

Chapter 4A

Effects of Tail Linkages on the Sequence-Specific Recognition of DNA by Hairpin Polyamide Conjugates

This project was done in collaboration with James Puckett (Dervan group, Caltech).

James Puckett synthesized polyamides **6-10** and prepared plasmid pJWP6.

Abstract

Hairpin pyrrole-imidazole polyamides are synthetic ligands that bind the minor groove of DNA at subnanomolar concentrations with programmable sequence specificity. While the empirically derived “pairing rules” have been well characterized for unlinked polyamide cores, far less is known about the binding affinity and sequence specificity of hairpin polyamide conjugates, which are now commonly used for improved nuclear localization in biological experiments. To further our understanding of C-terminal tail linkage effects on sequence specificity, the equilibrium association constants of five hairpin polyamide conjugates $\text{ImImPyPy-(R)}^{\text{H}_2\text{N}\gamma}\text{-ImPyPyPy-}\beta\text{-C}_3\text{-R}_1$ (β = β -alanine, C_3 = 1,3-diaminopropane linker) were measured by quantitative DNase I footprint titration experiments on plasmids containing the eight possible 5'-ANTGGTCA-3' and 5'-NATGGTCA-3' binding sites (N = A, T, C, or G). In addition, the binding affinities of five hairpin polyamide conjugates $\text{ImImPyPy-(R)}^{\text{H}_2\text{N}\gamma}\text{-ImPyPyPy-(+)-R}_2$ ((+) = 3,3'-diamino-*N*-methyldipropylamine linker) were analyzed on the same two plasmids. For both types of compounds, the linker moiety exhibited 3- to 8-fold specificity for A,T over G,C, and the R group showed a similar 2- to 9-fold preference for A,T, with amide linkages conferring greater specificity than thiourea linkages. These results indicate that linkers and functional R groups on the tails of hairpin polyamide conjugates have recognition properties that should be considered in the design of these molecules to target DNA binding sites.

4.1. Introduction

Aberrant gene expression has been implicated as the cause of many diseases.^{1,2} Small molecules that can predictably regulate gene transcription could hold promise in human medicine.³ Hairpin pyrrole-imidazole polyamides are synthetic ligands that target predetermined DNA sequences with affinities comparable to those of naturally occurring DNA-binding proteins in the nanomolar range.^{4,5} These cell-permeable small molecules have been shown to localize to the nucleus of living cells⁶⁻¹⁰ and regulate endogenous gene expression.¹¹

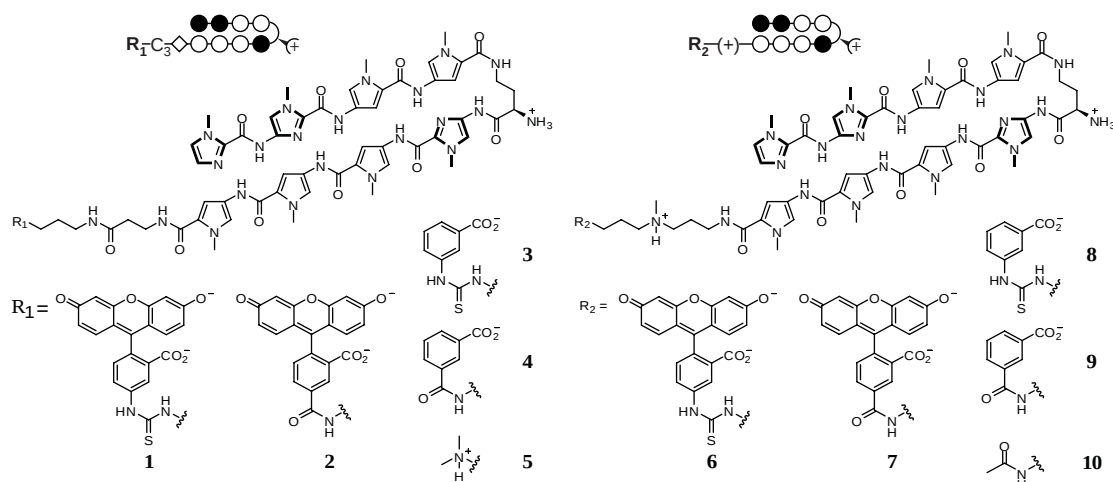
Polyamides selectively bind in the minor groove of B-form DNA according to a set of “pairing rules,” where each heterocyclic ring pair targets a specific Watson-Crick base pair.^{4,5} The antiparallel pairing of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) aromatic rings targets the C•G base pair (bp), while Im/Py discriminates G•C.^{12,13} The Py/Py pair recognizes A•T and T•A base pairs in the DNA minor groove.¹⁴ *N*-methylhydroxypyrrole (Hp) can be paired with Py to distinguish T•A from A•T.^{15,16} Structural data from NMR and X-ray crystallography studies have elucidated these binding preferences.^{12,14,17,18}

A covalent γ -aminobutyric acid (γ) linkage pre-organizes two polyamide subunits in the DNA minor groove, prepaying the entropic cost of antiparallel binding.¹⁹ Like the degenerate Py/Py pair, the γ residue at the hairpin “turn” is specific for A,T base pairs.²⁰ The use of chiral (*R*)-2,4-diaminobutyric acid to replace γ improves binding affinity and enforces a 5'→3' N→C binding orientation.²¹ As with the hairpin turn, a variety of different “tails” can be designed. When polyamides are synthesized on solid phase with Boc- β -Ala-PAM resin and cleaved with 3-dimethylamino-1-propylamine (Dp), the final compound contains a residual β -Dp tail at the C-terminus.²² Quantitative DNase I footprint titration experiments using the six-ring hairpin polyamide ImPyPy-(*R*)^{H₂N} γ -ImPyPy- β -Dp have demonstrated that both the β -alanine residue and the Dp tail are W-specific recognition elements, where W = A or T, at the N-1 position.²⁰ Oxime resin can be used to synthesize hairpin polyamides that do not contain β -alanine at the C-terminus and allow

rational targeting of more sites.²³ Interestingly, polyamide-fluorophore conjugates without β -alanine tend to have better nuclear uptake properties than analogs with β -alanine.^{9,10}

Hairpin polyamide cores can be functionalized at the turn and the tail to form polyamide conjugates suitable for a number of different applications. In particular, polyamide-fluorophore conjugates have been used to assay cellular uptake and nuclear localization in live cells using confocal microscopy.⁶⁻¹⁰ In one notable instance, such a compound has been shown to downregulate endogenous gene expression in cell culture.¹¹ However, there is no systematic data comparing the binding affinities and specificities of polyamide conjugates with their parent molecules at the N-2 and N-3 positions relative to the Py/Im polyamide core recognition element. Increased understanding of the effects on DNA binding would enable rational polyamide conjugate design in future biological studies.

