Effects of the A·T/T·A Degeneracy of Pyrrole−Imidazole Polyamide Recognition in the Minor Groove of DNA†

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

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

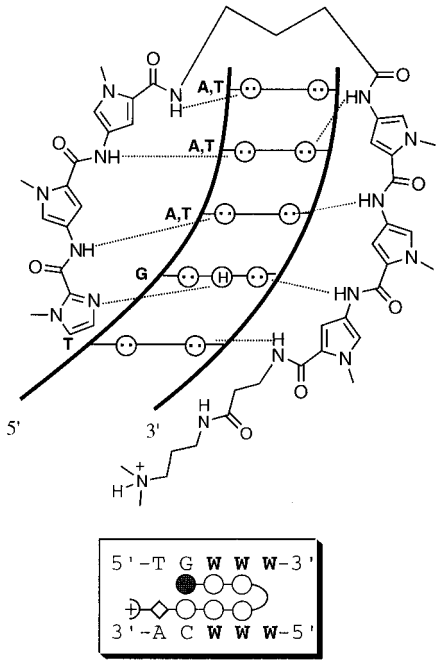
Received March 27, 1996; Revised Manuscript Received July 30, 1996⊗

ABSTRACT: Pairing rules have been developed to predict the sequence specificity of minor groove binding polyamides containing pyrrole (Py) and imidazole (Im) amino acids. An Im/Py pair distinguishes G·C from C·G and both of these from A·T/ T·A base pairs. A Py/Py pair appears not to distinguish A·T from T·A base pairs. To test the extent of this degeneracy, the affinity and binding orientation of the hairpin polyamide ImPyPy-γ-PyPyPy-β-Dp were measured for eight possible five base pair 5′-TG(A,T)$_3$-3′ match sites. Affinity cleavage experiments using a polyamide with an EDTA·Fe(II) moiety at the carboxy terminus, ImPyPy-γ-PyPyPy-β-Dp-EDTA·Fe(II), are consistent with formation of an oriented 1:1 hairpin polyamide complex at all eight 5′-TG(A,T)$_3$-3′ binding sites [20 mM HEPES, 200 mM NaCl, 50 mg/ml glycogen, pH 7.0, 22 °C, 5 mM DTT, 1 mM Fe(II)]. Quantitative DNase I footprint titration experiments reveal that ImPyPy-γ-PyPyPy-β-Dp binds all eight 5′-TG(A,T)$_3$-3′ target sites with only a 12-fold difference in the equilibrium association constants between the strongest site, 5′-TGTTT-3′ ($K_a = 2.1 \times 10^8$ M^{-1}), and the weakest site, 5′-TGAAT-3′ ($K_a = 1.8 \times 10^7$ M^{-1}) (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, 22 °C). This relatively small range indicates that the Py/Py pair is approximately degenerate for recognition of A,T base pairs, affording generality with regard to targeting sequences of mixed A·T/T·A composition.

Pyrrole−imidazole polyamide−DNA complexes (Wade et al., 1992, 1993; Mrksich et al., 1992) provide a paradigm for the design of artificial molecules for recognition of double-helical DNA. 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. The DNA-binding sequence specificity of these small molecules depends on the sequence of side-by-side amino acid pairings (Mrksich & Dervan, 1993a, 1995; Geierstanger et al., 1993, 1994a,b). A pairing of imidazole opposite pyrrole recognizes a G·C base pair, while a Py/Im combination targets a C·G base pair (Wade et al., 1992, 1993; Mrksich et al., 1992). A Py/Py pair has apparent degeneracy for A·T/T·A base pairs (Pelton & Wemmer, 1989, 1990; Wade et al., 1992; Chen et al., 1994).

Covalently linking polyamide subunits has led to designed ligands with increased affinity and specificity (Mrksich & Dervan, 1993b, 1994; Dwyer et al., 1993; Mrksich et al. 1994; Parks et al., 1996a,b; Trauger et al., 1996a,b). The polyamide ImPyPy-γ-PyPyPy-Dp containing γ-aminobutyric acid (γ) as an internal guide residue was found to specifically bind as a "hairpin" to a designated 5′-TGTTA-3′ target site with 300-fold enhancement relative to the binding affinities of the individual unlinked polyamide pair, ImPyPy and PyPyPy (Figures 1 and 2).

