Optimization of the Hairpin Polyamide Design for Recognition of the Minor Groove of DNA

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

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

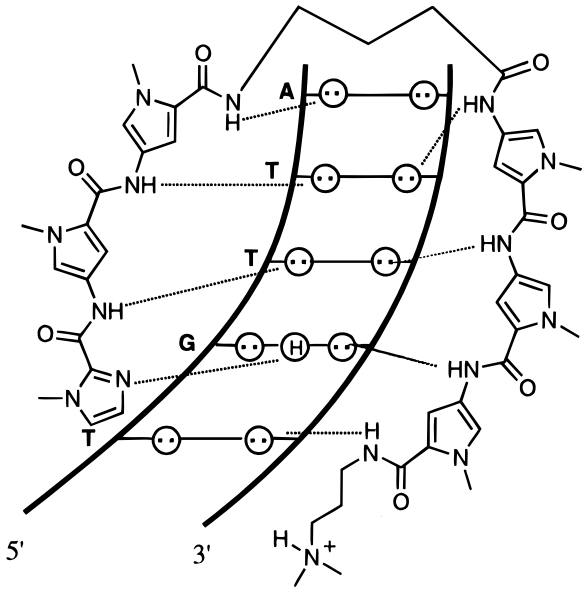
Received March 5, 1996[⊗]

Abstract: In order to optimize the hairpin design for ligands which bind the minor groove of DNA, a series of four pyrrole−imidazole polyamides substituted at the C-terminus with aliphatic amino acids have been prepared using solid phase synthetic methodology. Addition of a C-terminal β-alanine residue is found to enhance both the DNA binding affinity and sequence specificity, while addition of a C-terminal glycine residue is found to reduce DNA binding affinity and sequence specificity. These effects are modulated by the addition of an N-terminal acetyl group. Insertion of a C-terminal aliphatic amino acid residue makes the hairpin polyamide motif compatible with solid phase synthetic methods, allowing the rapid design of new polyamides for high-affinity specific recognition of a broad sequence repertoire in the minor groove of DNA.

Pyrrole−imidazole polyamide−DNA complexes combined with a convenient solid phase synthesis provide a new paradigm for the design of artificial molecules for the sequence specific recognition of double-helical DNA.[1,2] Polyamides containing *N*-methylimidazole (Im) and *N*-methylpyrrole (Py) amino acids can be combined in antiparallel side-by-side dimeric complexes with the minor groove of DNA.[1,3−5] The DNA sequence specificity of these small molecules can be controlled by the linear sequence of pyrrole and imidazole amino acids. An imidazole ring on one ligand complemented by a pyrrolecarboxamide ring on the second ligand recognizes a G·C base pair, while a pyrrolecarboxamide/ imidazole combination targets a C·G base pair.[1,4] A pyrrolecarboxamide/pyrrolecarboxamide pair is degenerate for A·T or T·A base pairs.[1,3−4]

Hairpin Polyamide. A simple polyamide hairpin motif with γ-aminobutyric acid (γ) serving as a turn monomer provides a synthetically accessible method of covalently linking polyamide units within the 2:1 motif (Figure 1).[6] Covalently linked polyamide heterodimers and homodimers have both increased affinities and sequence specificity.[6,7]

During the development of solid phase methods for the synthesis of the pyrrole−imidazole polyamides, it was necessary to identify an appropriate resin linkage. Because the reactivities of an imidazole aromatic ester, a pyrrole aromatic ester, and the ester of an aliphatic amino acid are different, separate cleavage chemistries would be necessary for polyamides with C-terminal imidazole and C-terminal pyrrole residues. Boc-β-alanine-Pam-resin has been reported to be cleaved in high yield by aminolysis.[8] Preloaded Boc-β-alanine-Pam-resin and Boc-



ImPyPy-γ-PyPyPy-Dp • TGTTA

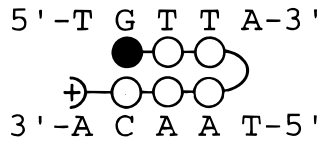


Figure 1. Top: Hairpin polyamide binding model for the complex formed between ImPyPy-γ-PyPyPy-Dp and a 5′-TGTTA-3′ sequence. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. Bottom: Schematic binding model; the imidazole and pyrrole rings are represented as shaded and unshaded circles, respectively.

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

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

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

(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.