In this study, we examine the effects of substitution at the tail on DNA affinity, specificity, and cellular uptake (Figure 4.1). The ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-R polyamide core sequence was used in all ten compounds.¹⁶ In the first series, five molecules of the type ImImPyPy-(R)^{H₂N} γ -ImPyPyPy- β -C₃-R₁ (β = β -alanine, C₃ = 1,3-diaminopropane linker) were synthesized using Boc- β -Ala-PAM resin. Fluorescein and 3-carboxyphenyl were conjugated to the polyamide using both thiourea and amide linkages. In the second series, five analogous conjugates of the type ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-(+)-R₂ ((+) = 3,3'-diamino-*N*-methyldipropylamine linker) were synthesized using oxime resin. The DNA binding energetics of all ten compounds were assayed by DNase I footprinting experiments on plasmids containing the eight possible 5'-ANTGGTCA-3' and 5'-NATGGTCA-3' binding sites (**N** = A, T, C, or G).



	R ₁	Linkage		R ₂	Linkage
	1 fluorescein	thiourea		6 fluorescein	thiourea
	2 fluorescein	amide		7 fluorescein	amide
	3 carboxyphenyl	thiourea		8 carboxyphenyl	thiourea
	4 carboxyphenyl	amide		9 carboxyphenyl	amide

Figure 4.1. Structures of polyamide conjugates **1-10**. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; β -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C₃”; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

4.2. Materials and methods

4.2.1. Polyamide synthesis

Polyamides were synthesized with Boc- β -Ala-PAM resin (compounds **1-5**, Peptides International) or oxime resin (compounds **6-10**, Novabiochem) according to published manual solid-phase synthesis protocols.^{22,23} The protected ^{Fmoc}HN γ -turn amine was deprotected with 20% piperidine in DMF and reprotected as the Boc derivative with a solution of Boc₂O (Fluka) and DIEA in DMF. The Boc-protected resin was cleaved with 1 mL of the appropriate amine (1,3-diaminopropane, compounds **1-4**; 3-dimethylamino-1-propylamine, compound **5**; 3,3'-diamino-*N*-methyldipropylamine, compounds **6-10**) at

37°C with agitation for 16 h. Products were purified by preparatory reverse-phase high-performance liquid chromatography (HPLC) on a Beckman Gold system using either a Waters Delta-Pak 25 × 100 mm, 15 μm 300 Å C₁₈ PrepPak Cartridge reverse-phase column or a Varian Dynamax 21.4 × 250 mm Microsorb 8 μm 300 Å C₈ reverse-phase column in 0.1% (w/v) TFA with acetonitrile as the eluent. The appropriate fractions were lyophilized after characterization by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization/time-of-flight mass spectrometry (MALDI-TOF-MS) and/or electrospray ionization mass spectrometry (ESI-MS).

Conjugates were formed by reacting the isothiocyanate (compounds **1**, **3**, **6**, and **8**), succinimidyl ester (compounds **2** and **7**), or anhydride (compound **10**) with the polyamide in a solution of DIEA (20 equiv) and DMF for 1 h at room temperature. Conjugates **4** and **9** were formed by preactivating isophthalic acid (3.0 equiv) with PyBOP (2.9 equiv, Novabiochem) in a solution of DIEA (20 equiv) and DMF at 37°C for 30 min, followed by reaction of the activated solution with the polyamide for 1 h at room temperature. Conjugates were deprotected with neat TFA (Halocarbon) and triethylsilane for 30 min at room temperature before purification by preparatory reverse-phase HPLC. Lyophilization of the appropriate fractions yielded the polyamide conjugates **1-10**, which were characterized as described above. Extinction coefficients were calculated according to standard protocols.²⁴ Chemicals not otherwise specified were from Aldrich.

ImImPyPy-(R)^{H₂N}_γ-ImPyPyPy-β-C₃-FITC (1). Cleaved from β-Ala-PAM resin with neat 1,3-diaminopropane and conjugated with fluorescein-5-isothiocyanate (FITC, Invitrogen). UV-vis λ_{max} = 312, 444 nm; ESI-MS *m/z* 1597.5 (C₇₆H₇₇N₂₄O₁₅S⁻ calculated [M-H]⁻ 1597.57)

