Extension of Sequence-Specific Recognition in the Minor Groove of DNA by Pyrrole–Imidazole Polyamides to 9–13 Base Pairs

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · JULY 1996

Impact Factor: 12.11 · DOI: 10.1021/ja960726o

4 AUTHORS, INCLUDING:

John Wesley Trauger
Novartis
32 PUBLICATIONS **2,359** CITATIONS

SEE PROFILE

Extension of Sequence-Specific Recognition in the Minor Groove of DNA by Pyrrole−Imidazole Polyamides to 9−13 Base Pairs

John W. Trauger, Eldon E. Baird, Milan Mrksich, and Peter B. Dervan*

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

Received March 6, 1996[⊗]

Abstract: The sequence-specific recognition of the minor groove of DNA by pyrrole−imidazole polyamides has been extended to 9−13 base pairs (bp). Four polyamides, ImPyPy-Py-PyPyPy-Dp, ImPyPy-G-PyPyPy-Dp, ImPyPy-β-PyPyPy-Dp, and ImPyPy-γ-PyPyPy-Dp (Im = *N*-methylimidazole, Py = *N*-methylpyrrole, Dp = *N*,*N*-dimethylaminopropylamide, G = glycine, β = β-alanine, and γ = γ-aminobutyric acid), were synthesized and characterized with respect to their DNA-binding affinities and specificities at sequences of composition 5′-(A,T)G-(A,T)$_5$C(A,T)-3′ (9 bp) and 5′-(A,T)$_5$G(A,T)C(A,T)$_5$-3′ (13 bp). In both sequence contexts, the β-alanine-linked compound ImPyPy-β-PyPyPy-Dp has the highest binding affinity of the four polyamides, binding the 9 bp site 5′-TGTTAAACA-3′ ($K_a = 8 \times 10^8$ M^{-1}) and the 13 bp site 5′-AAAAAGACAAAAA-3′ ($K_a = 5 \times 10^9$ M^{-1}) with affinities higher than the formally *N*-methylpyrrole-linked polyamide ImPyPy-Py-PyPyPy-Dp by factors of ∼8 and ∼85, respectively (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$, pH 7.0). The binding data for ImPyPy-γ-PyPyPy-Dp, which has been shown previously to bind DNA in a "hairpin" conformation, indicates that γ-aminobutyric acid does not effectively link polyamide subunits in an extended conformation. These results expand the binding site size targetable with pyrrole−imidazole polyamides and provide structural elements that will facilitate the design of new polyamides targeted to other DNA sequences.

Introduction

The development of 2:1 pyrrole−imidazole polyamide−DNA complexes provides a new model for the design of molecules for sequence-specific recognition in the minor groove of DNA. The polyamide ImPyPy-Dp was shown to specifically bind the mixed A,T/G,C sequence 5′-(A,T)G(A,T)C(A,T)-3′ as a side-by-side antiparallel dimer.[1] In this complex, each polyamide makes specific contacts with one strand on the floor of the minor groove such that the sequence specificity depends on the sequence of side-by-side amino acid pairings. A side-by-side pairing of imidazole opposite pyrrole recognizes G·C base pairs, while a side-by-side pairing of pyrrole opposite imidazole recognizes C·G base pairs.[1] A side-by-side pyrrole−pyrrole pairing is partially degenerate and targets both A·T and T·A base pairs.[1,2] The generality of the 2:1 model has been demonstrated by targeting other sequences of mixed A,T/G,C composition.[3−5] ImPyPy-Dp and distamycin (PyPyPy) bind simultaneously to a 5′-(A,T)G(A,T)$_3$-3′ site as an antiparallel heterodimer.[3] A PyImPy polyamide and distamycin bind 5′-(A,T)$_2$G(A,T)$_2$-3′ [4] and ImPyImPy-Dp targets 5′-(A,T)GCGC-(A,T)-3′.[5]

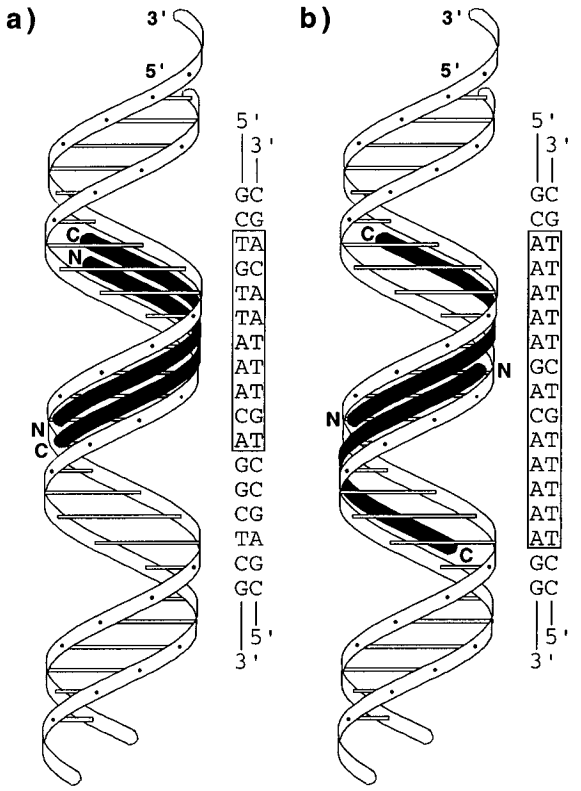

Figure 1. Ribbon models of (a) 9 bp "overlapped" and (b) 13 bp "slipped" 2:1 polyamide−DNA complexes.