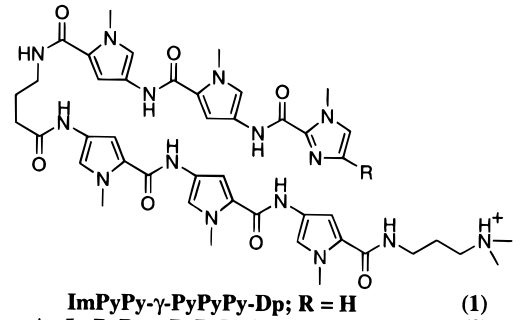
(4) (a) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572. (b) Geierstanger, B. H.; Jacobsen, J. P.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E.; *Biochemistry* **1994**, *33*, 3055. (c) Geierstanger, B. H.; Dwyer, T. J.; Bathini, Y.; Lown, J. W.; Wemmer, D.E. *J. Am. Chem. Soc.* **1993**,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

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

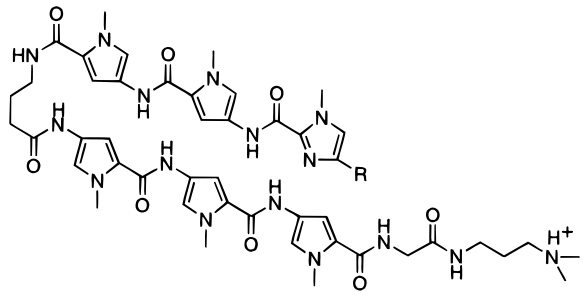
(6) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.

(7) (a) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 9892. (b) Dwyer, T. J.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 9900. (c) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 3663. (d) Chen, Y. H.; Lown, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 6995.

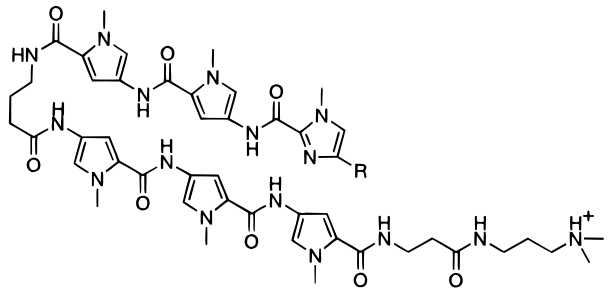
S0002-7863(96)00721-4 CCC: $12.00



ImPyPy-γ-PyPyPy-Dp; R = H (1)
Ac-ImPyPy-γ-PyPyPy-Dp; R = NHCOCH$_3$ (2)



ImPyPy-γ-PyPyPy-G-Dp; R = H (3)
Ac-ImPyPy-γ-PyPyPy-G-Dp; R = NHCOCH$_3$ (4)



ImPyPy-γ-PyPyPy-β-Dp; R = H (5)
Ac-ImPyPy-γ-PyPyPy-β-Dp; R = NHCOCH$_3$ (6)

Figure 2. Hairpin pyrrole−imidazole polyamides, with C-terminal Dp (**1** and **2**), G-Dp (**3** and **4**), and β-Dp (**5** and **6**) end groups.

glycine-Pam-resin are commercially available.[9] The addition of an aliphatic amino acid at the C-terminus of the pyrrole−imidazole polyamides provides a convenient alternative to developing new methods for linking and cleaving aromatic esters on a solid support.

A series of four pyrrole−imidazole polyamides containing either a C-terminal glycine or β-alanine residue were prepared for comparison of DNA binding properties (Figure 2). We report here that the polyamide synthesized from Boc-β-alanine-Pam-resin, ImPyPy-γ-PyPyPy-β-Dp (**5**), binds with both enhanced affinity and specificity relative to the parent compound, ImPyPy-γ-PyPyPy-Dp (**1**),[6] which lacks the C-terminal β-alanine residue. In addition, we find that N-terminal acetylation can alter the relative affinities of single base pair mismatches.

Identification of Binding Sites by MPE·Fe^{II} Footprinting. MPE·Fe^{II} footprinting[10] on the 3′- and 5′-^{32}P-end-labeled 135 base pair *Eco*RI/*Bsr*BI restriction fragments (25 mM Tris-acetate, 10 mM NaCl, 100 μM/base pair calf thymus DNA at pH 7.0 and 22 °C) reveals that the six polyamides **1**−**6** bind to