ImImPyPy-(R)^{H₂N}_γ-ImPyPyPy-β-C₃-FAM (2). Cleaved from β-Ala-PAM resin with neat 1,3-diaminopropane and conjugated with 5-carboxyfluorescein, succinimidyl ester (5-FAM, SE, Invitrogen). UV-vis λ_{max} = 315, 444 nm; MALDI-TOF-MS *m/z* 1568.74 (C₇₆H₇₈N₂₃O₁₆⁺ calculated [M+H]⁺ 1568.60)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy- β -C₃-CPITC (3). Cleaved from β -Ala-PAM resin with neat 1,3-diaminopropane and conjugated with 3-carboxyphenyl isothiocyanate (Trans World Chemicals). UV-vis $\lambda_{\max}$ = 316 nm; ESI-MS m/z 1387.6 (C₆₃H₇₁N₂₄O₁₂S⁻ calculated [M-H]⁻ 1387.54)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy- β -C₃-IPA (4). Cleaved from β -Ala-PAM resin with neat 1,3-diaminopropane and conjugated with isophthalic acid. UV-vis $\lambda_{\max}$ = 315 nm; MALDI-TOF-MS m/z 1358.66 (C₆₃H₇₂N₂₃O₁₃⁺ calculated [M+H]⁺ 1358.57)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy- β -Dp (5). Cleaved from β -Ala-PAM resin with neat 3-dimethylamino-1-propylamine. UV-vis $\lambda_{\max}$ = 315 nm; MALDI-TOF-MS m/z 1238.88 (C₅₇H₇₂N₂₃O₁₀⁺ calculated [M+H]⁺ 1238.58)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-(+)-FITC (6). Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with fluorescein-5-isothiocyanate (FITC, Invitrogen). UV-vis $\lambda_{\max}$ = 316, 444 nm; ESI-MS m/z 1597.6 (C₇₇H₈₁N₂₄O₁₄S⁻ calculated [M-H]⁻ 1597.61)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-(+)-FAM (7). Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with 5-carboxyfluorescein, succinimidyl ester (5-FAM, SE, Invitrogen). UV-vis $\lambda_{\max}$ = 316, 444 nm; MALDI-TOF-MS m/z 1568.52 (C₇₇H₈₂N₂₃O₁₅⁺ calculated [M+H]⁺ 1568.64)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-(+)-CPITC (8). Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with 3-carboxyphenyl isothiocyanate (Trans World Chemicals). UV-vis $\lambda_{\max}$ = 316 nm; ESI-MS m/z 1387.4 (C₆₄H₇₅N₂₄O₁₁S⁻ calculated [M-H]⁻ 1387.58)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-(+)-IPA (9). Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with isophthalic acid. UV-vis $\lambda_{\max}$ = 316 nm; MALDI-TOF-MS m/z 1358.31 (C₆₄H₇₆N₂₃O₁₂⁺ calculated [M+H]⁺ 1358.60)

ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-(+)-Ac (10). Cleaved from oxime resin with neat 3,3'-diamino-*N*-methyldipropylamine and conjugated with acetic anhydride. UV-vis $\lambda_{\max}$

= 316 nm; MALDI-TOF-MS m/z 1252.55 ($C_{58}H_{74}N_{23}O_{10}^+$ calculated $[M+H]^+$ 1252.60)

4.2.2. Plasmid construction

Plasmids pCFH6 and pJWP6 were constructed according to standard protocols for DNA manipulation.²⁵ Plasmid pCFH6 was constructed by annealing the two 87-mer DNA oligonucleotides 5'-GATCGTGTAATCAATGGTCATAGCTGTGTAATCATTGGTCATAGCTGTGT AATCACTGGTCATAGCTGTGTAATCAGTGGTCATAGC-3' and 5'-AGCTGCTATGACCACTGATTACACAGCTATGACCAGTGATTACACAGCTA TGACCAATGATTACACAGCTATGACCATTGATTACAC-3' (Integrated DNA Technologies). Plasmid pJWP6 was constructed by annealing the two 91-mer DNA oligonucleotides 5'-GATCGTGTAATCAAATGGTCATAGCTGTGTAATCATATGGTCATAGCTGT GTAATCACATGGTCATAGCTGTGTAATCAGATGGTCATAGC-3' and 5'-AGCTGCTATGACCATCTGATTACACAGCTATGACCATGTGATTACACAGC TATGACCATATGATTACACAGCTATGACCATTGATTACAC-3'. Annealed oligonucleotides were ligated into the BamHI/HindIII (Roche) restriction fragment of pUC19 (Sigma). DNA sequencing of the constructed plasmids was performed at the Sequence Analysis Facility at the California Institute of Technology.

4.2.3. Preparation of 5'-³²P-end-labeled DNA

The forward primer 5'-AATTCGAGCTCGGTACCCGGG-3' was ³²P-labeled at the 5'-end. The reverse primer 5'-CTGGCACGACAGGTTTCCCGAC-3' was used to amplify plasmids pCFH6 and pJWP6 as previously described.²⁴ PCR products (5'-³²P-end-labeled, 291 bp for pCFH6, 295 bp for pJWP6) were isolated according to standard protocols.²⁴

4.2.4. Quantitative DNase I footprint titrations

Quantitative DNase I footprint titration experiments were performed on the 5'-³²P-end-labeled PCR products of plasmids pCFH6 and pJWP6 according to standard protocols.²⁴ Radiolabeled DNA was equilibrated with polyamide solutions for 14-16 h at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂ at pH 7.0 prior to DNase I cleavage. Chemical sequencing reactions were performed according to published methods.^{26,27}

4.2.5. Cell cultures

Human cancer cell lines HeLa, MCF7, and PC3 were cultured in a 5% CO₂ atmosphere at 37°C in supplemented DMEM (HeLa, GIBCO) or RPMI medium 1640 (MCF7 and PC3, GIBCO).^{9,10} All media were supplemented with 10% fetal bovine serum (Omega Scientific) and 1% penicillin/streptomycin solution (Mediatech).

4.2.6. Confocal microscopy

Cell lines were trypsinized (Mediatech) for 5 min at 37°C, centrifuged for 10 min at 4°C at 100 *g*, and resuspended in fresh medium to a concentration of 1.33×10^5 cells/mL (HeLa) or 3.33×10^5 cells/mL (MCF7 and PC3).^{9,10} Incubations were performed by adding 150 µL of cells into culture dishes equipped with glass bottoms for direct imaging (MatTek). Cells were grown in the glass-bottom culture dishes for 24 h. The medium was then removed and replaced with 147 µL of fresh medium, followed by addition of 3 µL of the 100 µM polyamide solution for a final polyamide concentration of 2 µM. Cells were incubated in a 5% CO₂ atmosphere at 37°C for 12 h. Imaging was performed with a 40× oil-immersion objective lens on a Zeiss LSM 5 Pascal inverted laser scanning microscope. Polyamide-fluorescein conjugate fluorescence and visible light images were obtained using standard filter sets for fluorescein.^{9,10}

4.3. Results

4.3.1. DNA binding affinity and sequence specificity of polyamide conjugates

Ten hairpin polyamide conjugates were synthesized by solid-phase methods (Figure 4.1). Compounds **1-5**, synthesized on β -Ala-PAM resin, were of the type ImImPyPy-(R)^{H₂N} γ -ImPyPyPy- β -C₃-R₁ (β = β -alanine, C₃ = 1,3-diaminopropane linker). Compounds **6-10**, synthesized on oxime resin, were of the type ImImPyPy-(R)^{H₂N} γ -ImPyPyPy-(+)-R₂ ((+) = 3,3'-diamino-*N*-methyldipropylamine linker). Each series of five compounds contained thiourea-linked fluorescein, amide-linked fluorescein, thiourea-linked 3-carboxyphenyl, amide-linked 3-carboxyphenyl, and control (Dp or Ac) conjugates.

