higher affinity

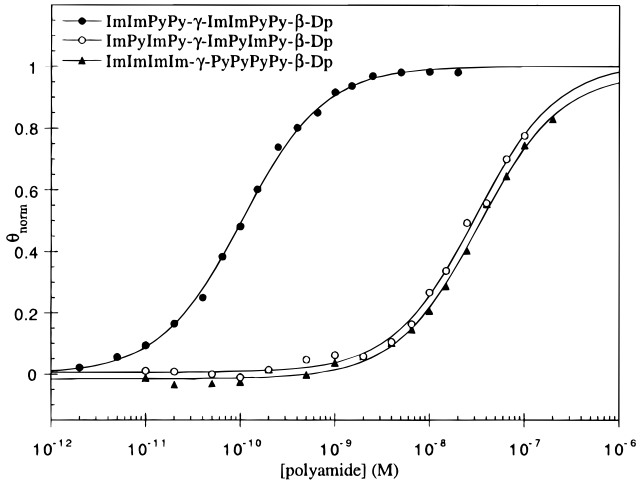


Figure 8. Data from the quantitative DNase I footprint titration experiments for the three hairpin polyamides to their respective match sites. The θ_{norm} points were obtained using photostimulable storage phosphor autoradiography and processed as described in the experimental section. The data for the binding of ImImPyPy-γ-ImImPyPy-β-Dp to 5′-caTGGCCAtg-3′ is indicated by filled circles (●), the binding of ImPyImPy-γ-ImPyImPy-β-Dp to 5′-caTGCGCAtg-3′ by open circles (○), and the binding of ImImImIm-γ-PyPyPyPy-β-Dp to 5′-caTGGGGAtg-3′ by filled triangles (▲). The solid curves are the best-fit Langmuir binding titration isotherms obtained from nonlinear least squares algorithm using eq 2, $n = 1$.

relative to the unlinked homodimer polyamides **4** and **5**, consistent with results from earlier systems.[5] Each of the three hairpin polyamides is specific for its individual target sequence. Polyamide **1** recognizes the two designed double mismatch sequences 5′-caTGCGCAtg-3′ and 5′-caTGGGGAtg-3′ with 250-fold and 350-fold lower affinity, respectively, compared to its recognition of the match site 5′-caTGGCCAtg-3′. Similarly, polyamide **2** is at least 4-fold specific for its match site compared to the double base pair mismatch sites, and polyamide **3** is at least 6-fold specific.[15]

Discussion

It was thought that pyrrole-imidazole polyamides, which sit deeply in the minor groove of DNA, would recognize G,C rich sequences with low binding affinity due to steric hinderance with the exocyclic amines of the guanine bases.[6] It has also been noted that the lower negative electrostatic potential of a G,C rich minor groove relative to an A,T rich minor groove might prohibit high affinity binding.[16] Quantitative DNase I footprint titrations reveal that the eight-ring hairpin polyamide ImImPyPy-γ-ImImPyPy-β-Dp binds with subnanomolar affinity ($K_a = 1 \times 10^{10}$ M^{-1}) to a 5′-TGGCCA-3′ sequence, an affinity that equals those seen for eight-ring hairpin polyamides that recognize six base pair sequences containing only one or two G,C base pairs.[1] Therefore, it seems likely that neither the electrostatic properties nor the steric bulk of the four exocyclic guanine amino groups interfere with polyamide binding at G,C rich sites.

Two isomeric polyamides, differing only in the ordering of the Im residues, ImPyImPy-γ-ImPyImPy-β-Dp and ImImImIm-γ-PyPyPyPy-β-Dp, bind their match sequences with far lower affinity than does ImImPyPy-γ-ImImPyPy-β-Dp. Sequence-dependent DNA structural features, such as intrinsic minor groove width, minor groove flexibility, and inherent curvature,

(15) Nonspecific DNase I inhibition by these polyamides procludes footprinting experiments at ligand concentrations greater than ~100 nM.
(16) Pullman, B. *Adv. Drug Res.* **1989**, *18*, 1.