The discrimination of G·C from C·G base pairs and both of these from A·T/T·A base pairs by pyrrole−imidazole polyamides has been demonstrated (Wade et al., 1992). A key issue is to determine if a polyamide of core sequence composition ImPyPy-γ-PyPyPy would bind all possible 5′-(A,T)G(A,T)$_3$-3′ target sequences. Alternatively, given the



ImPyPy-γ-PyPyPy-β-Dp•5'-TG(A,T)$_3$-3'

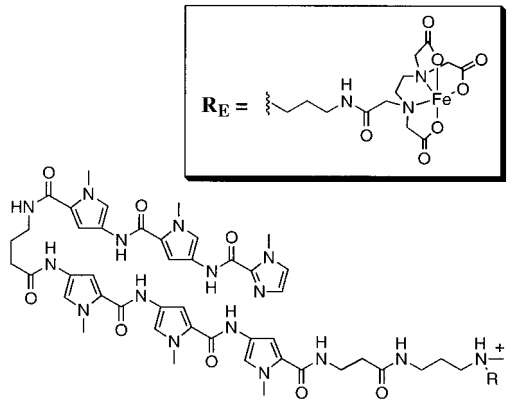
FIGURE 1: (Top) Model for the complex formed between the hairpin polyamide ImPyPy-γ-PyPyPy-β-Dp with a 5′-TG(A,T)$_3$-3′ site. Circles with dots represent lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. (Bottom) Schematic binding model. The imidazole and pyrrole rings are represented as shaded and unshaded spheres, respectively; the β-alanine residue is represented as an unshaded diamond. W is either A·T or T·A.

†This research was supported by National Institutes of Health Grant GM-27681, a National Science Foundation predoctoral fellowship to S.W., and a Howard Hughes Medical Institute predoctoral fellowship to E.E.B.

* To whom correspondence should be addressed.

⊗ Abstract published in *Advance ACS Abstracts,* September 1, 1996.

S0006-2960(96)00744-1 CCC: $12.00



R = CH_3 ImPyPy-γ-PyPyPy-β-Dp

R = R_E ImPyPy-γ-PyPyPy-β-Dp-EDTA•Fe(II)

FIGURE 2: Structure of the hairpin polyamides ImPyPy-γ-PyPyPy-β-Dp and ImPyPy-γ-PyPyPy-β-Dp-EDTA·Fe(II).

sequence-dependent variation in DNA groove width (Dickerson et al., 1994; Yoon et al., 1988), perhaps only a smaller subset of 5′-(A,T)G(A,T)$_3$-3′ sequences would be structurally compatible with polyamide−DNA complex formation. To address this question, a plasmid was designed containing eight binding sites of the form 5′-TG(A,T)$_3$-3′: 5′-TGTTT-3′, 5′-TGTTA-3′, 5′-TGTAA-3′, 5′-TGTAT-3′, 5′-TGAAA-3′, 5′-TGATT-′, 5′-TGAAT-3′, and 5′-TGATA-3′, with each site flanked by the same 12 base pair sequence (Figure 3). Quantitative DNase I footprint titration experiments afford a comparison of the equilibrium association constant for binding of each site by the polyamide ImPyPy-γ-PyPyPy-β-Dp, previously optimized for a hairpin motif (Parks et al., 1996). For controls, affinity cleavage experiments with ImPyPy-γ-PyPyPy-β-Dp-EDTA·Fe(II) confirm that the polyamide binds each five base pair site in a single orientation, supporting the hairpin model. We report here that ImPyPy-γ-PyPyPy-β-Dp binds *all sites* of the form 5′-TG(A,T)$_3$-3′ with a 12-fold range between the highest and lowest observed affinities.

MATERIALS AND METHODS

Materials. Sonicated, deproteinized calf thymus DNA, from Pharmacia, was dissolved in filter-sterilized water to a final concentration of 800 μM in base pairs and stored at 4 °C. Glycogen was purchased from Boehringer-Mannheim as a 20 mg/mL aqueous solution. Nucleotide triphosphates were purchased from Pharmacia and used as supplied. Nucleoside triphosphates labeled with ^{32}P ($\geq$3000 C$_i$/mmol) were obtained from Dupont-New England Nuclear. Cerenkov radioactivity was measured with a Beckman LS 2801 scintillation counter. All enzymes were purchased from Boehringer-Mannheim and were used according to the supplier's recommended protocol in the activity buffer provided. Plasmid pUC19 was obtained from Worthington Biochemical. A solution of 0.5 M EDTA, pH 8.0, was purchased from Ultrapure. Phosphoramidites were from Glen Research. The pH of buffer solutions was recorded using a digital pH/millivolt meter (model no. 611, Orion Research) and a ROSS semimicro combination pH electrode. General manipulations of duplex DNA (Sambrook et al., 1989) and oligonucleotides (Gait, 1984) were performed according to established procedures.