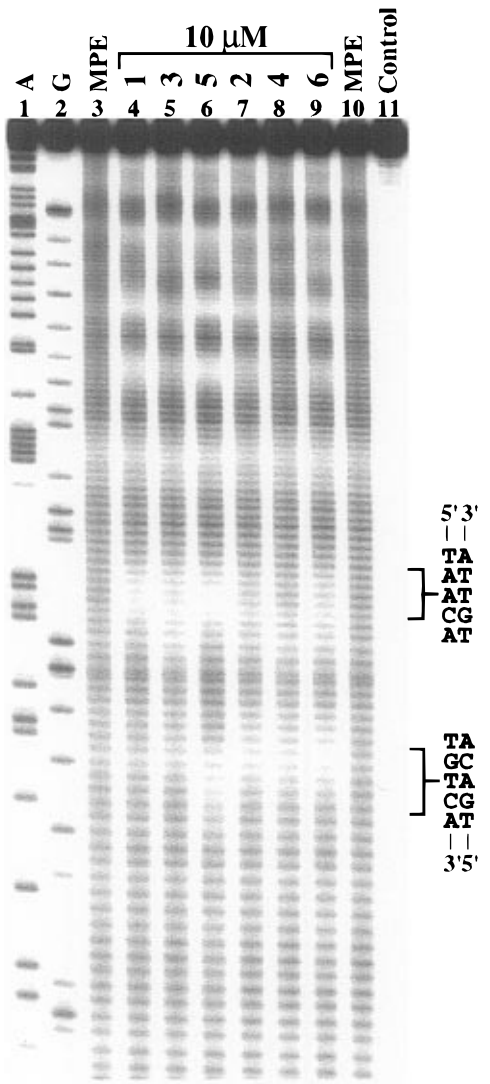


Figure 3. MPE·Fe^{II} footprinting of ImPyPy-γ-PyPyPy-Dp (**1**), ImPyPy-γ-PyPyPy-G-Dp (**3**), ImPyPy-γ-PyPyPy-β-Dp (**5**), AcImPyPy-γ-PyPyPy-Dp (**2**), AcImPyPy-γ-PyPyPy-G-Dp (**4**), and AcImPyPy-γ-PyPyPy-β-Dp (**6**) on a 135 base pair *Eco*RI/*Bsr*BI restriction fragment. All reactions contain 10 kcpm restriction fragment, 25 mM Tris-acetate, 10 mM NaCl, 100 μM calf thymus DNA (bp), and 5 mM DTT. Lane 1: A reaction. Lane 2: G reaction. Lanes 3 and 10: MPE·Fe^{II} standard. Lane 4: 10 μM **1**. Lane 5: 10 μM **3**. Lane 6: 10 μM **5**. Lane 7, 10 μM **2**. Lane 8: 10 μM **4**. Lane 9: 10 μM **6**. Lane 11: intact DNA.

the 5′-TGTTA-3′ match site (Figures 3 and 4). Remarkably, single base pair mismatch site preferences varied for the polyamides depending on the C- and N-terminal substitution: ImPyPy-γ-PyPyPy-Dp (**1**) and ImPyPy-γ-PyPyPy-β-Dp (**5**) bind a 5′-TGACA-3′ site, and AcImPyPy-γ-PyPyPy-Dp (**2**), ImPyPy-γ-PyPyPy-G-Dp (**3**), AcImPyPy-γ-PyPyPy-G-Dp (**4**), and AcImPyPy-γ-PyPyPy-β-Dp (**6**) bind to a 5′-ATTCG-3′ site.

Analysis of Energetics by Quantitative DNase I Footprint Titrations. Quantitative DNase I footprinting[11] on the 3′-^{32}P-end-labeled 135 base pair *Eco*RI/*Bsr*BI restriction fragment (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 and 22 °C) reveals the following association constants K_a for the polyamides studied: ImPyPy-γ-PyPyPy-β-Dp (**5**) ($K_a = 2.9 \times 10^8$ M^{-1}) > ImPyPy-γ-PyPyPy-Dp (**1**) ($K_a = 7.6 \times 10^7$ M^{-1}) > AcImPyPy-γ-PyPyPy-Dp (**2**) ($K_a = 6.4 \times 10^7$ M^{-1})

(8) Mitchell, A. R.; Kent, S. B. H.; Engelhard, M.; Merrifield, R. B. *J. Org. Chem.* **1978**, *43*, 2845.
(9) Available in bulk from Peptides International, Louisville, KY.
(10) (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.
(11) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132.

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

polyamide	match site 5′tTGTTAg-3′	single mismatch sites 5′-aTGACAt-3′	single mismatch sites 5′-cATTCGt-3′
ImPyPy-γ-PyPyPy-Dp	7.6×10^7 (0.8)	3.2×10^6 (0.7)	*c*
ImPyPy-γ-PyPyPy-G-Dp	3.3×10^6 (0.6)	*c*	1.0×10^6 (0.2)
ImPyPy-γ-PyPyPy-β-Dp	2.9×10^8 (0.5)	4.8×10^6 (1.1)	*c*
AcImPyPy-γ-PyPyPy-Dp	6.4×10^7 (3.0)	*c*	2.8×10^7 (0.5)
AcImPyPy-γ-PyPyPy-G-Dp	2.4×10^7 (2.2)	*c*	1.5×10^7 (1.2)
AcImPyPy-γ-PyPyPy-β-Dp	2.9×10^7 (0.5)	*c*	1.8×10^7 (0.2)

[a] Values reported are the mean values measured from at least three footprint titration experiments, with the standard deviation for each data set indicated in parentheses. [b] The assays were performed at 22 °C at pH 7.0 in the presence of 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. [c] The affinity for this site was not determined.