may differ between the three binding sites and could contribute to the range of binding affinities.[17] However, the observed binding affinities do *not* correlate with the number of purine−pyrimidine or purine−purine steps, suggesting that the *positions* of the Im amino acids are critical to high affinity DNA recognition. One possible explanation is that the Im residues located at the *C*-terminal end of each four-ring polyamide subunit (positions c, d, g, and h, Figure 3; for example, ImPy**ImPy**-γ-ImPy**ImPy**-β-Dp, where the bolded residues represent the internal amino acids) are somehow less capable of strong hydrogen bond formation than the *N*-terminal residues (a, b, e, and f, Figure 3). The mispositioning of the internal Im residues may be due to imperfect curvature match between the polyamide and the DNA or alternatively "rise-per-residue mismatch" between the polyamides and the DNA.[8,18] Further support for this model comes from studies of polyamides containing the four-ring subunit ImImImPy, which recognize DNA with lower affinity than do polyamides with ImPyPyPy subunits.[1,5d]

Implications for the Design of Minor Groove Binding Molecules. Pyrrole-imidazole polyamides provide a versatile chemical method for the targeting of any predetermined DNA sequence. However, prior to this study, subnanomolar recognition of sequences containing *only* G,C base pairs in the core sequence had not been demonstrated. Here, an eight-ring hairpin polyamide is used to recognize a 5′-TGGCCA-3′ sequence with subnanomolar affinity. This work represents an important step in the understanding of polyamide-DNA recognition, since it demonstrates that the positioning of the Im amino acids has a profound effect on the binding affinities of pyrrole-imidazole polyamides. If the rise-per-residue or curvature, or a combination of both effects, of the ligand is responsible for these position-dependent effects, then perhaps binding affinity can be restored by the design of hairpin polyamides where specific ring residues have been substituted by more flexible spacer amino acids.

Experimental Section

Dicyclohexylcarbodiimide (DCC), hydroxybenzotriazole (HOBt), 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), and 0.2 mmol/g of Boc-β-alanine-(4-carboxamidomethyl)benzyl ester-copoly(styrene-divinylbenzene) resin (Boc-β-Pam-Resin) 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. Boc-γ-aminobutyric acid was from NOVA Biochem, dichloromethane (DCM), and triethylamine (TEA) was reagent grade from EM, thiophenol (PhSH), dimethylaminopropylamine from Aldrich, trifluoroacetic acid (TFA) from Halocarbon, phenol from Fisher, and ninhydrin 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. 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 no. 2 sintered glass frit were made as described by Kent.[19] ^{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 Instituite of Technology. HPLC analysis was performed on either 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. 18MΩ water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered.

Enzymes were purchased from Boehringer-Mannheim and used with their supplied buffers. Deoxyadenosine and thymidine 5′-[α-^{32}P]-triphosphates were obtained from Amersham, and deoxyadenosine 5′-[γ-^{32}P]triphosphate was purchased from I.C.N. Sonicated, deproteinized calf thymus DNA was acquired from Pharmacia. RNase free water was obtained from USB and used for all footprinting reactions. All other reagents and materials were used as received. All DNA manipulations were performed according to standard protocols.[20]

ImImPyPy-γ-ImImPyPy-β-Dp (1). ImImPyPy-γ-ImImPyPy-β-PAM-resin was prepared from 0.2 mmol/g of Boc-β-PAM-resin by machine assisted synthesis. The γ-Im step was introduced using Boc-γ-Im acid (HBTU, DIEA); all other residues were added as appropriate activated Boc protected monomer units. A sample of resin (250 mg, 0.16 mmol/g[21]) was placed in a 20 mL glass scintillation vial, 2 mL of dimethylaminopropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImImPyPy-γ-ImImPyPy-β-Dp (26 mg, 45% recovery) as a white powder. UV λ_{max}(H_2O) 248, 312 (66 000); ^{1}H NMR (DMSO-d_6) δ 10.34 (m, 2 H); 10.32 (m, 2 H); 9.73 (m, 2 H); 9.5 (br s, 1 H), 9.32 (s, 1 H); 8.10 (m, 3 H); 7.55 (m, 2 H); 7.52 (s, 1 H); 7.44 (s, 1 H); 7.23 (m, 2 H), 7.14 (m, 4 H); 7.06 (d, 1 H, J = 1.4 Hz); 6.86 (m, 2 H); 3.98 (m, 9 H); 3.95 (s, 3 H); 3.81 (m, 6 H); 3.77 (m, 6 H); 3.31 (m, 2 H); 3.17 (t, 2 H, J = 5.5 Hz); 3.06 (m, 2 H, J = 5.7 Hz); 2.93 (m, 2 H, J = 4.7 Hz); 2.74 (d, 6 H, J = 4.4 Hz); 2.30 (m, 4 H); 1.74 (m, 4 H); MALDI-TOF-MS, 1224.9 (1225.3 calc. for M + H).