The binding affinity and sequence specificity of a noncovalent antiparallel homodimeric or heterodimeric polyamide−DNA

[⊗] 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.; Geirstanger, 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) (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.; Sundaralingam, M. *Nature Struct. Biol.* **1994**, *1*, 169.

(3) (a) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 2572−2576. (b) Geierstanger, B. H.; Jacobsen, J.-P.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Biochemistry* **1994**, *33*, 3055.

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

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

S0002-7863(96)00726-3 CCC: $12.00

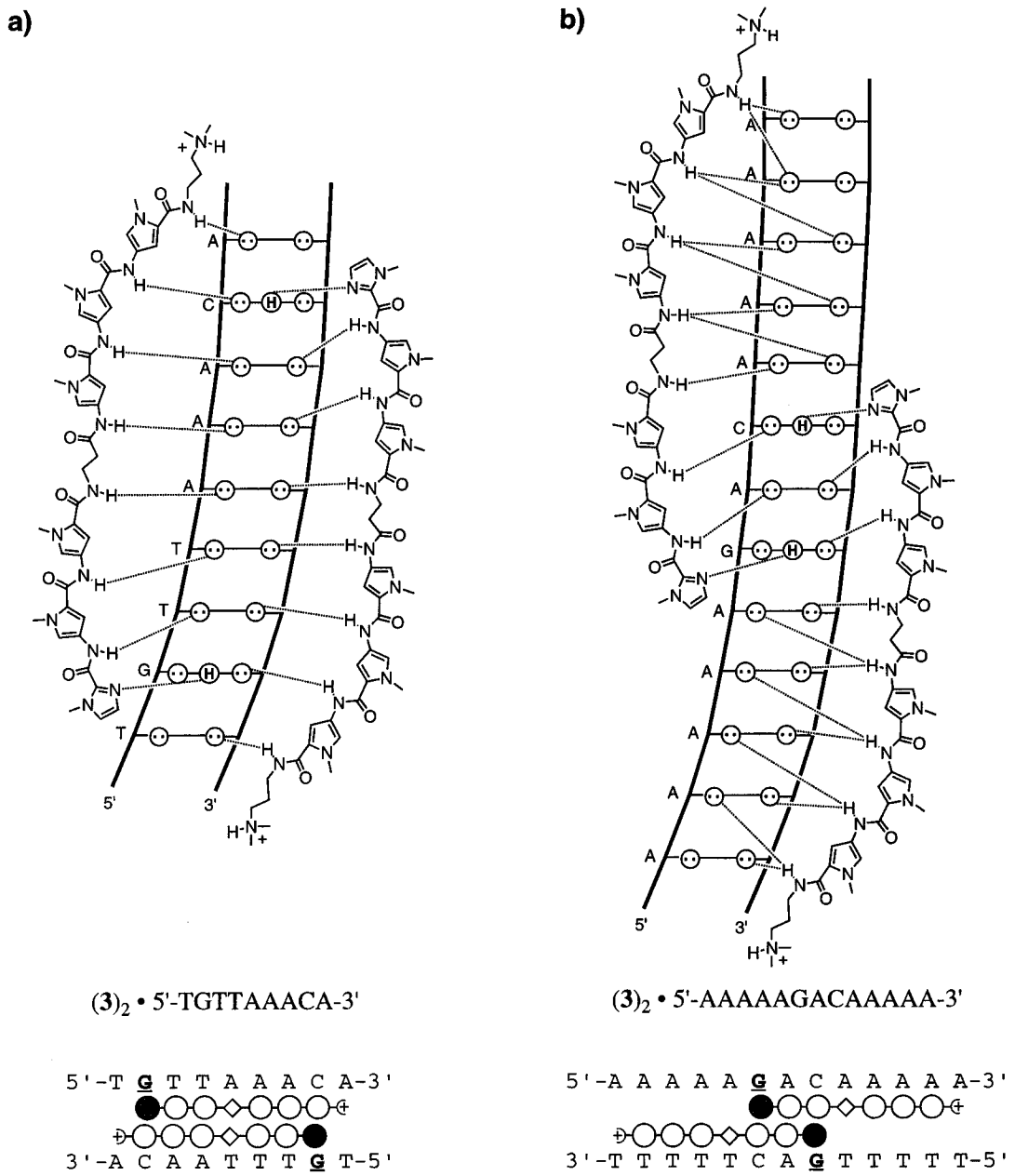

Figure 2. (Top) Complexes of ImPyPy-β-PyPyPy-Dp with the targeted sites (a) 5′-TGTTAAACA-3′ (9 bp "overlapped") and (b) 5′-AAAAAGACAAAAA-3′ (13 bp, "slipped"). Circles with dots represent lone pairs on N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dashed lines. (Bottom) Complexes of ImPyPy-X-PyPyPy-Dp, where X = Py, G, and β, with (a) 5′-TGTTAAACA-3′ and (b) 5′-AAAAAGACAAAAA-3′. The shaded and light circles represent imidazole and pyrrole rings, respectively, and the diamond represents the internal amino acid X. The specifically targeted guanines are highlighted.