Plasmid pCFH6 was designed to contain four binding sites of the form 5'-ANTGGTCA-3' (N = A, T, C, or G) (Figure 4.2). The four Watson-Crick base pairs were permuted at the 5'-end of the 7-bp binding site to test sequence specificity at the N-2 position of the C-terminal tail. Following the results of footprinting experiments on plasmid pCFH6 (Figure 4.3 and Table 4.1), plasmid pJWP6 was designed to contain four binding sites of the form 5'-NATGGTCA-3' with permutation at the 5'-end of the 8-bp binding site to test sequence specificity at the N-3 position of the C-terminal tail (Figure 4.2).

Quantitative DNase I footprint titration experiments were performed on the 291 bp, 5'-³²P-end-labeled PCR product of plasmid pCFH6 and the 295 bp, 5'-³²P-end-labeled PCR product of pJWP6.²⁴ Each polyamide was assayed on a plasmid containing four potential match sites in which a single Watson-Crick base pair was permuted. On plasmid pCFH6, the thiourea-linked polyamide-fluorescein conjugate **1** bound the 7-bp binding sites 5'-ATGGTCA-3' and 5'-TTGGTCA-3' with equilibrium association constants of $6.3 \times 10^9 \text{ M}^{-1}$ and $5.1 \times 10^9 \text{ M}^{-1}$, respectively (Figure 4.3 and Table 4.1). Compound **1** bound the other two designed sites (C•G and G•C) with $K_a = 1.8 \times 10^9 \text{ M}^{-1}$, indicating that this molecule exhibits 4-fold specificity for A•T and 3-fold specificity for T•A. The amide-linked polyamide-fluorescein conjugate **2** showed decreased binding affinity but strong sequence specificity, as its equilibrium association constant of $3.2 \times 10^9 \text{ M}^{-1}$ displayed an

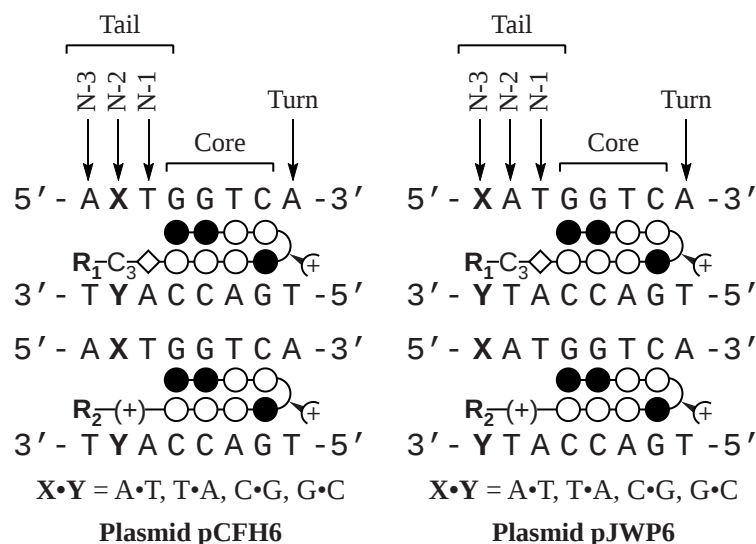


Figure 4.2. Designed binding sites on plasmids pCFH6 and pJWP6. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; β -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C₃”; the 3,3’-diamino-*N*-methyldipropylamine linker is shown as “(+)”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

8-fold preference for A•T over C,G. Both of the carboxyphenyl conjugates **3** and **4** bound their match sites with a binding affinity of greater than $1 \times 10^{10} \text{ M}^{-1}$ and exhibited similar 4-fold specificity. As expected, the smaller carboxyphenyl conjugates bound more tightly than fluorescein conjugates. The binding profiles of compounds **3** and **4** were very similar even though these conjugates were created with different linkages. The control compound **5** had the highest equilibrium association constant in the series, as its $K_a = 3.2 \times 10^9 \text{ M}^{-1}$ for the A•T site was 5-fold higher than the K_a for the C,G sites. The A,T preference seen in **5** is consistent with the trend observed for the analogous six-ring polyamide ImImPy-(R)^{H₂N} γ -ImPyPy- β -Dp, which exhibited 23-fold and 11-fold specificity for A•T and T•A, respectively.²⁰ The binding energetics data for compounds **1-4** demonstrate that the 1,3-diaminopropane linker also displays A,T specificity.