Synthesis of ImPyPy-γ-PyPyPy-β-Dp and ImPyPy-γ-PyPyPy-β-Dp-EDTA. Polyamides optimized for hairpin formation (Parks et al., 1996) were synthesized from β-alanine−PAM resin using solid-phase methods (Baird & Dervan, 1996) and characterized by a combination of analytical HPLC, ^{1}H NMR spectroscopy, and MALDI-TOF mass spectroscopy. MS: mass observed for ImPyPy-γ-PyPyPy-β-Dp, 978.0; 978.1 calculated; 1295.4 observed for ImPyPy-γ-PyPyPy-β-Dp-EDTA; 1295.3 calculated.

Construction of Plasmid DNA. Oligodeoxynucleotides were synthesized by standard automated solid support chemistry using an Applied Biosystems Model 380B DNA synthesizer and *O*-cyanoethyl *N*,*N*-diisopropylphosphoramidites. Plasmid pDEH1 was prepared by hybridization of two complementary sets of synthetic oligonucleotides: (1) 5′-CTAGACCACCATTGTTTGACCACCCACATT-GTTAGACCACCCACATTGTAAGACCACCC-ACATTGTATGACCACC-3′; (2) 5′-GTCATACAATGT-GGGTGGTCTTACAATGTGGGTGGTCTAACA-ATGTGGGTGGTCAAACAATGTGGGTGGT-3′; (3) 5′-CACATTGAAAGACCACCCACATTGATTGA-CCACCCACATTGAATGACCACCCACATTGA-TAGACCACCCACATTGCA-3′; (4) 5′-ATGTGGGTGG-TCTATCAATGTGGGTGGTCATTCAATGT-GGGTGGTCAATCAATGTGGGTGGTC-TTTCAATGTGGGTG-3′. Oligonucleotides 2 and 3 were phosphorylated with dATP and T4 polynucleotide kinase and then annealed to their respective complementary strands, 1 and 4. The two sets of duplexes were then ligated to the large pUC19 *Xba*I/*Pst* I restriction fragment using T4 DNA ligase. The ligated plasmid was then used to transform Epicurean Coli XL-1 Blue Supercompetent cells. Colonies were selected for α-complementation on 25 mL of Luria−Bertani medium agar plates containing 50 mg/mL ampicillin and treated with X-GAL and IPTG solutions. Large-scale plasmid purification was performed using Qiagen purification kits. The presence of the desired insert was determined by dideoxy sequencing using a USB Sequenase version 2.0 kit. Plasmid DNA concentration was determined at 260 nm using the relation 1 OD unit = 50 mg/mL duplex DNA. The plasmid pDEH1 was digested with *Eco*RI, labeled at the 3′ end using Sequenase version 2.0, and digested with *Pvu*II. The 370 base pair fragment was isolated by nondenaturing gel electrophoresis and used in all experiments described here. For affinity cleaving reactions, pDEH1 also was 5′-^{32}P labeled. First, the plasmid pDEH1 was digested with *Eco*RI and then dephosphorylated with calf intestine alkaline phosphatase. The digested plasmid then was 5′-^{32}P labeled with [γ-^{32}P]dATP using T4 polynucleotide kinase and digested with *Pvu*II. The 5′-labeled fragment was isolated by nondenaturing gel electrophoresis and used in affinity cleaving experiments. Chemical sequencing adenine-specific reactions were carried out as previously described (Iverson & Dervan, 1987).

Quantitative DNase I Footprint Titrations. All reactions were executed in a total volume of 400 μL (Brenowitz et al., 1986). A polyamide stock solution or H_2O (for reference lanes) was added to an assay buffer containing 3′-^{32}P-radiolabeled restriction fragment (20 000 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−1 μM polyamide ImPyPy-γ-PyPyPy-β-Dp or (ii) no polyamide (for reference lanes). The solutions were allowed to

FIGURE 3: Partial sequence of the 370 base pair *Eco*RI/*Pvu*II restriction fragment. Eight five base pair binding sites having the sequence 5′-TG(A,T)$_3$-3′ proximal to the ^{32}P label at the *Eco*RI site were analyzed by quantitative footprint titration analysis.

equilibrate for 24 h at 22 °C. Footprinting reactions were initiated by the addition of 10 μL of a stock solution of DNase I (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 50 μL of a solution containing 2.25 M NaCl, 150 mM EDTA, 23 μM base pair calf thymus DNA, and 0.6 mg/mL glycogen and ethanol precipitated. The reactions were resuspended in 1 × TBE/80% formamide loading buffer, denatured by heating at 85 °C for 15 min, and placed on ice. The reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross-linking, 7 M urea) in 1 × TBE at 2000 V for 1.5 h. Gels were dried on a slab dryer and then exposed to a storage phosphor screen at 22 °C.