complex can be increased by covalently linking the two polyamides.[6,7] The DNA-binding properties of the polyamides ImPyPy-G-PyPyPy-Dp, ImPyPy-β-PyPyPy-Dp, and ImPyPy-γ-PyPyPy-Dp, in which the terminal carboxyl group of ImPyPy and the terminal amine of PyPyPy-Dp are connected with glycine (G), β-alanine (β), and γ-aminobutyric acid (γ), respectively, were recently reported.[7,8] The γ-aminobutyric acid-linked polyamide bound the designated target site 5′-TGTTA-3′ with high affinity and sequence specificity and exhibited a binding isotherm in quantitative footprinting experiments consistent with formation of an intramolecular "hairpin" complex in which the polyamide folds back on itself.[7] Modeling suggested that the glycine- and β-alanine-linked polyamides do not favorably bind as "hairpins" in the minor groove of DNA. Moreover, these polyamides exhibited cooperative binding isotherms in quantitative footprinting experiments, consistent with two polyamides binding in extended conformations as *intermolecular dimers*.[7,8] It appears that the glycine- and β-alanine-linked polyamides disfavor binding in the hairpin conformation and prefer to bind in an extended conformation.[7,8] In a formal sense, there are multiple extended binding motifs (and hence multiple binding site sequences) for polyamides of sequence composition Im(Py)$_x$-Dp, as discussed below.

"Overlapped" and "Slipped" Binding Modes. We report here the DNA-binding affinities of four polyamides having the

(6) (a) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 9892−9899. (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) Singh, M. P.; Plouvier, B.; Hill, G. C.; Gueck, J.; Pon, R. T.; Lown, J. W. *J. Am. Chem. Soc.* **1994**, *116*, 2006.

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

(8) Mrksich, M. Ph.D. Thesis, California Institute of Technology, 1994.

general sequence ImPyPy-X-PyPyPy-Dp, where X = Py, G, β, or γ, to the 9 bp site 5′-TGTTAAACA-3′ and to the 13 bp sites 5′-AAAAAGACAAAAA-3′ and 5′-ATATAGACATATA-3′. The polyamides having internal *N*-methylpyrrole, glycine, and β-alanine residues were anticipated to bind the 9 and 13 bp sites in an extended conformation. It was not clear at the outset if the γ-aminobutyric acid-linked polyamide ImPyPy-γ-PyPyPy-Dp would bind exclusively in an extended or "hairpin" conformation to the targeted sites. For ImPyPy-X-PyPyPy-Dp polyamides binding in an extended conformation, the polyamide−DNA complexes expected to form at the 9 bp and 13 bp target sites represent two distinct binding modes, which we refer to as "overlapped" and "slipped", respectively. In the "overlapped" (9 bp) binding mode, two ImPyPy-X-PyPyPy-Dp polyamides bind directly opposite one another (Figures 1a and 2a). The "slipped" (13 bp) binding mode integrates the 2:1 and 1:1 polyamide−DNA binding motifs at a single site. In this binding mode, the ImPyPy moieties of two ImPyPy-X-PyPyPy-Dp polyamides bind the central 5′-AGACA-3′ sequence in a 2:1 manner as in the ImPyPy homodimer,[1] and the PyPyPy moieties of the polyamides bind the all-A,T flanking sequences as in the 1:1 complexes of distamycin (Figures 1b and 2b). The structure of the complex formed by the polyamide ImPyPy-G-PyPyPy-Dp with a 13 bp target site has been characterized by 2D NMR.[9]

In the 9 bp "overlapped" and 13 bp "slipped" binding sites described above, the G·C and C·G base pairs are separated by one and five A,T base pairs, respectively. While we have concentrated here on these sites, we note that "partially slipped" sites of 10, 11, and 12 bp in which the G·C and C·G base pairs are separated by two, three, and four A,T base pairs, respectively, are also potential binding sites of the polyamides studied here.

Studies of the energetics of distamycin binding have shown that while the binding affinities are similar for complexation to poly[d(A-T)]·poly[d(A-T)] and poly[d(A)]·poly[d(T)], the origins of these binding affinities are different.[10] Binding to the alternating copolymer is enthalpy driven, while binding to the homopolymer is entropy driven.[10] However, not all 5 bp sites $(A,T)_5$ are bound with equal affinity. By quantitative footprinting experiments, the distamycin analog Ac-PyPyPy-Dp (Ac = acetyl) was shown to bind the sites 5′-AATAA-3′ and 5′-TTAAT-3′ with 2-fold and 14-fold lower affinity, respectively, relative to the site 5′-AAAAA-3′.[1c] On the basis of this result, we anticipated that the 13 bp site 5′-AAAAAGACAAAAA-3′ may be bound with higher affinity than the 13 bp site 5′-ATATAGACATATA-3′.

The binding affinities of polyamides **1**−**4** (Figure 3) for the three targeted sites 5′-TGTTAAACA-3′, 5′-AAAAAGACAAAAA-3′, and 5′-ATATAGACATATA-3′ were determined by quantitative DNase I footprint titration experiments.

Results

Synthesis of Polyamides. The polyamides ImPyPy-Py-PyPyPy-Dp (**1**),[11] ImPyPy-G-PyPyPy-Dp (**2**),[7] ImPyPy-β-PyPyPy-Dp (**3**),[7] and ImPyPy-γ-PyPyPy-Dp (**4**)[7] were prepared as described previously.