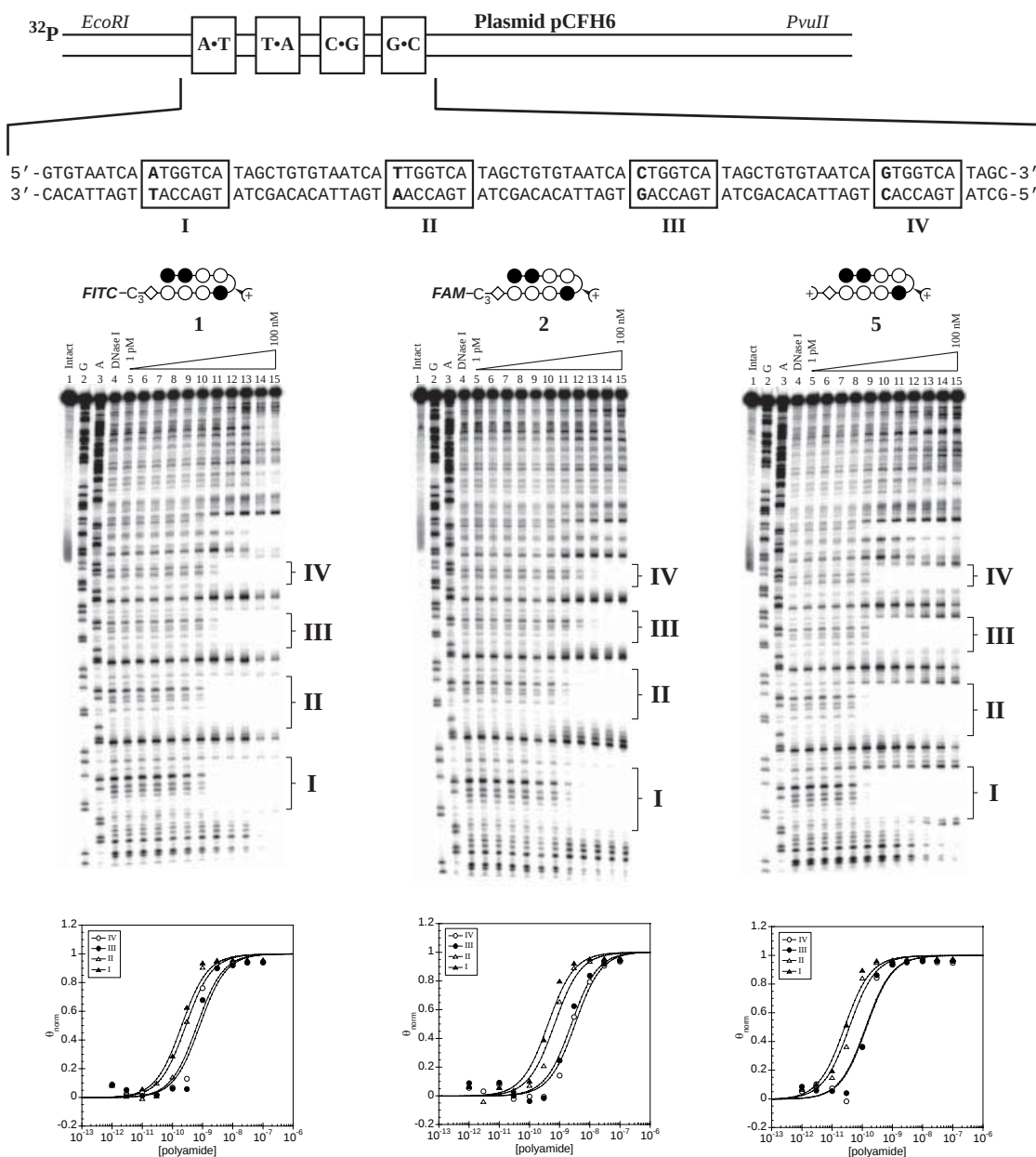


Figure 4.3. Quantitative DNase I footprint titration experiments for polyamide conjugates 1, 2, and 5 on the 291 bp, 5'-³²P-end-labeled PCR product of plasmid pCFH6: lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-15, 1 pM, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, and 100 nM polyamide, respectively. Binding isotherms for the four designed sites are shown below each footprinting gel; θ_{norm} values were calculated according to published methods.²⁴ Imidazole and pyrrole are shown as filled and non-filled circles, respectively; β -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as "C₃"; the dimethylaminopropylamine tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

Table 4.1. Equilibrium association constants^a

pCFH6	R	Linkage	5'-AATGGTCA-3'	5'-ATTGGTCA-3'	5'-ACTGGTCA-3'	5'-AGTGGTCA-3'
1	fluorescein	thiourea	6.3 (± 1.9) $\times 10^9$ [4]	5.1 (± 1.7) $\times 10^9$ [3]	1.7 (± 0.5) $\times 10^9$ [1]	1.8 (± 0.5) $\times 10^9$ [1]
2	fluorescein	amide	3.2 (± 1.1) $\times 10^9$ [8]	2.1 (± 0.6) $\times 10^9$ [5]	4.0 (± 1.2) $\times 10^8$ [1]	4.1 (± 1.0) $\times 10^8$ [1]
3	carboxyphenyl	thiourea	1.3 (± 0.5) $\times 10^{10}$ [4]	1.2 (± 0.5) $\times 10^{10}$ [3]	3.7 (± 1.8) $\times 10^9$ [1]	4.1 (± 1.8) $\times 10^9$ [1]
4	carboxyphenyl	amide	1.5 (± 0.4) $\times 10^{10}$ [4]	1.2 (± 0.4) $\times 10^{10}$ [3]	4.1 (± 1.4) $\times 10^9$ [1]	4.8 (± 1.9) $\times 10^9$ [1]
5	—	—	3.7 (± 0.5) $\times 10^{10}$ [5]	2.5 (± 0.1) $\times 10^{10}$ [3]	8.7 (± 1.4) $\times 10^9$ [1]	8.0 (± 0.7) $\times 10^9$ [1]
6	fluorescein	thiourea	1.7 (± 0.2) $\times 10^{10}$ [3]	1.6 (± 0.2) $\times 10^{10}$ [3]	6.8 (± 0.3) $\times 10^9$ [1]	6.0 (± 0.7) $\times 10^9$ [1]
7	fluorescein	amide	6.5 (± 1.0) $\times 10^9$ [7]	6.6 (± 1.2) $\times 10^9$ [7]	2.0 (± 0.3) $\times 10^9$ [2]	9.7 (± 2.0) $\times 10^8$ [1]
8	carboxyphenyl	thiourea	8.8 (± 1.3) $\times 10^9$ [3]	8.8 (± 1.0) $\times 10^9$ [3]	3.5 (± 1.1) $\times 10^9$ [1]	2.6 (± 0.5) $\times 10^9$ [1]
9	carboxyphenyl	amide	1.0 (± 0.2) $\times 10^{10}$ [3]	1.0 (± 0.2) $\times 10^{10}$ [3]	5.1 (± 0.4) $\times 10^9$ [1]	3.3 (± 0.4) $\times 10^9$ [1]
10	acetyl	amide	3.1 (± 0.5) $\times 10^{10}$ [3]	2.9 (± 0.3) $\times 10^{10}$ [3]	1.5 (± 0.1) $\times 10^{10}$ [1]	1.1 (± 0.1) $\times 10^{10}$ [1]

^a K_a (M^{-1}) values reported are the mean values from at least three DNase I footprint titration experiments. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl₂, and 5 mM CaCl₂ at pH 7.0. Specificity is indicated in brackets and calculated as K_a (site) / K_a (min).