ImPyImPy-γ-ImPyImPy-β-Dp (2). ImPyImPy-γ-ImPyImPy-β-PAM-resin was prepared from 0.2 mmol/g Boc-β-PAM-resin by manual polyamide synthesis. The Py-Im and γ-Im steps were introduced using Boc-γ-Im acid and Boc-Py-Im acid (HBTU, DIEA); all other residues were added as appropriate activated Boc protected monomer units. A sample of resin (250 mg, 0.16 mmol/g[21]) was placed in a 20 mL glass scintillation vial, 2 mL of dimethylaminopropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImPyImPy-γ-ImPyImPy-β-Dp (19 mg, 32% recovery) as a white powder. UV λ_{max}(H_2O) 246, 312 (66 000); ^{1}H NMR (DMSO-d_6) δ 10.33 (m, 2 H); 10.25 (m, 2 H); 10.04 (m, 2 H); 9.95 (s, 1 H); 9.5 (br s, 1 H), 8.10 (m, 3 H); 7.57 (m, 2 H); 7.48 (s, 1 H); 7.42 (s, 1 H); 7.40 (s, 1 H); 7.23 (m, 3 H), 7.17 (d, 1 H; J = 1.5 Hz); 7.03 (d, 1 H, J = 1.5 Hz); 6.98 (m, 3 H); 4.02 (s, 3 H); 3.99 (m, 6 H); 3.97 (s, 3 H); 3.88 (m, 6 H); 3.83 (m, 6 H); 3.42 (m, 2 H); 3.18 (t, 2 H, J = 5.2 Hz); 3.06 (m, 2 H, J = 5.5 Hz); 2.80 (m, 2 H, J = 4.7 Hz); 2.76 (d, 6 H, J = 4.4 Hz); 2.38 (m, 4 H); 1.93 (m, 4 H); MALDI-TOF-MS, 1225.2. (1225.3 calc. for M + H).

ImImImIm-γ-PyPyPyPy-β-Dp (3). ImImImIm-γ-PyPyPyPy-β-PAM-resin was prepared from 0.2 mmol/g of Boc-β-PAM-resin by manual polyamide synthesis. All residues were added as appropriate

(17) (a) Wu, H.; Crothers, D. M. *Nature* **1984**, *308*, 509. (b) Steitz, T. A. *Q. Rev. Biophys.* **1990**, *23*, 205. (c) Goodsell, D. S.; Kopka, M. L.; Cascio, D.; Dickerson, R. E. *Proc. Natl. Sci. U.S.A.* **1993**, *90*, 2930. (d) Paolella, D. N.; Palmer, R.; Schepartz, A. *Science* **1994**, *264*, 1130. (e) Kahn, J. D.; Yun, E.; Crothers, D. M. *Nature* **1994**, *368*, 163. (f) Geierstanger, B. H.; Wemmer, D. E. *Ann. Rev. Biochem.* **1995**, *24*, 463. (g) Hansen, M. R.; Hurley, L. H. *Acc. Chem. Res.* **1996**, *29*, 249.

(18) (a) Koopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *J. Mol. Biol.* **1985**, *183*, 553. (b) Coll, M.; Fredrick, C. A.; Wang, A. H.; Rich, A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8385.