Figure 4. MPE·Fe^{II} protection patterns for ImPyPy-γ-PyPyPy-Dp (**1**), ImPyPy-γ-PyPyPy-G-Dp (**3**), ImPyPy-γ-PyPyPy-β-Dp (**5**), AcImPyPy-γ-PyPyPy-Dp (**2**), AcImPyPy-γ-PyPyPy-G-Dp (**4**), and AcImPyPy-γ-PyPyPy-β-Dp (**6**) on a 135 base pair *Eco*RI/*Bsr*BI restriction fragment. Bar heights are proportional to the protection from cleavage at each band. Boxes represent binding sites determined by the published model.[10]

> AcImPyPy-γ-PyPyPy-G-Dp (**4**) ($K_a = 4.0 \times 10^7$ M^{-1}) > AcImPyPy-γ-PyPyPy-β-Dp (**6**) ($K_a = 2.9 \times 10^7$ M^{-1}) > ImPyPy-γ-PyPyPy-G-Dp (**3**) ($K_a = 3.3 \times 10^6$ M^{-1}) (Table 1). Equilibrium association constants for the single base pair mismatch site 5′-TGACA-3′ for ImPyPy-γ-PyPyPy-Dp (**1**) and ImPyPy-γ-PyPyPy-β-Dp (**5**) are $K_a = 3.2 \times 10^6$ M^{-1} and $K_a = 4.8 \times 10^6$ M^{-1}, respectively. AcImPyPy-γ-PyPyPy-Dp (**2**), ImPyPy-γ-PyPyPy-G-Dp (**3**), AcImPyPy-γ-PyPyPy-G-Dp (**4**), and AcImPyPy-γ-PyPyPy-β-Dp (**6**) bind to 5′-ATTCG-3′ with association constants of $K_a = 2.8 \times 10^7$ M^{-1}, 1.0×10^6 M^{-1}, 1.2×10^7 M^{-1}, and 1.8×10^7 M^{-1}, respectively (Figures 5 and 6).

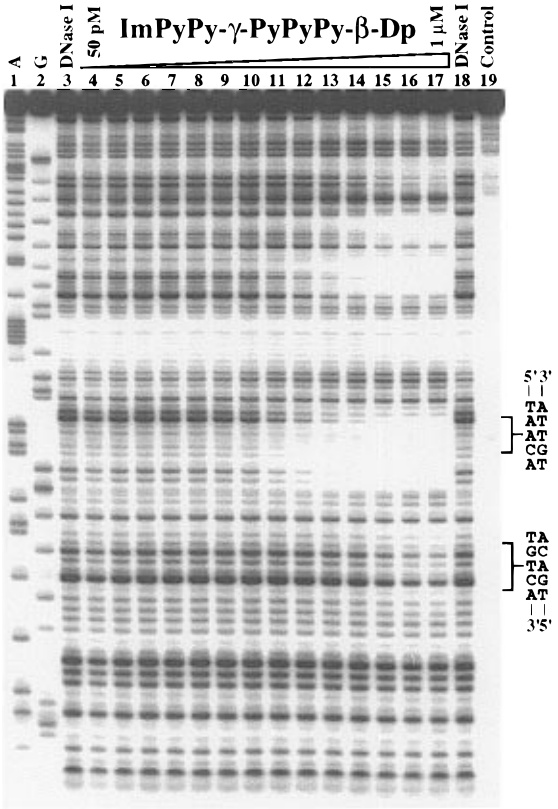


Figure 5. Quantitative DNase I footprint titration experiment with ImPyPy-γ-PyPyPy-β-Dp (**5**) on the 3′-^{32}P-labeled 135 base pair *Eco*RI/*Bsr*BI restriction fragment from plasmid pMM5. Lane 1: A reaction. Lane 2: G reaction. Lanes 3 and 18: DNase I standard. Lanes 4−17: 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, 500 nM, 1 μM ImPyPy-γ-PyPyPy-β-Dp, respectively. Lane 19: intact DNA. The 5′-TGTTA-3′ and 5′-TGACA-3′ binding sites which were analyzed are shown on the right. All reactions contain 10 kcpm restriction fragment, 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