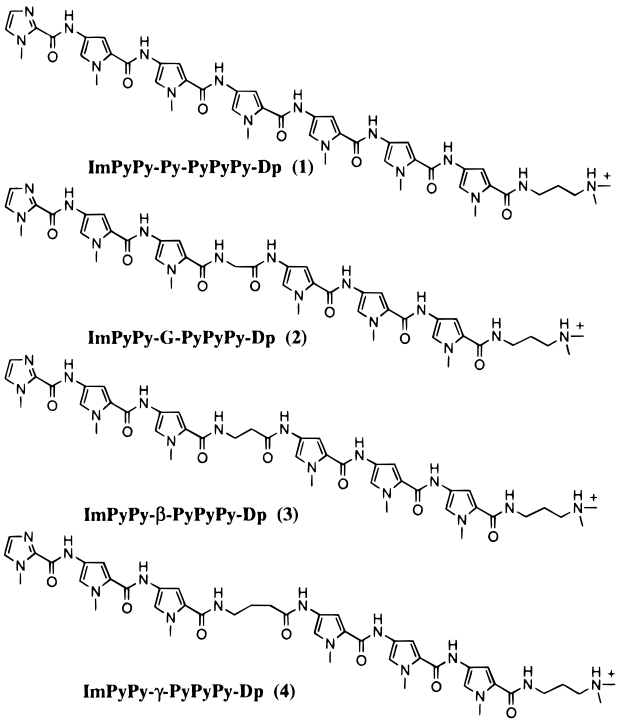

Figure 3. Structures of the four polyamides ImPyPy-X-PyPyPy-Dp, where X = *N*-methylpyrrole (Py), glycine (G), β-alanine (β), and γ-aminobutyric acid (γ).

Footprinting. Quantitative DNase I footprinting[12,13] on the 3′-^{32}P-labeled 281 bp pJT3 *Afl*II/*Fsp*I restriction fragment (Figures 4 and 5) (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, 22 °C) reveals that, of the four polyamides ImPyPy-X-PyPyPy-Dp, three (X = Py, G, β) bind to both the 9 bp "overlapped" site 5′-TGTTAAACA-3′ and the 13 bp "slipped" site 5′-AAAAAGACAAAAA-3′ with equilibrium association constants (K_a) greater than 5×10^7 M^{-1} (Table 1) and display cooperative binding isotherms (eq 2, $n = 2$) at these sites *consistent with binding as intermolecular dimers* (Figure 6). The fact that the polyamides ImPyPy-G-PyPyPy-Dp and ImPyPy-β-PyPyPy-Dp bind in the 9 bp "overlapped" binding mode indicates that the internal glycine and β-alanine amino acids are accommodated opposite a second ligand in a 2:1 polyamide−DNA complex.

The polyamide ImPyPy-γ-PyPyPy-Dp binds the site 5′-TGTTAAACA-3′ with an equilibrium association constant K_a 1×10^8 M^{-1} and also binds the site 5′-AAAAAGACAAAAA-3′ with lower affinity ($K_a = 6 \times 10^6$ M^{-1}). This compound displays binding isotherms (eq 2, $n = 1$) at these sites (Figure 6), *consistent with binding as an intramolecular hairpin* to the 5 bp "matched" sites 5′-TGTTA-3′ and 5′-AAACA-3′ (Figure 7) and to the 5 bp "single base pair mismatch" site 5′-AGACA-3′.[7] Significantly, it appears from these results that ImPyPy-γ-PyPyPy-Dp does not effectively link polyamide subunits in an extended conformation.

Binding Affinities. Comparison of the binding affinities of the four polyamides ImPyPy-X-PyPyPy-Dp, where X = Py, G, β, and γ, reveals that the internal amino acid X has a dramatic effect on complex stabilities (Table 1, Figure 6). The formally *N*-methylpyrrole-linked polyamide ImPyPy-Py-PyPyPy-Dp binds

(9) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Nature Struct. Biol.* **1996**, *3*, 321.

(10) Breslauer, K. J.; Remeta, D. P.; Chou, W.-Y.; Ferrante, R.; Curry, J.; Zaunczkowski, D.; Snyder, J. G.; Marky, L. A. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8922. (b) Marky, L. A.; Breslauer, K. J. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 4359. (c) Marky, L. A.; Kupke, K. J. *Biochemistry* **1989**, *28*, 9982.

(11) Kelly, J. J.; Baird, E. E.; Dervan, P. B., *Proc. Natl. Acad. Sci. U.S.A.*, in press.

(12) (a) Galas, D.; Schmitz, A. *Nucleic Acids Res.* **1978**, *5*, 3157. (b) Fox, K. R.; Waring, M. J. *Nucleic Acids Res.* **1984**, *12*, 9271.

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

Figure 4. Sequence of the 281 bp pJT3 *Afl*II/*Fsp*I restriction fragment. The three binding sites that were analyzed in quantitative footprint titration experiments are indicated.