Compounds **6-10** have an additional positive charge due to the protonation of the 3,3'-diamino-*N*-methyldipropylamine linker at physiological pH. The thiourea-linked fluorescein conjugate **6** bound the A•T site with $K_a = 1.7 \times 10^{10} M^{-1}$ (3-fold specificity), and the amide-linked fluorescein conjugate **7** bound the T•A site with $K_a = 6.6 \times 10^9 M^{-1}$ (7-fold specificity). When compared with **1** and **2**, oxime compounds **6** and **7** bound DNA more tightly but with comparable specificity. Compounds **8** and **9** bound A,T sites with $K_a = 8.8 \times 10^9 M^{-1}$ and $1.0 \times 10^{10} M^{-1}$, respectively. The acetylated control compound **10** displayed an equilibrium association constant of $3.1 \times 10^{10} M^{-1}$, and, for all three of these compounds **8-10**, their 3-fold preference for A,T was slightly less than that shown by compounds **3-5**. These experiments indicate that the 3,3'-diamino-*N*-methyldipropylamine linker is A,T-specific; however, the linker likely spans the two base pairs 5'-WT-3'.

Based on the conclusion that both linkers specify for A,T base pairs, plasmid pJWP6 was designed to test whether the conjugated R group can serve as a recognition element, by placing an A•T base pair at the N-2 position and varying the N-3 base pair across the four Watson-Crick base pairs. Control polyamide **5** exhibited the lowest specificity for T•A ($4.5 \times 10^{10} \text{ M}^{-1}$) over G•C ($2.7 \times 10^{10} \text{ M}^{-1}$) (Figure 4.4 and Table 4.2). The other nine conjugates showed at least 2-fold specificity for A,T, and, as with the previous plasmid, amide-linked conjugates displayed stronger specificity than thiourea-linked compounds. One highlight was the amide-linked fluorescein compound **7**, whose equilibrium association constant of $6.0 \times 10^9 \text{ M}^{-1}$ for A•T was nearly an order of magnitude greater than the K_a of $6.8 \times 10^8 \text{ M}^{-1}$ for G•C. In general, rational design of the linker and R group leads to the recognition of an eighth base pair at the N-3 position.

4.3.2. Nuclear localization of polyamide-fluorescein conjugates

Human cancer cell lines HeLa, MCF7, and PC3 were cultured for cell uptake studies using confocal laser scanning microscopy. For compounds **1**, **6**, and **7**, the nuclear staining exceeded that of the medium in all cell lines (Table 4.3). For compound **2**, there was some nuclear staining in HeLa cells, although it was similar to that of the medium, and little nuclear staining was observed in MCF7 and PC3 cells. The cellular uptake data for compound **6** were previously reported.¹⁰ Because FITC and FAM were considered to be interchangeable previously, the results for compound **7** were expected. However, the positive nuclear localization data for the β -alanine-linked fluorescein conjugate **1** was a pleasant surprise, suggesting that 1,3-diaminopropane could be a viable linker for β -alanine-linked conjugates synthesized on PAM resin. The mixed results for compound **2** are a reminder that thiourea and amide linkages are often different in terms of cellular uptake as well as sequence specificity.