Discussion

Binding Affinities. All six polyamides of core sequence composition ImPyPy-γ-PyPyPy, but varying at the N- and C-terminus, bind the designated target site 5′-TGTTA-3′ as its high-affinity site, but the overall equilibrium association constants vary by 2 orders of magnitude. Among the six polyamides, ImPyPy-γ-PyPyPy-β-Dp (**5**) binds the targeted 5′-TGTTA-3′ site with the highest affinity.[12] This suggests that addition of a C-terminal linker β-alanine residue to facilitate solid phase polyamide synthesis is not merely an acceptable strategy but, serendipitously, *designates an optimized hairpin polyamide*. ImPyPy-γ-PyPyPy-β-Dp (**5**) binds with an associa-

(12) Affinity cleaving experiments reveal that ImPyPy-γ-PyPyPy-β-Dp binds the 5′-TGTTA-3′ target site with a single orientation.[13]

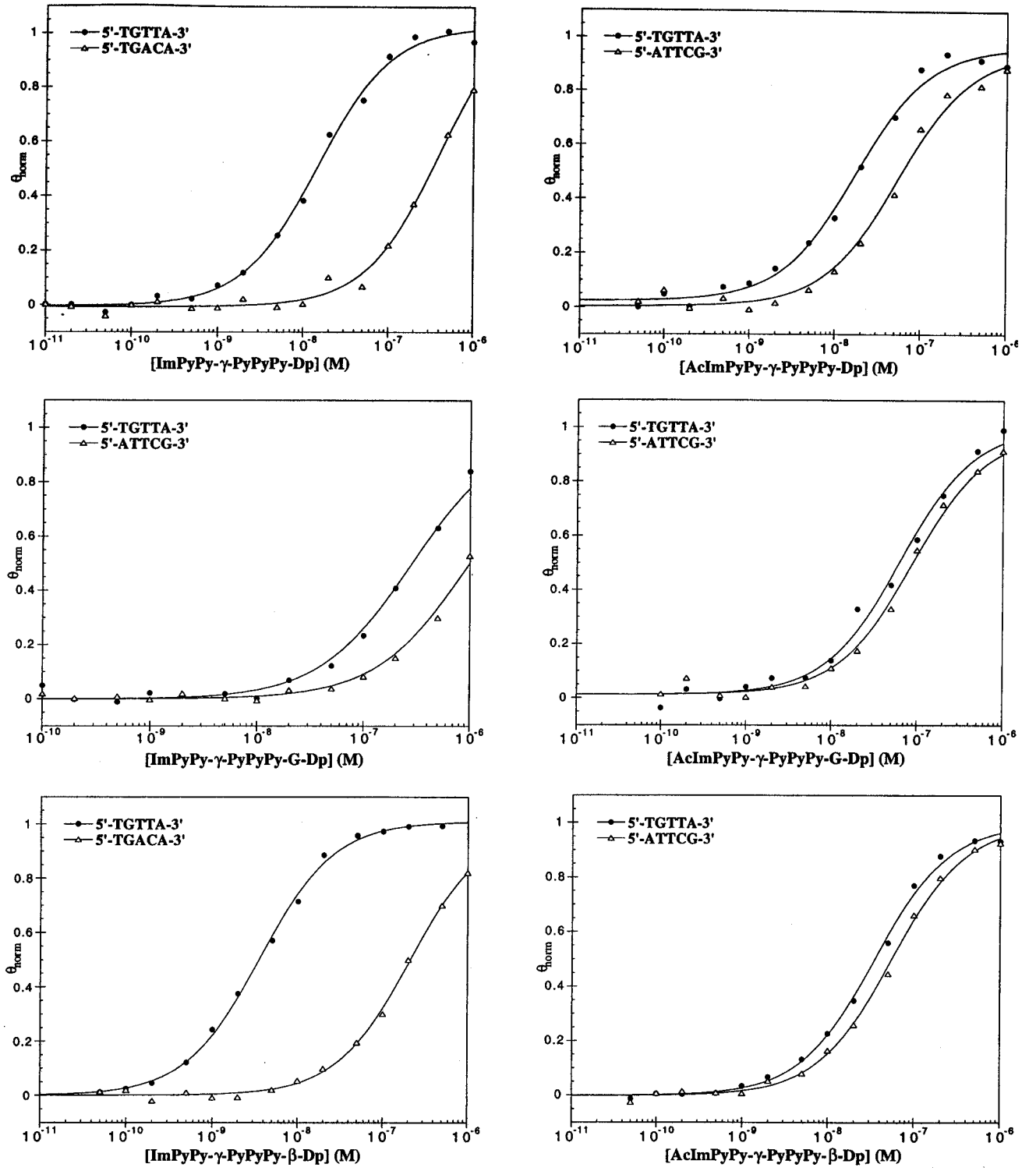


Figure 6. Data for the quantitative DNase I footprint titration experiments for the six polyamides in complex with the designated 5′-TGTTA-3′ target site. The θ_{norm} points were obtained using photostimulable storage phosphor autoradiography and processed as described in the Experimental Section. The data points for ImPyPy-γ-PyPyPy-Dp (**1**), ImPyPy-γ-PyPyPy-G-Dp (**3**), ImPyPy-γ-PyPyPy-β-Dp (**5**), AcImPyPy-γ-PyPyPy-Dp (**2**), AcImPyPy-γ-PyPyPy-G-Dp (**4**), and AcImPyPy-γ-PyPyPy-β-Dp (**6**) are indicated by filled circles for the 5-TGTTA-3′ match site, and open triangles for the 5-TGACA-3′ or 5′-ATTCG-3′ single base pair mismatch sites. The solid curves are the best-fit Langmuir binding titration isotherms obtained from nonlinear least squares algorithm using eq 2.

tion constant, $K_a = 3 \times 10^8\ M^{-1}$, a factor of 4 greater than that of the parent polyamide, ImPyPy-γ-PyPyPy-Dp (**1**), $K_a = 8 \times 10^7\ M^{-1}$. A C-terminal glycine residue *reduces* binding affinity at the 5′-TGTTA-3′ match site by a factor of 88 for ImPyPy-γ-PyPyPy-G-Dp (**3**) binding relative to ImPyPy-γ-PyPyPy-β-Dp (**5**). The glycine residue may create an unfavorable steric clash placing the glycine carbonyl toward the floor of the minor groove. A high-resolution NMR study of the 2:1 polyamide−DNA complex for ImPyPy-G-PyPyPy-Dp indicates considerable distortion of the polyamides to avoid placing the glycine carbonyl in the floor of the minor groove.[14] In contrast to glycine, a C- terminal β-alanine residue presents a steric surface which resembles that of the C-terminus of the original hairpin