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

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

(21) Resin substitution can be calculated as L_{new}(mmol/g) = L_{old}/(1 + L_{old}(W_{new} − W_{old}) × 10^{-3}), where L is the loading (mmol of amine per g of resin), and W is the weight (g mol^{-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.

activated Boc protected monomer units. A sample of resin (250 mg, 0.16 mmol/g[21]) was placed in a 20 mL glass scintillation vial, 2 mL of dimethylaminopropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImImImIm-γ-PyPyPyPy-β-Dp (12 mg, 21% recovery) as a white powder. UV λ_{max}(H_2O) 246, 314 (66 000); ^{1}H NMR (DMSO-d_6) δ 9.91 (m, 2 H); 9.89 (m, 3 H); 9.83 (s, 1 H); 9.60 (s, 1 H); 9.5 (br s, 1 H); 8.34 (m, 1 H); 8.10 (m, 2 H); 7.63 (m, 2 H); 7.50 (s, 1 H); 7.42 (s, 1 H); 7.19 (m, 2 H), 7.13 (m, 3 H); 7.04 (m, 2 H); 6.86 (m, 2 H); 3.98 (m, 6 H); 3.96 (s, 3 H); 3.93 (s, 3 H); 3.81 (m, 6 H); 3.77 (s, 3 H); 3.73 (s, 3 H); 3.30 (m, 2 H); 3.10 (t, 2 H, J = 5.3 Hz) 3.09 (m, 2 H, J = 5.5 Hz); 2.91 (m, 2 H, J = 4.6 Hz); 2.71 (d, 6 H, J = 4.2 Hz); 2.32 (m, 4 H); 1.70 (m, 4 H); MALDI-TOF-MS, 1225.6 (1225.3 calc. for M + H).

ImImPyPy-β-Dp (4). ImImPyPy-β-PAM-resin was prepared from 0.2 mmol/g of Boc-β-PAM-resin by machine assisted synthesis. Residues were added as appropriate activated Boc protected monomer units. A sample of resin (250 mg, 0.18 mmol/g[21]) was placed in a 20 mL glass scintillation vial, 2 mL of dimethylaminopropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImImPyPy-β-Dp (14 mg, 54% recovery) as a white powder. UV λ_{max}(H_2O) 248, 308 (33 000); ^{1}H NMR (DMSO-d_6) δ 10.35 (s, 1 H); 9.91 (s, 1 H); 9.72 (s, 1 H); 9.3 (br s, 1 H), 8.03 (m, 2 H); 7.55 (s, 1 H); 7.44 (s, 1 H); 7.24 (d, 1 H, J = 1.6 Hz); 7.14 (d, 1 H, J = 1.6 Hz), 7.11 (d, 1 H, J = 1.7 Hz); 7.05 (d, 1 H, J = 1.6 Hz); 6.84 (d, 1 H, J = 1.6 Hz); 3.97 (s, 3 H); 3.81 (s, 3 H); 3.77 (s, 6 H); 3.45 (q, 2 H, J = 5.6 Hz); 3.36 (q, 2 H, J = 6.1 Hz); 3.07 (q, 2 H, J = 6.2 Hz); 2.70 (d, 6 H, J = 4.9 Hz); 2.31 (t, 2 H, J = 6.9 Hz); 1.70 (quintet, 2 H, J = 7.3 Hz); MALDI-TOF-MS, 649.1 (649.6 calc. for M + H).

ImPyImPy-β-Dp (5). ImPyImPy-β-PAM-resin was prepared from 0.2 mmol/g Boc-β-PAM-resin by manual polyamide synthesis. The Py-Im step was introduced using Boc-Py-Im acid (HBTU, DIEA) all other residues were added as appropriate activated Boc protected monomer units. A sample of resin (250 mg, 0.18 mmol/g[21]) was placed in a 20 mL glass scintillation vial, 2 mL of dimethylaminopropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImPyImPy-β-Dp (5 mg, 19% recovery) as a white powder. UV λ_{max}(H_2O) 252, 312 (33 000); ^{1}H NMR (DMSO-d_6) δ 10.35 (s, 1 H); 10.25 (s, 1 H); 10.05 (s, 1 H); 9.3 (br s, 1 H), 8.10 (m, 2 H); 7.58 (s, 1 H); 7.42 (s, 1 H); 7.40 (d, 1 H, J = 1.5 Hz); 7.24 (d, 1 H, J = 1.5 Hz), 7.17 (d, 1 H; J = 1.5 Hz); 7.07 (d, 1 H; J = 1.5 Hz); 6.93 (d, 1 H, J = 1.5 Hz); 4.01 (s, 3 H); 3.99 (s, 3 H); 3.88 (s, 3 H); 3.82 (s, 3 H); 3.41 (q, 2 H, J = 5.8 Hz); 3.31 (m, 2 H); 3.03 (m, 2 H); 2.73 (d, 6 H, J = 4.4 Hz); 2.32 (t, 2 H, J = 6.1 Hz); 1.74 (m, 2 H); MALDI-TOF-MS, 650.3 (649.6 calc. for M + H).