Quantitation by Storage Phosphor Technology Autoradiography. Photostimuatable storage phosphor imaging plates (Kodak Storage Phosphor Screen SO230 obtained from Molecular Dynamics) were pressed flat against dried gel samples and exposed in the dark at 22 °C for 12−24 h. A Molecular Dynamics 400S PhosphorImager was used to obtain all data from the storage screens (Johnston et al., 1990). The data were analyzed by performing volume integration of the target site and reference blocks using the ImageQuant version 3.3 software running on a Compaq Pentium 80.

Quantitative DNase I Footprint Titration Data Analysis. 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 occupancy (θ_{app}) of the sites was calculated using the equation:

$$\theta_{app} = 1 - \frac{I_{site}/I_{ref}}{I_{site}^{0}/I_{ref}^{0}} \quad (1)$$

where I_{site}^{0} and I_{ref}^{0} are the site and reference intensities, respectively, from a DNase I control lane to which no polyamide was added.

The ($[L]_{tot}$, θ_{app}) data were fit to a Langmuir binding isotherm (eq 2, $n = 1$) by minimizing the difference between θ_{app} and θ_{fit}, using the modified Hill equation:

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

where $[L_{tot}]$ is the total polyamide concentration, K_a is the equilibrium association constant, and θ_{min} and θ_{max} are the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. The data were fit using a nonlinear least-squares fitting procedure of KaleidaGraph software (version 3.0.1, Abelbeck Software) 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. Four sets of acceptable data were used in determining each association constant. All lanes from a gel were used unless a visual inspection revealed a data point to be obviously flawed relative to neighboring points. The data were normalized using the equation:

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

At higher concentrations of polyamide, (≥1 μM for ImPyPy-γ-PyPyPy-β-Dp), the reference site becomes partially protected due to nonspecific DNA binding, resulting in low θ_{app} values. For this reason, higher concentrations were not used.

Affinity Cleaving Titrations. All affinity cleavage reactions (Schultz & Dervan, 1983, 1984) were performed in a total volume of 400 μL. A polyamide stock solution (ImPyPy-γ-PyPyPy-β-Dp-EDTA) or H_2O (for reference lanes) was added to an assay buffer containing either 3′-^{32}P-labeled or 5′-^{32}P-labeled restriction fragment (20 000 cpm), affording final solution conditions of 20 mM HEPES, 200 mM NaCl, 50 mg/mL glycogen, pH 7, and either (i) 0.1 nM−1 μM polyamide ImPyPy-γ-PyPyPy-β-Dp-EDTA or (ii) no polyamide (for reference lanes). Solutions were incubated at 22 °C for 24 h. Then, 20 μL of a 20 mM ferrous ammonium sulfate solution was added, and the solution was allowed to equilibrate for 20 min at 22 °C. Affinity cleaving reactions were initiated by the addition of 40 μL of a 50 mM dithiothreitol solution and reacted for 11 min at 22 °C. Reactions were stopped by the addition of 10 μL of a solution containing 2.8 mg/mL glycogen and 112 μM base pair calf thymus DNA and ethanol precipitated. Reactions were resuspended in 1 × TBE/80% formamide loading buffer, denatured by heating at 85 °C for 15 min, and placed on ice. Reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross-linking, 7 M urea) in 1 × TBE at 2000 V for either 1.5 h (for visualization of the first four binding sites) or 3 h (for visualization of the last four binding sites). Gels were dried on a slab dryer and then exposed to a storage phophor screen at 22 °C.

Data Analysis of Affinity Cleavage Titrations. The data were analyzed by performing volume integrations of the cleavage bands and reference bands using the ImageQuant version 3.3 software running on a Compaq Pentium 80. Background-corrected volume integration of rectangles en-