(13) White, S.; Baird, E. E.; Dervan, P. B. *Biochemistry*, submitted.
(14) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Nature Struct. Biol.* **1996**, *3*, 321.

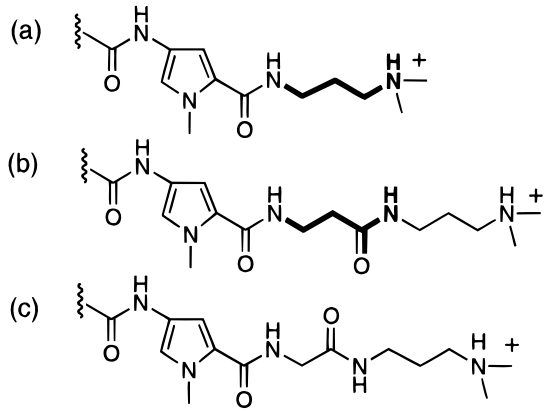


Figure 7. Modification of the C-terminus of pyrrole−imidazole polyamides: (a) Dp, (b) G-Dp, (c) β-Dp.

polyamide (Figure 7). The modest increased binding affinity of the C-terminal β-alanine polyamide may result from an additional hydrogen bond between the β-alanine carboxamide and a "sixth" base pair of the binding site. Evidence for such an interaction must await further structural studies.

Specificity. Although the pyrrole−imidazole polyamides can tolerate a variety of substitutions at the termini with regard to match site binding affinity, the overall specificity for mismatch sites is significantly modulated. The most specific was ImPyPy-γ-PyPyPy-β-Dp (**5**), which binds the target 5′-TGTTA-3′ match site with 60-fold specificity relative to a single base pair 5′-TGACA-3′ mismatch site. This can be compared with the parent polyamide ImPyPy-γ-PyPyPy-Dp (**1**), which has a 24-fold specific binding relative to the same two DNA sites.[6] An N-terminal acetyl group or a C-terminal glycine group reduces the observed relative sequence preference to 1- and 3-fold. Furthermore, MPE·Fe^{II} footprinting reveals the sequence of the mismatch binding site changes from 5′-TGACA-3′ to 5′-ATTCG-3′ (Figure 4). For example, the acetylated polyamide, Ac-ImPyPy-γ-PyPyPy-β-Dp (**6**), binds the 5′-TGTTA-3′ match site with 10-fold reduced affinity compared to ImPyPy-γ-PyPyPy-β-Dp (**5**), while binding a 5′-ATTCG-3′ mismatch site with a 1.6-fold *increase* in affinity relative to the original 5′-TGACA-3′ mismatch site. Binding at the 5′-TGACA-3′ single base pair mismatch is greatly reduced by introducing the acetyl group, or C-terminal glycine group in all cases. The steric bulk of the acetyl group or glycine residue may simply be preventing the polyamide from sitting deeply in the minor groove, reducing the affinity for the match, while increasing the tolerance at the mismatch. Alternatively, there may be a favorable interaction between the acetyl group or glycine and the terminal G·C base pair of the 5′-ATTCG-3′ mismatch site. Structural studies will be necessary to describe this interaction.