ImImPyPy-γ-ImImPyPy-β-Dp-NH$_2$ (1-NH$_2$). A sample of ImImPyPy-γ-ImImPyPy-β-PAM-resin (250 mg, 0.16 mmol/g[21]) was placed in a 20 mL glass scintillation vial, 2 mL of 3,3′-diamino-*N*-methyldipropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImImPyPy-γ-ImImPyPy-β-Dp-NH$_2$ (30 mg, 54% recovery) as a white powder. ^{1}H NMR (DMSO-d_6) δ 10.35 (m, 2 H); 10.32 (m, 2 H); 9.71 (m, 2 H); 9.6 (br s, 1 H), 9.30 (s, 1 H); 8.0 (m, 3 H); 7.8 (br s, 3 H); 7.55 (m, 2 H); 7.48 (s, 1 H); 7.41 (s, 1 H); 7.22 (d, 2 H); 7.14 (m, 4 H); 7.04 (d, 1 H, J = 1.2 Hz); 6.86 (d, 2 H); 4.00 (m, 9 H); 3.94 (s, 3 H); 3.80 (m, 12 H); 3.57 (q, 2 H, J = 4.5 Hz), 3.2- 3.0 (m, 8 H), 2.83 (q, 2 H, J = 4.5 Hz), 2.71 (d, 3 H, J = 4.8 Hz), 2.46 (t, 2 H, J = 6.0 Hz), 2.32 (t, 2 H, J = 5.9 Hz), 1.87 (quintet, 2 HJ = 5.7 Hz), 1.76 (m, 4 H). MALDI-TOF-MS, 1267.9 (1268.4 calc. for M + H).

ImPyImPy-γ-ImPyImPy-β-Dp-NH$_2$ (2-NH$_2$). A sample of ImPyImPy-γ-ImPyImPy-β-PAM-resin (250 mg, 0.16 mmol/g[23]) was placed in a 20 mL glass scintillation vial, 2 mL of 3,3′-diamino-*N*-methyldipropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by reversed phase HPLC to provide ImPyImPy-γ-ImPyImPy-β-Dp-NH$_2$ (24 mg, 43% recovery) as a white powder. ^{1}H NMR (DMSO-d_6) δ 10.33 (m, 2 H); 10.22 (m, 2 H); 10.01 (m, 2 H); 9.96 (s, 1 H); 9.8 (br s, 1 H), 8.0 (m, 3 H); 7.8 (br s, 3H); 7.55 (m, 2 H); 7.48 (s, 1 H); 7.41 (s, 1 H); 7.23 (d, 2 H); 7.17 (d, 2 H); 7.02 (d, 1 H, J = 1.6 Hz); 6.96 (m, 4 H); 4.03 (s, 3 H); 3.99 (m, 6 H); 3.96 (s, 3 H); 3.90 (m, 6 H); 3.81 (m, 6 H); 3.51 (q, 2 H, J = 5.1 Hz); 3.2−3.0 (m, 8 H); 2.80 (q, 2 H, J = 5.1 Hz), 2.71 (d, 3 H, J = 4.6 Hz), 2.41 (t, 2 H, J = 5.9 Hz), 2.28 (t, 2 H, J = 5.9 Hz), 1.88 (quintet, 2 H; J = 6.2 Hz), 1.73 (m, 4 H). MALDI-TOF-MS, 1268.0 (1268.4 calc. for M + H).