Table 1. Association Constants (M^{-1}) for Polyamides ImPyPy-X-PyPyPy-Dp, Where X = *N*-Methylpyrrole (Py), Glycine (G), β-Alanine (β), and γ-Aminobutyric Acid (γ)[a,b]

	polyamide			
binding site	ImPyPy-Py-PyPyPy-Dp	ImPyPy-G-PyPyPy-Dp	ImPyPy-β-PyPyPy-Dp	ImPyPy-γ-PyPyPy-Dp
5′-TGTTAAACA-3′	9.7 ($\pm$2.3) $\times$ 10^7	1.4 ($\pm$0.1) $\times$ 10^8	7.8 ($\pm$0.6) $\times$ 10^8	1.4 ($\pm$0.3) $\times$ 10^8
5′-AAAAAGACAAAAA-3′	5.4 ($\pm$1.5) $\times$ 10^7	1.1 ($\pm$0.1) $\times$ 10^8	$\geq$4.7 ($\pm$0.7) $\times$ 10^9	6.4 ($\pm$0.6) $\times$ 10^6
5′-ATATAGACATATA-3′	3.6 ($\pm$0.5) $\times$ 10^7	6.6 ($\pm$0.4) $\times$ 10^6	1.0 ($\pm$0.1) $\times$ 10^9	4.6 ($\pm$0.5) $\times$ 10^6

[a] The reported association constants are the mean values obtained from three DNase I footprint titration experiments. The standard deviation for each value is indicated in parentheses. [b] 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 $CaCl_2$.

5′-TGTTAAACA-3′ and 5′-AAAAAGACAAAAA-3′ with equilibrium association constants $K_a = 1 \times 10^8$ M^{-1} and 5×10^7 M^{-1}, respectively. The glycine-linked polyamide ImPyPy-G-PyPyPy-Dp binds both 5′-TGTTAAACA-3′ and 5′-AAAAAGACAAAAA-3′ with affinities similar (equal and 2-fold higher, respectively) to those of ImPyPy-Py-PyPyPy-Dp. In contrast, the β-alanine linked polyamide ImPyPy-β-PyPyPy-Dp binds 5′-TGTTAAACA-3′ and 5′-AAAAAGACAAAAA-3′ with affinities higher than those of ImPyPy-Py-PyPyPy-Dp by factors of approximately 8 ($K_a = 8 \times 10^8$ M^{-1}) and 85 ($K_a = 5 \times 10^9$ M^{-1}), respectively. Relative to ImPyPy-Py-PyPyPy-Dp, the hairpin-forming polyamide ImPyPy-γ-PyPyPy-Dp binds 5′-TGTTAAACA-3′ (a matched hairpin binding site) and 5′-AAAAAGACAAAAA-3′ (a mismatched hairpin binding site) with equal and 8-fold-lower affinities, respectively.

Specificity for 5′-AAAAAGACAAAAA-3′ versus 5′-ATATAGACATATA-3′. Comparison of the binding affinities of the four polyamides at the 13 bp sites 5′-AAAAAGACAAAAA-3′ and 5′-ATATAGACATATA-3′ indicates that the specificity between the sites depends on the internal amino acid (Table 1). ImPyPy-G-PyPyPy-Dp and ImPyPy-β-PyPyPy-Dp are approximately 20-fold and $\geq$5-fold specific respectively for 5′-AAAAAGACAAAAA-3′ versus 5′-ATATAGACATATA-3′. The polyamides ImPyPy-Py-PyPyPy-Dp and ImPyPy-γ-PyPyPy-Dp bind 5′-AAAAAGACAAAAA-3′ and 5′-ATATAGACATATA-3′ with similar affinities.

Discussion

Implications for the Design of Minor Groove Binding Molecules. The results presented here reveal that β-alanine is an optimal linker for joining two three-ring subunits in an extended conformation, providing a useful structural motif for the design of new polyamides targeted to sites longer than 8 bp. Recently, it has been shown that as the length of a polyamide having the general sequence $Im(Py)_x$-Dp increases beyond five rings (corresponding to a 7 bp binding site), the binding affinity ceases to increase with increasing polyamide length, indicating that the *N*-methylimidazole and *N*-methylpyrrole residues fail to maintain an appropriate base-pair register across the entire length of the polyamide−DNA complex.[11] The higher binding affinities observed here for ImPyPy-β-PyPyPy-Dp relative to ImPyPy-Py-PyPyPy-Dp indicate that the flexible β-alanine linker relieves the register mismatch, allowing both three-ring subunits to bind optimally. Notably, higher binding affinities are observed for ImPyPy-β-PyPyPy-Dp versus ImPyPy-Py-PyPyPy-Dp despite the higher conformational entropy and lower aromatic surface area of the β-alanine-linked polyamide. The observation here that β-alanine effectively links polyamide subunits within 2:1 polyamide−DNA complexes is consistent with the previously reported finding that β-alanine effectively links polyamide subunits within 1:1 polyamide−DNA complexes.[14]

From the standpoint of binding specificity, the observation here that a single compound can bind in multiple binding modes is problematic. The next generation of pyrrole−imidazole polyamides targeted to binding sites greater than 8 bp in length should incorporate constraints that specify a single binding mode.

The previously described γ-aminobutyric acid-based "hairpin" motif[7] complements the β-alanine-based "extended" motif described here. For ImPyPy-γ-PyPyPy-Dp, the binding isotherms and affinities observed here are consistent with the previous report that γ-aminobutyric acid effectively links polyamide subunits in a "hairpin" conformation[7] and indicate that γ-aminobutyric acid does not effectively link polyamide subunits in an extended conformation. Significantly, these observations indicate that (1) extended binding modes will not compromise the sequence specificity of hairpin-forming, γ-aminobutyric acid-linked polyamides, and (2) β-alanine and γ-aminobutyric acid linkers could be used within a single polyamide with predictable results.