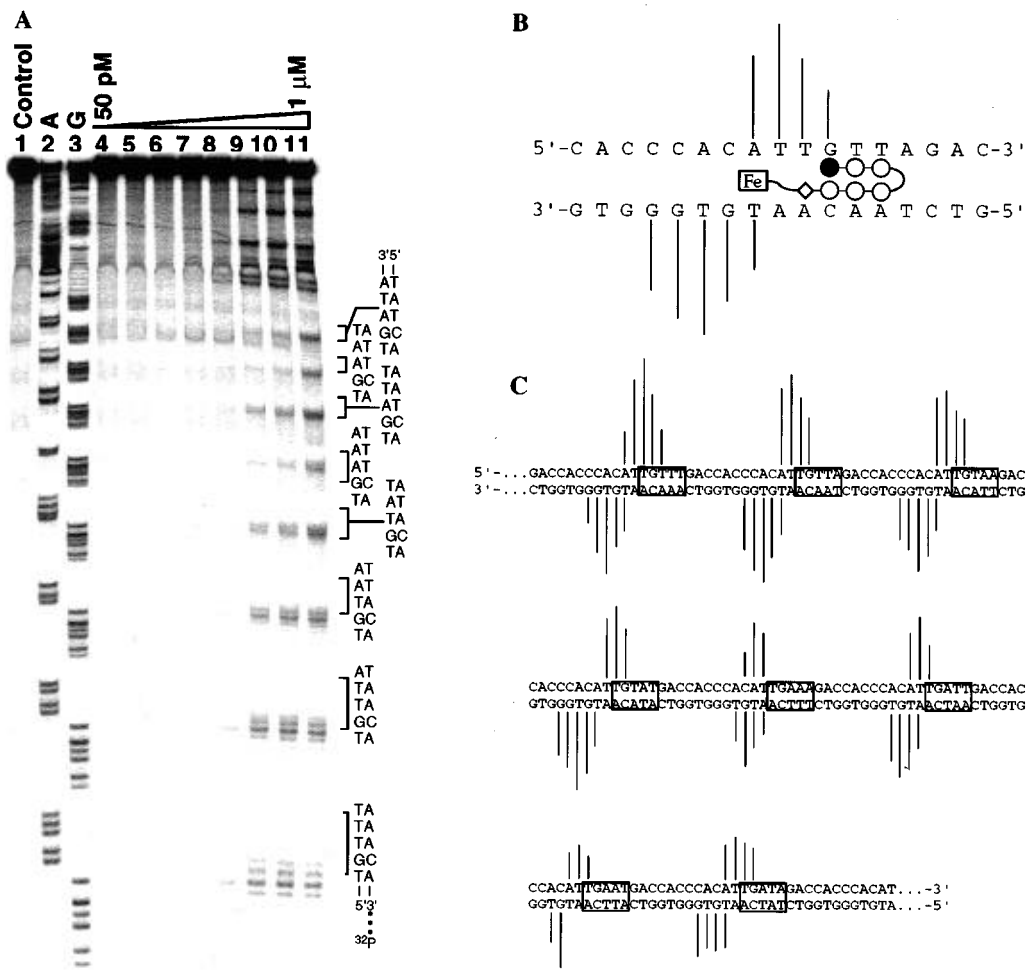


FIGURE 4: (A) Storage phosphor autoradiogram of a 8% denaturing polyacrylamide gel used to separate the fragments generated by affinity cleaving experiments performed with ImPyPy-γ-PyPyPy-β-Dp-EDTA·Fe(II): lane 1, digestion products obtained in the absence of polyamide; lanes 2 and 3, A and G sequencing lanes; lanes 4−11, digestion products obtained in the presence of **2**, 50 pM, 100 pM, 1 nM, 5 nM, 10 nM, 50 nM, 100 nM, and 1 μM. The targeted binding sites are indicated on the right side of the autoradiograms. All reactions contain 20 kcpm of 3′-^{32}P restriction fragment, 20 mM HEPES, 200 mM NaCl, and 50 mg/mL glycogen, pH 7. (B) Schematic binding model of affinity cleaving at the 5′-TGTTA-3′ binding site. The imidazole and pyrrole rings are represented as shaded and unshaded spheres, respectively; the β-alanine residue is represented as an unshaded diamond; the position of the iron is represented as a rectangle. (C) Cleavage of the 370 bp restriction fragment from pDEH1 at 100 nM concentration of polyamide. Lines are proportional to the integrated densities of the cleavage bands.

compassing cleavage bands was normalized to a maximum value of 3.95.

RESULTS AND DISCUSSION

Hairpin Motif. Affinity cleavage experiments using a polyamide with Fe(II)·EDTA at the carboxy terminus confirm that ImPyPy-γ-PyPyPy-β-Dp binds each discrete site with a *single orientation*. Cleavage experiments with ImPyPy-γ-PyPyPy-β-Dp-EDTA·Fe(II) were performed on the 370 base pair restriction fragment radiolabeled at either the 5′ or 3′ end [20 mM HEPES, 200 mM NaCl, 50 mg/mL glycogen, pH 7.0, 22 °C, 5 mM DTT, 1 mM Fe(II)]. A single cleavage locus is observed proximal to the 5′ side of each of the eight 5′-TG(A,T)$_3$-3′ binding sites, indicating that the carboxy terminus of the polyamide is located at the 5′ side of each binding site. A 3′-shifted asymmetric cleavage pattern is consistent with the location of the 1:1 polyamide complex in the minor groove (Figure 4).