ImImImIm-γ-PyPyPyPy-β-Dp-NH$_2$ (3-NH$_2$). A sample of ImImImIm-γ-PyPyPyPy-β-PAM-Resin (250 mg, 0.16 mmol/g[21]) was placed in a 20 mL glass scintillation vial, 2 mL of 3,3′-diamino-*N*-methyldipropylamine was added, and the mixture was allowed to stand at 55 °C for 18 h. Resin was removed by filtration through a disposable propylene filter, and the resulting solution diluted with water to a total volume of 8 mL and purified directly by preparatory reversed phase HPLC to provide ImImImIm-γ-PyPyPyPy-β-Dp (18 mg, 17% recovery) as a white powder. ^{1}H NMR (DMSO-d_6) δ 9.90 (m, 4 H); 9.81 (s, 1 H); 9.60 (s, 2 H); 9.5 (br s, 1 H); 8.30 (m, 1 H); 8.0 (m, 2 H); 7.8 (br s, 3H) 7.65 (m, 2 H); 7.55 (s, 1 H); 7.40 (s, 1 H); 7.17 (m, 2 H), 7.10 (m, 2 H); 7.03 (m, 3 H); 6.80 (m, 2 H); 4.01 (m, 6 H); 3.95 (s, 3 H); 3.90 (s, 3 H); 3.80 (m, 6 H); 3.77 (s, 3 H); 3.73 (s, 3 H); 3.55 (q, 2 H, J = 5.4 Hz); 3.2−3.0 (m, 8 H); 2.84 (q, 2 H, J = 5.4 Hz), 2.71 (d, 3 H, J = 4.8 Hz), 2.44 (t, 2 H, J = 5.8 Hz), 2.20 (t, 2 H, J = 6.0 Hz), 1.86 (quintet, 2 H; J = 6.1 Hz), 1.76 (m, 4 H). MALDI-TOF-MS, 1267.4 (1268.4 calc. for M + H).

ImImPyPy-γ-ImImPyPy-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 ImImPyPy-γ-ImImPyPy-β-Dp-NH$_2$ (**1-NH$_2$**) (10 mg, 7 μmol) dissolved in 750 μL of DMSO. The mixture was heated at 55 °C for 25 min, treated with 3 mL 0.1M NaOH, and heated at 55 °C for 10 min. TFA (0.1%) was added to adjust the total volume to 8 mL, and the solution purified directly by preparatory HPLC chromatography to provide **1-E** as a white powder. (3 mg, 26% recovery) MALDI-TOF-MS, 1543.0 (1543.6 calc. for M + H).

ImPyImPy-γ-ImPyImPy-β-Dp-EDTA (2-E). Polyamide **2-E** was prepared from ImPyImPy-γ-ImPyImPy-β-Dp-NH$_2$ (**2-NH$_2$**) (10 mg, 7 μmol) as described for **1-E**. (1 mg, 9% recovery) MALDI-TOF-MS, 1543.2 (1543.6 calc. for M + H).

ImImImIm-γ-PyPyPyPy-β-Dp-EDTA (3-E). Polyamide **3-E** was prepared from ImImImIm-γ-PyPyPyPy-β-Dp-NH$_2$ (**3-NH$_2$**) (10 mg, 7 μmol) as described for **1-E**. (2.5 mg, 22% recovery) MALDI-TOF-MS, 1542.6 (1543.6 calc. for M + H).

Construction of Plasmid DNA. The plasmid pSES9hp was constructed by hybridization of the inserts, 5′- GATCCTATGTCAGTCATGGGGATGACTGTCAGTCATGGCCATGACTGTCAGTCATGCGCATGACTGTCAGTCTTAAGC-3′ and 5′- GATACAGTCAGTACCCCTACTGACAGTCAGTACCGGTACTGACAGTCAGTACGCGTACTGACAGTCAGAATTCGTCGA-3′. The hybridized insert was ligated into linearized pUC19 *Bam*HI/*Hind*III plasmid using T4 DNA ligase. The resultant constructs were used to transform Top10F′ OneShot competent cells from Invitrogen. Ampicillin-resistant white colonies were selected from 25 mL Luria-Bertani medium agar plates containing 50 μg/mL ampicillin and treated with XGAL and IPTG solutions. Large-scale plasmid purification was performed with Qiagen Maxi purification kits. Dideoxy sequencing was used to verify the presence of the desired insert. Concentration of the prepared plasmid was determined at 260 nm using the relationship of 1 OD unit = 50 μg/mL duplex DNA.