The results reported here expand the binding site size targetable with pyrrole−imidazole polyamides and provide structural motifs that will facilitate the design of new pyrrole−imidazole polyamides targeted to other sequences.

Experimental Section

Materials. Restriction endonucleases were purchased from either New England Biolabs or Boeringer-Mannheim and used according to the manufacturer's protocol. *Escherichia coli* XL-1 Blue competent cells were obtained from Stratagene. Sequenase (version 2.0) was obtained from United States Biochemical, and DNase I was obtained from Pharmacia. [α-^{32}P]Thymidine 5′-triphosphate ($\geq$3000 Ci/mmol) and [α-^{32}P]deoxyadenosine 5′-triphosphate ($\geq$6000 Ci/mmol) were purchased from DuPont NEN.

Construction of Plasmid DNA. Plasmid pJT3 was prepared by standard methods. Briefly, plasmid pJT1 was prepared by hybridization of two complementary sets of synthetic oligonucleotides: 5′-CCGG-

(14) (a) Youngquist, R. S.; Dervan, P. B. *J. Am. Chem. Soc.* **1987**, *109*, 7564. (b) Griffin, J. H.; Dervan, P. B. Unpublished results.

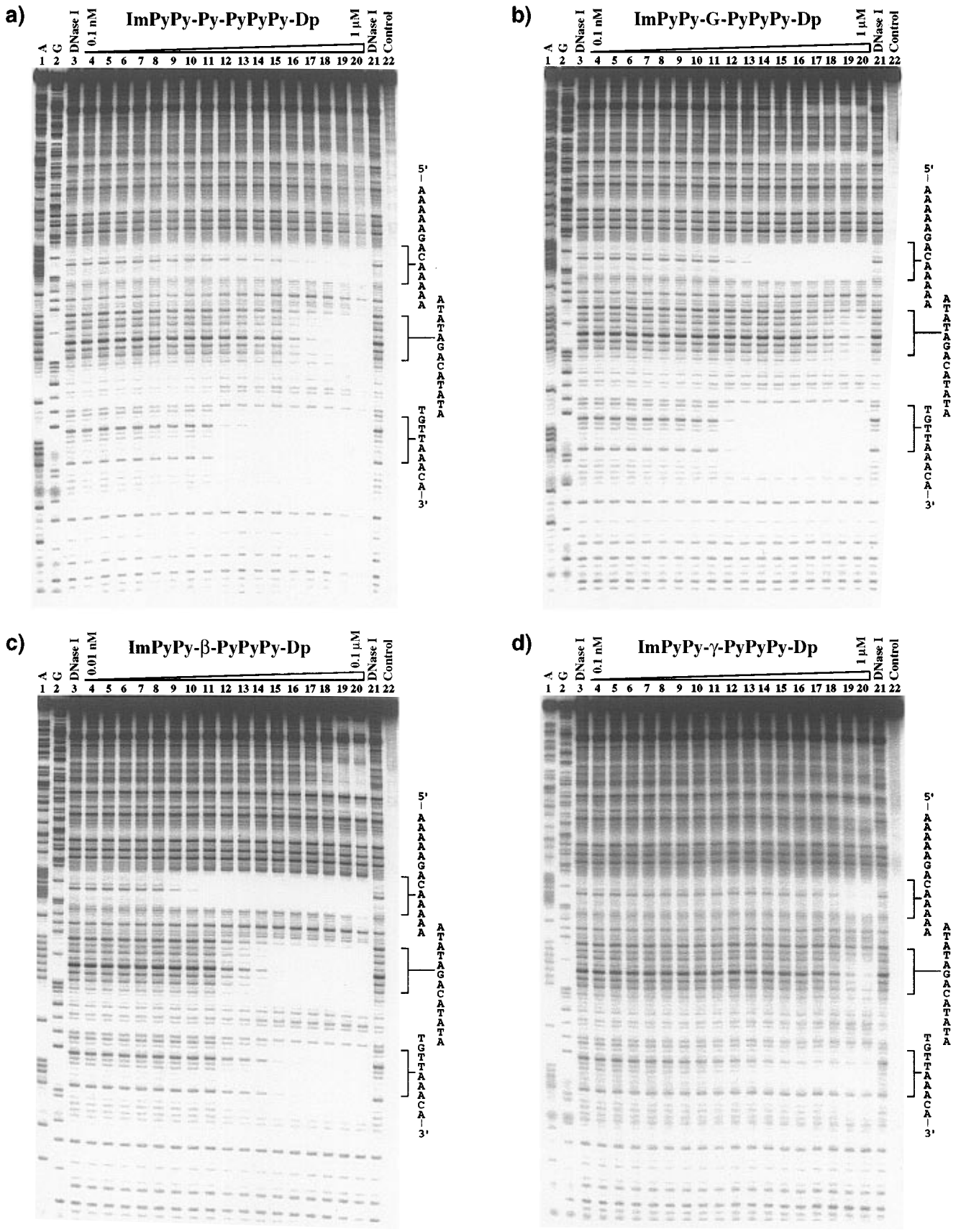

Figure 5. Storage phosphor autoradiograms of 8% denaturing polyacrylamide gels used to separate the fragments generated by DNase I digestion in quantitative footprint titration experiments. Lanes 1 and 2, A and G sequencing lanes. Lanes 3 and 