Implications for the Design of Minor Groove Binding Molecules. Six hairpin polyamides of similar core sequence composition bind a common designated 5′-TGTTA-3′ target site, but substitution at the N- and C-termini modulates specificity for mismatch sites. The addition of a C-terminal β-alanine residue is compatible with solid phase methods and enhances both the affinity and sequence specificity of the pyrrole−imidazole hairpin. This result sets the stage for the elucidation of the binding-site size limits of the hairpin motif.

Experimental Section

Synthesis of Polyamides. ImPyPy-γ-PyPyPy-Dp (**1**) and AcImPyPy-γ-PyPyPy-Dp (**2**) were prepared as previously described.[6] Four polyamides, ImPyPy-γ-PyPyPy-G-Dp (**3**), AcImPyPy-γ-PyPyPy-G-Dp (**4**), ImPyPy-γ-PyPyPy-β-Dp (**5**), and AcImPyPy-γ-PyPyPy-β-Dp (**6**), containing C-terminal amino acids were prepared by solid phase methodology.[2] All polyamides were found to be soluble in aqueous solution up to at least 1 mM concentration.

ImPyPy-γ-PyPyPy-G-Dp (3): (12 mg, 40% recovery); HPLC, room temperature 26.9; UV λ_{max} (H_2O) 246 (41 100), 306 (51 300) ^{1}H NMR (DMSO-d_6) δ 10.50 (s, 1 H), 9.95 (s, 1 H), 9.93 (s, 1 H), 9.92 (s, 1 H), 9.86 (s, 1 H), 9.2 (br s, 1 H), 8.29 (t, 1 H, J = 4.4 Hz), 8.07 (t, 1 H, J = 5.2 Hz), 8.03 (t, 1 H, J = 5.4 Hz), 7.39 (s, 1 H), 7.27, (d, 1 H, J = 1.6 Hz), 7.22 (m, 2 H), 7.16 (m, 2 H), 7.04 (m, 2 H), 6.92 (d, 1 H, J = 1.6 Hz), 6.89 (d, 1 H, J = 1.7 Hz), 6.86 (d, 1 H, J = 1.6 Hz), 3.97 (s, 3 H), 3.82 (m, 6 H), 3.81 (s, 3 H), 3.78 (m, 6 H), 3.70 (d, 2 H, J = 5.7 Hz), 3.20 (q, 2 H, J = 5.7), 3.11 (q, 2 H, J = 4.2 Hz), 3.00 (q, 2 H, J = 4.4 Hz), 2.76 (d, 6 H, J = 4.7 Hz), 2.24 (t, 2 H, J = 4.8 Hz), 1.77 (m, 4 H); MALDI-TOF-MS, 964.3 (964.1 calcd for M + H).

AcImPyPy-γ-PyPyPy-G-Dp (4): 13.1 mg, 30% yield; HPLC, room temperature 24.0, UV λ_{max} (H_2O) 246 (35 900), 312 (48 800); ^{1}H NMR (DMSO-d_6) δ 10.23 (s, 1 H), 9.98 (s, 1 H), 9.32 (s, 1 H), 9.90 (m, 2 H), 9.84 (s, 1 H), 9.2 (br s, 1 H), 8.27 (t, 1 H, J = 5.0 Hz), 8.05 (m, 2 H), 7.41 (s, 1 H), 7.25 (d, 1 H, J = 1.4 Hz), 7.22 (m, 2 H), 7.16 (m, 2 H), 7.12 (d, 1 H, J = 1.7 Hz), 7.05 (d, 1 H, J = 1.5 Hz), 6.94 (d, 1 H, J = 1.6 Hz), 6.89 (d, 1 H, J = 1.7 Hz) 6.87 (d, 1 H, J = 1.6 Hz), 3.93 (s, 3 H), 3.83 (s, 3 H), 3.82 (m, 6 H), 3.81 (s, 3 H), 3.79 (s, 3 H), 3.71 (d, 2 H, J = 5.1 Hz), 3.19 (q, 2 H, J = 5.8 Hz), 3.12 (q, 2 H, J = 5.0 Hz), 3.01 (q, 2 H, J = 4.2 Hz), 2.74 (d, 6 H, J = 4.6 Hz), 2.26 (t 2 H, J = 4.6 Hz), 2.00 (s, 3 H), 1.75 (m, 4 H); MALDI-TOF-MS, 1021.6 (1021.1 calcd for M + H).