Preparation of 3′- and 5′-End-Labeled Restriction Fragments. The plasmid pSES9hp was linearized with *Eco*RI and *Pvu*II and then

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

treated with Klenow fragment, deoxyadenosine 5′-[α-^{32}P]triphosphate and thymidine 5′-[α-^{32}P]triphosphate for 3′ labeling. Alternatively, pSES9hp was linearized with *Eco*RI, treated with calf alkaline phosphatase, and then 5′ labeled with T4 polynucleotide kinase and deoxyadenosine 5′-[γ-^{32}P]triphosphate. The 5′ labeled fragment was then digested with *Pvu*II. The labeled fragment (3′ or 5′) was loaded onto a 5% nondenaturing polyacrylamide gel, and the desired 282 base pair band was visualized by autoradiography and isolated. Chemical sequencing reactions were performed according to published methods.[22]

MPE·Fe(II) Footprinting.[12] All reactions were carried out in a volume of 40 μL. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were 25 mM Tris-acetate buffer (pH 7.0), 10 mM NaCl, 100 μM/base pair calf thymus DNA, and 30 kcpm 3′- or 5′-radiolabeled DNA. The solutions were allowed to equilibrate for 4 h. A fresh 50 μM MPE·Fe(II) solution was made from 100 μL of a 100 μM MPE solution and 100 μL of a 100 μM ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2·6H_2O$) solution. MPE·Fe(II) solution (5 μM) was added to the equilibrated DNA, and the reactions were allowed to equilibrate for 5 min. Cleavage was initiated by the addition of dithiothreitol (5 mM) and allowed to proceed for 14 min. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 5 min, and placed on ice, and half of each tube (∼15 kcpm) was immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V.

Affinity Cleaving.[13] All reactions were carried out in a volume of 40 μL. A polyamide stock solution or water (for reference lanes) was added to an assay buffer where the final concentrations were 25 mM Tris-acetate buffer (pH 7.0), 10 mM NaCl, 100 μM/base pair calf thymus DNA, and 20 kcpm 3′- or 5′-radiolabeled DNA. The solutions were allowed to equilibrate for 4 h. A fresh solution of ferrous ammonium sulfate ($Fe(NH_4)_2(SO_4)_2·6H_2O$) (10 μM) was added to the equilibrated DNA, and the reactions were allowed to equilibrate for 15 min. Cleavage was initiated by the addition of dithiothreitol (10 mM) and allowed to proceed for 30 min. Reactions were stopped by ethanol precipitation, resuspended in 100 mM tris-borate-EDTA/80% formamide loading buffer, denatured at 85 °C for 5 min, and placed on ice, and the entire sample was immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V.

DNase I Footprinting.[14] All reactions were carried out in a volume of 400 μL. We note explicitly that no carrier DNA was used in these reactions until after DNase I cleavage. 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 a minimum of 12 h at 22 °C. Cleavage was initiated by the addition of 10 μL of a DNase I stock solution (diluted with 1 mM DTT to give a stock concentration of 0.28 u/mL) and allowed to proceed for 5 min at 22 °C. The reactions were stopped by adding 50 μL of a solution containing 2.25 M NaCl, 150 mM EDTA, 0.6 mg/mL glycogen, and 30 μM base-pair 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 5 min, placed on ice, and immediately loaded onto an 8% denaturing polyacrylamide gel (5% cross-link, 7 M urea) at 2000 V for 1 h. The gels were dried under vacuum at 80 °C, then quantitated using storage phosphor technology.

Equilibrium association constants were determined as previously described.[5a] The data were analyzed by performing volume integrations of the 5′-TGGCCA-3′, 5′-TGCGCA-3′, and 5′-TGGGGA-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, $n = 1$ for polyamides **1**−**3**, $n = 2$ for polyamides **4** and **5**) by minimizing the difference between θ_{app} and θ_{fit}, using the modified Hill equation

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

where $[L]_{tot}$ corresponds to the total polyamide concentration, K_a corresponds to the equilibrium 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, θ_{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. Photostimulable storage phosphorimaging plates (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) were pressed flat against gel samples and exposed in the dark at 22 °C for 12−20 h. A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens. The data were analyzed by performing volume integrations of all bands using the ImageQuant v. 3.2.

Acknowledgment. We are grateful to the National Institutes of Health (GM-2768) for research support, the National Institutes of Health for a research traineeship award (GM-08501) to S.E.S., and the Howard Hughes Medical Institute for a predoctoral fellowship to E.E.B. We thank G. M. Hathaway for MALDI-TOF mass spectrometry.

JA971184T