21: DNase I digestion products obtained in the absence of polyamide. Lanes 4−20: DNase I digestion products obtained in the presence of 0.1 nM (0.01 nM), 0.2 nM (0.02 nM), 0.5 nM (0.05 nM), 1 nM (0.1 nM), 1.5 nM (0.15 nM), 2.5 nM (0.25 nM), 4 nM (0.4 nM), 6.5 nM (0.65 nM), 10 nM (1 nM), 15 nM (1.5 nM), 25 nM (2.5 nM), 40 nM (4 nM), 65 nM (6.5 nM), 100 nM (10 nM), 200 nM (20 nM), 500 nM (50 nM), 1 μM (0.1 μM) (concentrations used for polyamide ImPyPy-β-PyPyPy-Dp only are in parentheses). Lane 22: intact DNA. The targeted binding sites are indicated on the right side of the autoradiograms. All reactions contain 15 kcpm restriction fragment, 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

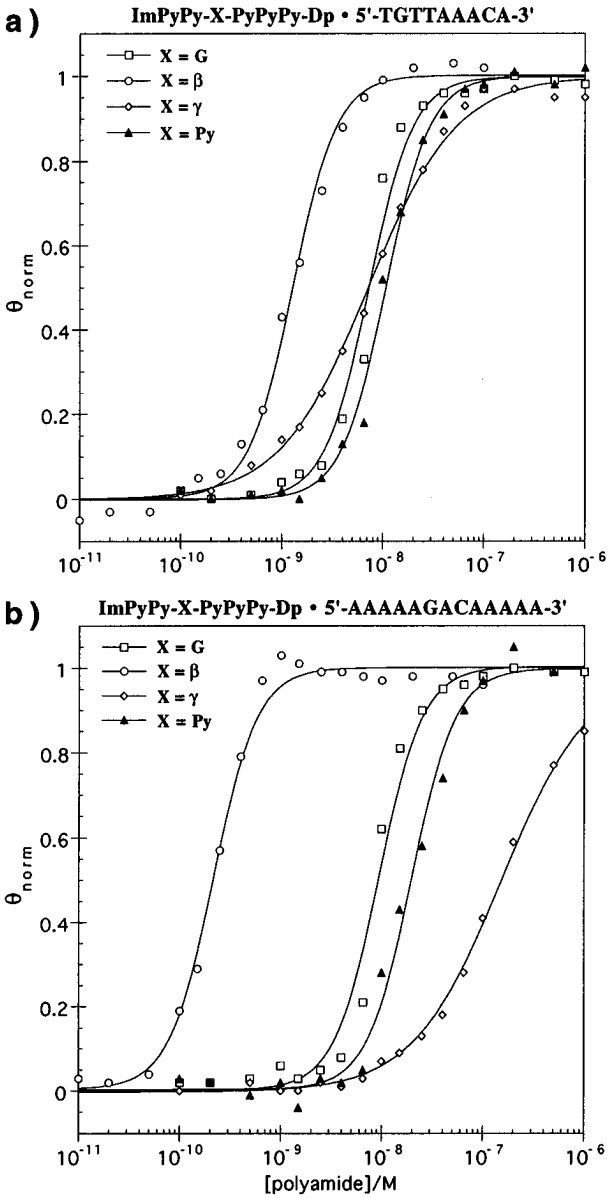

Figure 6. Data obtained from quantitative DNase I footprint titration experiments showing the effect of the internal amino acid X on binding of the polyamides ImPyPy-X-PyPyPy-Dp to the sites (a) 5′-TGTTAAACA-3′ and (b) 5′-AAAAAGACAAAAA-3′, where X = G (□), β (●), γ (◇), and Py (▲). The (θ_{norm}, $[L]_{tot}$) data points were obtained as described in the Experimental Section. Each data point is the average value obtained from three quantitative footprint titration experiments.

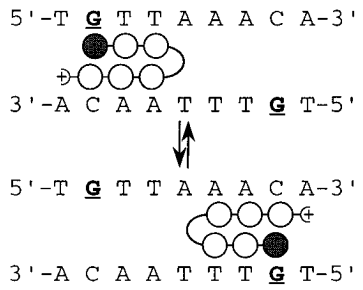

Figure 7. Model for the complex formed by ImPyPy-γ-PyPyPy-Dp with 5′-TGTTAAACA-3′ (5 bp, "hairpin").