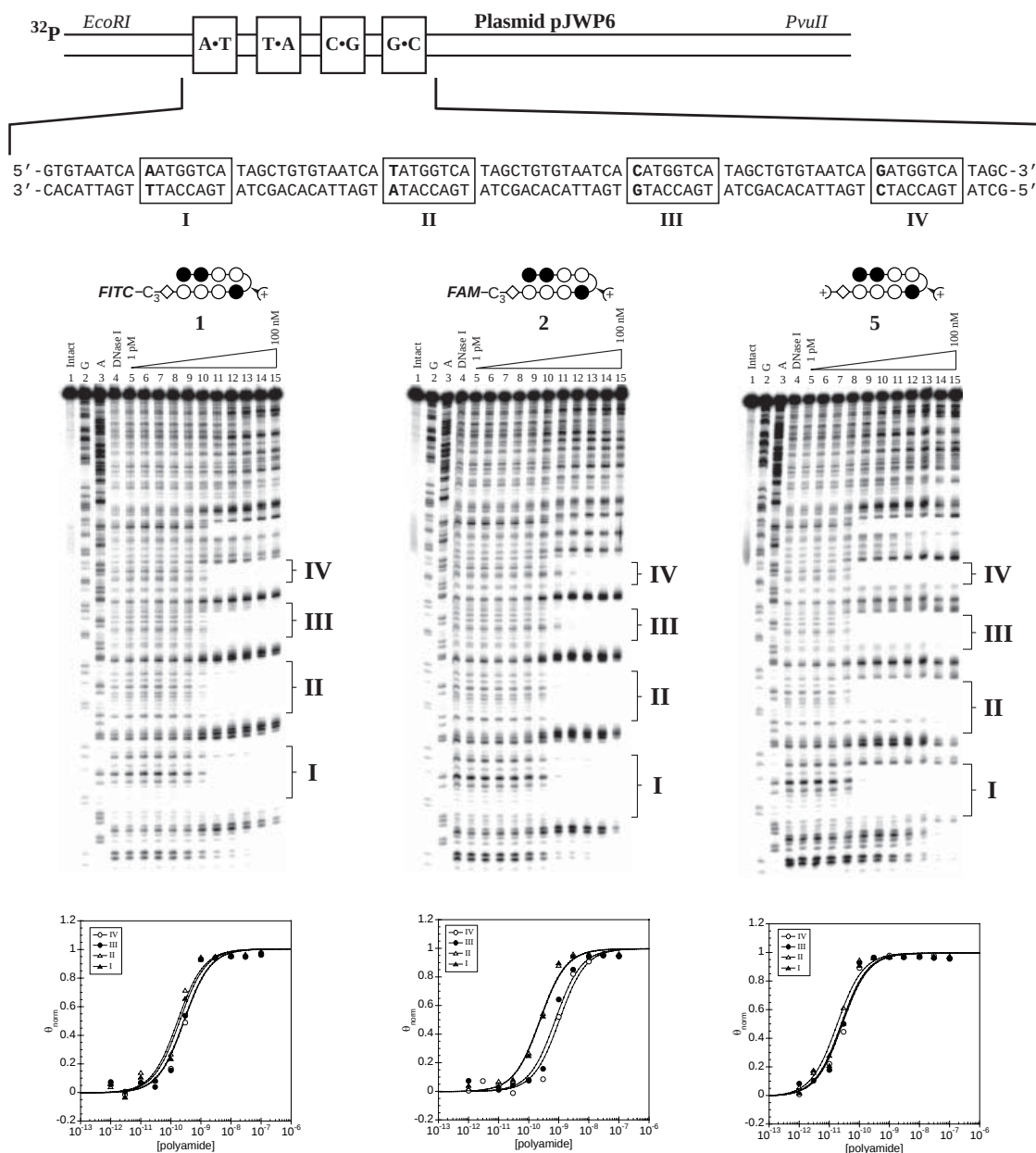


Figure 4.4. Quantitative DNase I footprint titration experiments for polyamide conjugates 1, 2, and 5 on the 295 bp, 5'- ^{32}P -end-labeled PCR product of plasmid pJWP6: lane 1, intact DNA; lane 2, G reaction; lane 3, A reaction; lane 4, DNase I standard; lanes 5-15, 1 pM, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, and 100 nM polyamide, respectively. Binding isotherms for the four designed sites are shown below each footprinting gel; θ_{norm} values were calculated according to published methods.²⁴ Imidazole and pyrrole are shown as filled and non-filled circles, respectively; β -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as "C₃"; the dimethylaminopropylamine tail is shown as a half-circle with a plus; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

Table 4.2. Equilibrium association constants^a

pJWP6	R	Linkage	5'-AATGGTCA-3'	5'-TATGGTCA-3'	5'-CATGGTCA-3'	5'-GATGGTCA-3'
1	fluorescein	thiourea	5.1 (± 0.8) $\times 10^9$ [2]	5.4 (± 1.1) $\times 10^9$ [2]	3.6 (± 0.7) $\times 10^9$ [1]	3.2 (± 0.6) $\times 10^9$ [1]
2	fluorescein	amide	3.7 (± 1.0) $\times 10^9$ [4]	3.5 (± 1.4) $\times 10^9$ [4]	1.2 (± 0.2) $\times 10^9$ [1]	8.8 (± 1.1) $\times 10^8$ [1]
3	carboxyphenyl	thiourea	1.4 (± 0.6) $\times 10^{10}$ [2]	1.5 (± 0.6) $\times 10^{10}$ [2]	8.3 (± 2.3) $\times 10^9$ [1]	7.1 (± 2.3) $\times 10^9$ [1]
4	carboxyphenyl	amide	1.5 (± 0.2) $\times 10^{10}$ [3]	1.5 (± 0.1) $\times 10^{10}$ [3]	6.1 (± 1.1) $\times 10^9$ [1]	4.7 (± 1.9) $\times 10^9$ [1]
5	–	–	3.1 (± 1.5) $\times 10^{10}$ [1]	4.5 (± 1.9) $\times 10^{10}$ [2]	3.0 (± 1.2) $\times 10^{10}$ [1]	2.7 (± 1.2) $\times 10^{10}$ [1]
6	fluorescein	thiourea	1.6 (± 0.4) $\times 10^{10}$ [2]	2.3 (± 0.4) $\times 10^{10}$ [3]	1.2 (± 0.2) $\times 10^{10}$ [1]	8.9 (± 1.1) $\times 10^9$ [1]
7	fluorescein	amide	6.0 (± 3.9) $\times 10^9$ [9]	4.9 (± 3.1) $\times 10^9$ [7]	2.1 (± 0.6) $\times 10^9$ [3]	6.8 (± 0.9) $\times 10^8$ [1]
8	carboxyphenyl	thiourea	1.9 (± 0.4) $\times 10^{10}$ [2]	2.2 (± 0.4) $\times 10^{10}$ [3]	1.1 (± 0.1) $\times 10^{10}$ [1]	8.3 (± 1.5) $\times 10^9$ [1]
9	carboxyphenyl	amide	3.2 (± 0.2) $\times 10^{10}$ [6]	2.2 (± 0.5) $\times 10^{10}$ [4]	1.1 (± 0.1) $\times 10^{10}$ [2]	5.1 (± 1.2) $\times 10^9$ [1]
10	acetyl	amide	3.6 (± 0.5) $\times 10^{10}$ [3]	3.2 (± 0.3) $\times 10^{10}$ [3]	1.4 (± 0.1) $\times 10^{10}$ [1]	1.1 (± 0.1) $\times 10^{10}$ [1]

^a K_a (M^{-1}) values reported are the mean values from at least three DNase I footprint titration experiments. Assays were performed at 22°C in a buffer of 10 mM Tris-HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0. Specificity is indicated in brackets and calculated as $K_a(\text{site}) / K_a(\text{min})$.

4.4. Discussion

DNase I footprinting experiments indicate that hairpin polyamide conjugates **1-4** and **6-9** have decreased binding affinity relative to their parent polyamide compounds **5** and **10**, respectively. This trend becomes more severe as the steric bulk of the conjugated R group increases, so fluorescein conjugates are impacted more than 3-carboxyphenyl conjugates. Control polyamides **5** and **10** bind their preferred match sites with an equilibrium association constant of $3 \times 10^{10} M^{-1}$, while 3-carboxyphenyl and fluorescein conjugates bind in the mid- $10^9 M^{-1}$ range. The subnanomolar binding affinities of these conjugates, coupled with good sequence specificity, make these compounds strong leads for continued biological applications.