The observation of a single cleavage locus is consistent only with an oriented 1:1 complex and rules out any 2:1 overlapped or extended binding motifs (Trauger et al., 1996c). A 1:1 oriented but extended motif would require at least an eight base pair binding site, which is inconsistent with high-resolution MPE footprinting data (Parks et al., 1996a). The hairpin structure is supported by direct NMR structure studies on a six-ring hairpin polyamide of sequence composition ImPyPy-γ-PyPyPy binding to a core five base pair 5′-TGTTA-3′ site (de Clairac et al., 1996).

Relative Energetics. DNase I footprinting on the 3′-^{32}P-end-labeled 370 base pair *Eco*RI/*Pvu*II restriction fragment from the plasmid pDEH1 (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, 22 °C) reveals that the equilibrium association constants for ImPyPy-γ-PyPyPy-β-Dp binding each of the five base pair sites, 5′-TG(A,T)$_3$-3′, range from $K_a = 2.1 \times 10^8$ M^{-1} to $K_a = 1.8 \times 10^7$ M^{-1} in decreasing order: 5′-TGTTT-3′ > 5′-TGTTA-3′ > 5′-TGTAA-3′ > 5′-TGTAT-3′ > 5′-TGATT-3′ > 5′-TGATA-3′ > 5′-TGAAA-3′ > 5′-TGAAT-3′ (Table 1, Figure 5). The polyamide displays a binding isotherm (eq 2, $n = 1$) consistent with binding as an intramolecular hairpin at all sites (Figure 6).

The affinities of ImPyPy-γ-PyPyPy-β-Dp for binding sites of the type 5′-TG(A,T)$_3$-3′ may be grouped into two sets according to sequence composition: 5′-TG**T**(A,T)$_2$-3′ and

Table 1: Equilibrium Association Constants and Binding Energies for ImPyPy-γ-PyPyPy-β-Dp[a,b]

binding site	K_a (M^{-1})	K_a (rel)[c]	ΔG (kcal/mol)
5′-TGTTT-3′	2.1 (±0.7) × 10^8	12	−11.2
5′-TGTTA-3′	1.5 (±0.4) × 10^8	8.4	−11.0
5′-TGTAA-3′	7.3 (±1.0) × 10^7	4.1	−10.6
5′-TGTAT-3′	4.7 (±0.8) × 10^7	2.6	−10.4
5′-TGATT-3′	3.9 (±1.0) × 10^7	2.2	−10.3
5′-TGATA-3′	2.5 (±0.9) × 10^7	1.4	−10.0
5′-TGAAA-3′	2.2 (±0.9) × 10^7	1.2	−9.9
5′-TGAAT-3′	1.8 (±0.8) × 10^7	1	−9.8

[a] Values reported are the mean values measured from four 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 $MgCl_2$, 10 mM KCl, and 5 mM $CaCl_2$. [c] K_a(rel) = K_a(5′-TGWWW-3′)/K_a(5′-TGAAT-3′); W is either A or T.