AcImPyPy-γ-PyPyPy-β-Dp (6): 9.2 mg, 31% yield; UV λ_{max} (H_2O) 246 (42 800), 312 (50 400); HPLC, room temperature 24.9; ^{1}H NMR (DMSO-d_6) δ 10.25 (s, 1 H), 10.01 (s, 1 H), 9.92 (m, 3 H), 9.86 (s, 1 H), 9.3 (br s, 1 H), 8.10 (m, 3 H), 7.42 (s, 1 H), 7.25 (d, 1 H, J = 1.5 Hz), 7.20 (d, 1 H, J = 1.6 Hz), 7.16 (m, 3 H), 7.12 (d, 1 H, J = 1.4 Hz), 7.03 (d, 1 H J = 1.7), 6.89 (d, 1 H, J = 1.6 Hz), 6.86 (m, 2 H), 3.92 (s, 3 H), 3.83 (s, 3 H), 3.82 (s, 3 H), 3.80 (s, 6 H), 3.78 (s, 3 H), 3.35 (q, 2 H, J = 5.5 Hz), 3.20 (q, 2 H, J = 3.8 Hz), 3.08 (q, 2 H, J = 3.3 Hz), 2.97 (q, 2 H, J = 3.8 Hz), 2.75 (d, 6 H, J = 4.8 Hz), 2.34 (t, 2 H, J = 5.0 Hz), 2.24 (t, 2 H, J = 4.4 Hz), 2.00 (s, 3 H), 1.71 (m, 4 H); MALDI-TOF-MS, 1035.4 (1035.1 calcd for M + H).

DNA Reagents and Materials. Sonicated, deproteinized calf thymus DNA was obtained from Pharmacia. Enzymes were purchased from Boehringer-Mannheim and used with the buffers supplied. Deoxyadenosine 5′-[α-^{32}P]triphosphate, thymidine 5′-[α-^{32}P]triphosphate, and adenosine 5′-[γ-^{32}P]triphosphate were obtained from Amersham. Storage phosphor autoradiography was performed using a Molecular Dynamics 400S PhosphorImager and ImageQuant software. The 135 base pair 3′- and 5′-end-labeled *Eco*RI/*Bsr*BI restriction fragments from plasmid pMM5 were prepared and purified as follows. Plasmid DNA was linearized using *Eco*RI, followed by treatment with either Klenow, deoxyadenosine 5′-[α-^{32}P]triphosphate, and thymidine 5′-[α-^{32}P]triphosphate for 3′ labeling or calf alkaline phosphatase and subsequent 5′ end labeling with T4 polynucleotide kinase and γ-[^{32}P]-dATP. The linearized plasmid DNA was digested with *Bsr*BI, and the 135 base pair *Eco*RI/*Bsr*BI restriction fragment was isolated by nondenaturing 5% polyacrylamide gel electrophoresis (PAGE). The gel bands were visualized by autoradiography, isolated, and filtered to remove the polyacrylamide. The resulting solution was further purified by phenol extraction followed by ethanol precipitation. Chemical sequencing reactions were performed according to published protocols.[15] Standard protocols were used for all DNA manipulations.[16]

MPE·Fe^{II} Footprinting. All reactions were carried out in a total volume of 40 μL with final concentrations of species as indicated in parentheses. The ligands were added to solutions of radiolabeled restriction fragment (10 000 cpm), calf thymus DNA (100 μM bp), Tris-acetate (25 mM, pH 7.0), and NaCl (10 mM) and incubated for 1 h at 22 °C. A 50 μM MPE·Fe^{II} solution was prepared by mixing 100 μL of a 100 μM MPE solution with a freshly prepared 100 μM ferrous ammonium sulfate solution. Footprinting reactions were initiated by the addition of MPE·Fe^{II} (5 μM), followed 5 min later by the addition of dithiothreitol (5 mM), and allowed to proceed for 15 min at 22 °C. Reactions were stopped by ethanol precipitation, resuspended in 100

(15) (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.

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

mM Tris-borate-EDTA/80% formamide loading buffer, and electrophoresed on 8% polyacrylamide denaturing gels (5% cross-link, 7 M urea) at 2000 V for 1 h. The gels were analyzed using storage phosphor technology.

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

Equilibrium association constants were determined as previously described.[6,11] The data were analyzed by performing volume integrations of the 5′-TGTTA-3′, 5′-TGACA-3′, and 5′-ATTCG-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:

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

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

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

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

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

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

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

JA960721R