GAACGTAGCGTACCGGTGCAAAAAGACAAAAAGG-CTCGA-3′ and 5′-GGCGTCGAGCCTTTTTGTCTTTTTGCACC-GGTACGCTACGTTC-3′, and 5′-CGCCGCATATAGACATATAG-GGCCCAGCTCGTCCTAGCTAGCGTCGTAGCGTCTTAAG-3′ and 5′-TCGACTTAAGACGCTACGACGCTAGCTAGGACGAGCT-GGGCCCTATATGTCTATATGC-3′. The resulting oligonucleotide duplexes were phosphorylated with ATP and T4 polynucleotide kinase and ligated to the large pUC19 *Ava*I/*Sal*I restriction fragment using T4 DNA ligase. *E. coli* XL-1 Blue competent cells were then transformed with the ligated plasmid, and plasmid DNA from ampicillin-resistant white colonies isolated using a Promega Maxi-Prep kit. Plasmid pJT3 was prepared in a similar manner, except that the following synthetic oligonucleotides were hybridized and cloned into the large *Apa*I/*Sal*I fragment of pJT1: 5′-GTCGACGCTGTTAAA-CAGGCTCGACGCCAGCTCGTCCTAGCTAGCGT-CGTAGCGTCTTAAGAG-3′ and 5′-TCGACTCTTAAGACGCTAC-GACGCTAGCTAGGACGAGCTGGCGTCGAGCCT-GTTTAACAGCGTCGACGGCC-3′. The presence of the desired insert was determined by restriction analysis and dideoxy sequencing. Plasmid DNA concentration was determined at 260 nm using the relation 1 OD unit = 50 μg/mL duplex DNA.

Preparation of ^{32}P-End-Labeled Restriction Fragments. Plasmid pJT3 was digested with *Afl*II, labeled at the 3′-end using Sequenase (version 2.0), and digested with *Fsp*I. The 281 bp restriction fragment was isolated by nondenaturing gel electrophoresis and used in all quantitative footprinting experiments described here. Chemical sequencing reactions were performed as described.[15,16] Standard techniques were employed for DNA manipulations.[17]

Quantitative DNase I Footprint Titration.[12,13] All reactions were executed in a total volume of 40 μL. We note explicitly that no carrier DNA was used in these reactions. A polyamide stock solution or H_2O (for reference lanes) was added to an assay buffer containing radiolabeled restriction fragment (15 000 cpm), affording final solution conditions of 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, and either (i) 0.1 nM to 1 μM polyamide, for all polyamides except ImPyPy-β-PyPyPy-Dp, (ii) 0.01 nM to 0.1 μM polyamide for ImPyPy-β-PyPyPy-Dp, or (iii) no polyamide (for reference lanes). The solutions were allowed to equilibrate for 5 h at 22 °C. Footprinting reactions were initiated by the addition of 4 μL of a DNase I stock solution (at the appropriate concentration to give ~55% intact DNA) containing 1 mM dithiothreitol and allowed to proceed for 7 min at 22 °C. The reactions were stopped by the addition of 10 μL of a solution containing 1.25 M NaCl, 100 mM EDTA, and 0.2 mg/mL glycogen, and ethanol precipitated. The reaction mixtures were resuspended in 1X TBE/80% formamide loading buffer, denatured by heating at 85 °C for 10 min, and placed on ice. The reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross-link, 7 M urea) in 1X TBE at 2000 V. Gels were dried and exposed to a storage phosphor screen (Molecular Dynamics).

Quantitation and Data Analysis. Data from the footprint titration gels were obtained using a Molecular Dynamics 400S PhosphorImager followed by quantitation using ImageQuant software (Molecular Dynamics). Background-corrected volume integration of rectangles encompassing the footprint sites and a reference site at which DNase I reactivity was invariant across the titration generated values for the site intensities (I_{site}) and the reference intensity (I_{ref}). The apparent fractional occupancies (θ_{app}) of the sites were calculated using the equation

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

where I°_{site} and I°_{ref} are the site and reference intensities, respectively, from a control lane to which no polyamide was added.

The ($[L]_{tot}$, θ_{app}) data points were fitted to a general Hill equation (eq 2) by minimizing the difference between θ_{app} and θ_{fit}:

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

where $[L]_{tot}$ is the total polyamide concentration, K_a is the apparent association constant, and θ_{min} and θ_{max} are the experimentally deter-

(15) Iverson, B. L.; Dervan, P. B. *Nucleic Acids Res.* **1987**, *15*, 7823−7830.
(16) Maxam, A. M.; Gilbert, W. S. *Methods Enzymol.* **1980**, *65*, 499−560.
(17) Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.

mined site saturation values when the site is unoccupied or saturated, respectively. The data were fitted using a nonlinear least squares fitting procedure with K_a, $\theta_{max,}$ and θ_{min} as the adjustable parameters, and with either $n = 2$ or $n = 1$ depending on which value of n gave the better fit. We note explicitly that treatment of the data in this manner does not represent an attempt to model a binding mechanism. Rather, we have chosen to compare values of the association constant, a parameter that represents the concentration of polyamide at which the binding site is half-saturated. The binding isotherms were normalized using the following equation:

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

Three sets of data were used in determining each association constant.

The method for determining association constants used here involves the assumption that $[L]_{tot} \approx [L]_{free}$, where $[L]_{free}$ is the concentration of polyamide free in solution (unbound). For very high association constants this assumption becomes invalid, resulting in underestimated association constants. In the experiments described here, the DNA concentration is estimated to be ~50 pM. As a consequence, measured association constants of 1×10^9 M^{-1} and 5×10^9 M^{-1} underestimate the true association constants by factors of approximately ≤1.5 and 1.5−2, respectively.

Acknowledgment. We are grateful to the National Institutes of Health (Grant GM-27681) for research support, to the National Science Foundation for a predoctoral fellowship to J.W.T., and to the Howard Hughes Medical Institute for a predoctoral fellowship to E.E.B.

JA960726O