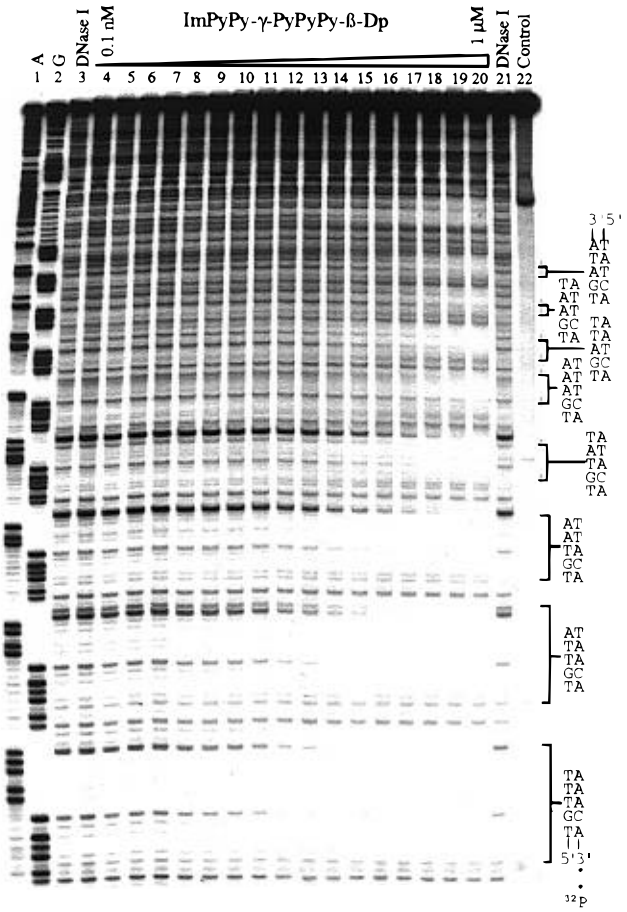


FIGURE 5: Quantitative DNase I footprint titration experiment with ImPyPy-γ-PyPyPy-β-Dp on the 3′-^{32}P-labeled 370 bp *Eco*RI/*Pvu*II restriction fragment from plasmid pDEH1. The eight binding sites 5′-TG(A,T)$_3$-3′ are shown on the right side of the storage phosphor autoradiogram. All reactions contain 20 kcpm of restriction fragment, 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$, pH 7.0. Lane 1, A reaction; lane 2, G reaction; lanes 3 and 21, DNase I standard; lanes 4−20, 0.1 nM, 0.2 nM, 0.5 nM, 1 nM, 1.5 nM, 2.5 nM, 4.0 nM, 6.5 nM, 10 nM, 15 nM, 25 nM, 40 nM, 65 nM, 100 nM, 200 nM, 500 nM, and 1 μM ImPyPy-γ-PyPyPy-β-Dp, respectively.

5′-TG**A**(A,T)$_2$-3′. ImPyPy-γ-PyPyPy-β-Dp binds 5′-TG-**T**(A,T)$_2$-3′ sites with between 2-fold and 12-fold higher affinity than 5′-TG**A**(A,T)$_2$-3′ sites. Overall, sequence composition results in a difference of up to 1.5 kcal/mol for recognition of 5′-TG(A,T)$_3$-3′ sites. X-ray diffraction data suggest that a G-A step acts to narrow the minor groove of

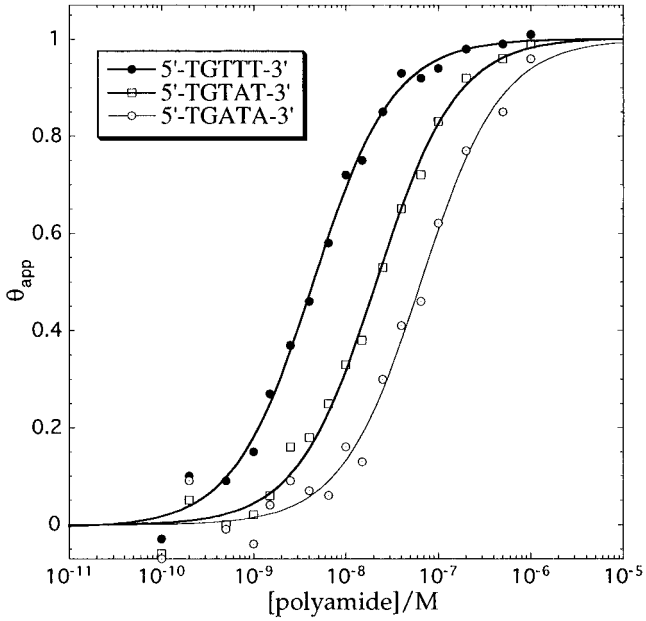


FIGURE 6: Data for the quantitative DNase I footprint titration experiments for ImPyPy-γ-PyPyPy-β-Dp in complex with three binding sites: 5′-TGTTT-3′, 5′-TGTAT-3′, and 5′-TGATA-3′. The θ_{app} points were obtained using phosphostimuatable storage phosphor autoradiography and processed as described in Materials and Methods. The data points for the 5′-TGTTT-3′ site are indicated by filled circles, for the 5′-TGTAT-3′ site by open squares, and for the 5′-TGATA-3′ site by open circles. The solid lines are the best fit Langmuir binding titration isotherms obtained from a nonlinear least-squares algorithm using eq 2.

B-form DNA (Larsen et al., 1991; Narayana et al., 1991; Yoon et al., 1988; Yanagi et al., 1991; Dickerson et al., 1994). A decrease in minor groove width and flexibility would act to disfavor binding of a hairpin polyamide, which prefers a wide, flexible minor groove for favorable binding (Mrksich *et al.*, 1992).

Implications for the Design of Minor Groove Binding Polyamides. The results reported here indicate that A·T and T·A base pairs are degenerate in the hairpin polyamide−DNA motif with a 12-fold difference in binding affinities. These results indicate that at least a 10-fold *range* of binding affinities and sequence specificities will be observed for a polyamide binding to a designated set of match sites containing A·T base pairs. The similiarity of the polyamide binding affinities for the eight 5′-TG(A,T)$_3$-3′ match sites reflects a limit to the specificity of the hairpin polyamide binding motif. Because G·C is distinct from C·G, the most specific recognition will be G,C rich sequences. In order to increase the specificity of pyrrole−imidazole polyamides for sequences of rich A,T composition, methods of recognition to discriminate A·T and T·A base pairs are needed.

REFERENCES

Baird, E. E., & Dervan, P. B. (1996) *J. Am. Chem. Soc. 118*, 6141.
Brenowitz, M., Senear, D. F., Shea, M. A., & Ackers, G. K. (1986) *Methods Enzymol. 130*, 132.
Chen, X., Ramakrishnan, B., Rao, S. T., & Sundaralingham, M. (1994) *Struct. Biol. Nat. 1*, 169.
Dickerson, R. E., Goodsell, D. S., & Neidle, S. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 3579.
Dwyer, T. J., Geierstanger, B. H., Mrksich, M., Dervan, P. B., & Wemmer, D. E. (1993) *J. Am. Chem. Soc. 115*, 9900.
Gait, M. J. (1984) *Oligonucleotide Synthesis: A Practical Approach*, IRL Press, Oxford.

Geierstanger, B. H., Dwyer, T. J., Bathini, Y., Lown, J. W., & Wemmer, D. E. (1993) *J. Am. Chem. Soc. 115,* 4474.
Geierstanger, B. H., Mrksich, M., Dervan, P. B., & Wemmer, D. E. (1994a) *Science 266,* 646.
Geierstanger, B. H., Jacobsen, J. P., Mrksich, M., Dervan, P. B., & Wemmer, D. E. (1994b) *Biochemistry 33,* 3055.
Iverson, B. L., & Dervan, P. B. (1987) *Nucleic Acids Res. 15,* 7823.
Johnston, R. F., Picket, S. C., & Barker, D. L. (1990) *Electrophoresis 11,* 355.
Larsen, T. A., Kopka, M. L., & Dickerson, R. E. (1991) *Biochemistry 30,* 4443.
Mrksich, M., & Dervan, P. B. (1993a) *J. Am. Chem. Soc. 115,* 2572.
Mrksich, M., & Dervan, P. B. (1993b) *J. Am. Chem. Soc. 115,* 9892.
Mrksich, M., & Dervan, P. B. (1994) *J. Am. Chem. Soc. 116,* 3663.
Mrksich, M., Wade, W. S., Dwyer, T. J., Geierstanger, B. H., Wemmer, D. E., & Dervan, P. B. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89,* 7586.
Mrksich, M., & Dervan, P. B. (1995) *J. Am. Chem. Soc. 117,* 3325.
Mrksich, M., Parks, M. E., & Dervan, P. B. (1994) *J. Am. Chem. Soc. 116,* 7983.
Narayana, N., Ginell, S. L., Russu, I. M., & Berman, H. M. (1991) *Biochemistry 30,* 4449.
Parks, M. E., Baird, E. E., & Dervan, P. B. (1996a) *J. Am. Chem. Soc. 118,* 6147.
Parks, M. E., Baird, E. E., & Dervan, P. B. (1996b) *J. Am. Chem, Soc. 118,* 6153.
Pelton, J. G., & Wemmer, D. E. (1989) *Proc. Natl. Acad. Sci. USA 86,* 5723.
Pelton, J. G., & Wemmer, D. E. (1990) *J. Am. Chem. Soc. 112,* 1393.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning,* 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Trauger, J. W., Baird, E. E., & Dervan, P. B. (1996a) *Chem. Biol. 3,* 369.
Trauger, J. W., Baird, E. E., & Dervan, P. B. (1996b) *Nature 382,* 559.
Trauger, J. W., Baird, E. E., Mrksich, M., & Dervan, P. B. (1996c) *J. Am. Chem. Soc. 118,* 6160.
Wade, W. S., Mrksich, M., & Dervan, P. B. (1992) *J. Am. Chem. Soc. 114,* 8783.
Wade, W. S., Mrksich, M., & Dervan, P. B. (1993) *Biochemistry 32,* 11385.
Yanagi, K., Privé, G. G., & Dickerson, R. E. (1991) *J. Mol. Biol. 217,* 201.
Yoon, C., Privé, G. G., Goodsell, D. S., & Dickerson, R. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85,* 6332.

BI960744I