Table 4.3. Nuclear localization of polyamide-fluorescein conjugates^a

	Compound	Linkage	HeLa	MCF7	PC3
	1	thiourea	++	++	++
	2	amide	+	–	–
	6	thiourea	++	++	++
	7	amide	++	++	++

^a Uptake profile of polyamide-fluorescein conjugates in three cell lines. ++, Nuclear staining exceeds that of the medium; +, nuclear staining less than or equal to that of the medium, but still prominent; –, very little nuclear staining, with the most fluorescence seen in the cytoplasm and/or medium. Imidazole and pyrrole are shown as filled and non-filled circles, respectively; β -alanine is shown as a diamond; the 1,3-diaminopropane linker is shown as “C₃”; the 3,3'-diamino-*N*-methyldipropylamine linker is shown as “(+)”; and the chiral diaminobutyric acid turn residue is shown as a semicircle connecting the two subunits linked to a half-circle with a plus.

For conjugates synthesized on Boc- β -Ala-PAM resin, it appears that the 1,3-diaminopropane linker is A,T-specific, with the C₃ linker extending out only one additional Watson-Crick base pair, as evidenced by the footprinting of control compound 5. For oxime compounds, the 3,3'-diamino-*N*-methyldipropylamine linker prefers 5'-WT-3' base pair steps over 5'-GT-3' or 5'-CT-3'. In general, β -C₃ and 3,3'-diamino-*N*-methyldipropylamine linkers lead to compounds with similar binding affinities, although they might contain different charges at physiological pH. Notably, amide-linked compounds have higher sequence specificity, and amide-linked fluorescein conjugates have an especially strong A,T preference. Thus, it is now possible to take advantage of the inherent W-specific recognition properties of linkers and R groups to design high-affinity, high-specificity DNA-binding compounds taking into account binding preferences at the N-2 and N-3 positions.

The positive cellular uptake data is highly encouraging as hairpin polyamide conjugates are used in cell culture experiments. The nuclear localization of compound **1** suggests that the β -alanine residue is not lethal when the appropriate amine linker is utilized, so PAM resin can be used in the synthesis of this class of conjugates going forward. The uptake of compounds **6** and **7** is consistent with the observed trend that conjugates synthesized on oxime resin (without the β -alanine residue) tend to localize well in human cancer cell lines. Having the ability to cleave from either resin with different linkers broadens the range of cell-permeable hairpin polyamide conjugates that are suitable for regulation of gene transcription.

Acknowledgements. We are grateful to the National Institutes of Health for research support. James Puckett synthesized polyamides **6-10** and prepared plasmid pJWP6.

References

1. Pandolfi, P.P. (2001) Transcription therapy for cancer. *Oncogene*, **20**, 3116-3127.
2. Darnell, J.E. (2002) Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer*, **2**, 740-749.
3. Arkin, M.R. and Wells, J.A. (2004) Small-molecule inhibitors of protein-protein interactions: Progressing towards the dream. *Nat. Rev. Drug Discov.*, **3**, 301-317.
4. Dervan, P.B. (2001) Molecular recognition of DNA by small molecules. *Bioorg. Med. Chem.*, **9**, 2215-2235.
5. Dervan, P.B. and Edelson, B.S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.*, **13**, 284-299.
6. Belitsky, J.M., Leslie, S.J., Arora, P.S., Beerman, T.A. and Dervan, P.B. (2002) Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. *Bioorg. Med. Chem.*, **10**, 3313-3318.
7. Gygi, M.P., Ferguson, M.D., Mefford, H.C., Lund, K.P., O'Day, C., Zhou, P., Friedman, C., van den Engh, G., Stolowitz, M.L. and Trask, B.J. (2002) Use of fluorescent sequence-specific polyamides to discriminate human chromosomes by microscopy and flow cytometry. *Nucleic Acids Res.*, **30**, 2790-2799.
8. Crowley, K.S., Phillion, D.P., Woodard, S.S., Schweitzer, B.A., Singh, M., Shabany, H., Burnette, B., Hippenmeyer, P., Heitmeier, M. and Bashkin, J.K.

- (2003) Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. *Bioorg. Med. Chem. Lett.*, **13**, 1565-1570.
9. Best, T.P., Edelson, B.S., Nickols, N.G. and Dervan, P.B. (2003) Nuclear localization of pyrrole-imidazole polyamide-fluorescein conjugates in cell culture. *Proc. Natl. Acad. Sci. USA*, **100**, 12063-12068.
 10. Edelson, B.S., Best, T.P., Olenyuk, B., Nickols, N.G., Doss, R.M., Foister, S., Heckel, A. and Dervan, P.B. (2004) Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res.*, **32**, 2802-2818.
 11. Olenyuk, B.Z., Zhang, G.J., Klco, J.M., Nickols, N.G., Kaelin, W.G. and Dervan, P.B. (2004) Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc. Natl. Acad. Sci. USA*, **101**, 16768-16773.
 12. Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. and Dervan, P.B. (1992) Antiparallel Side-by-Side Dimeric Motif for Sequence-Specific Recognition in the Minor Groove of DNA by the Designed Peptide 1-Methylimidazole-2-Carboxamide Netropsin. *Proc. Natl. Acad. Sci. USA*, **89**, 7586-7590.
 13. Wade, W.S., Mrksich, M. and Dervan, P.B. (1992) Design of Peptides That Bind in the Minor Groove of DNA at 5'-(a,T)G(a,T)C(a,T)-3' Sequences by a Dimeric Side-by-Side Motif. *J. Am. Chem. Soc.*, **114**, 8783-8794.

14. Pelton, J.G. and Wemmer, D.E. (1989) Structural Characterization of a 2-1 Distamycin A.D(Cgcaaattggc) Complex by Two-Dimensional Nmr. *Proc. Natl. Acad. Sci. USA*, **86**, 5723-5727.
15. Trauger, J.W., Baird, E.E. and Dervan, P.B. (1996) Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature*, **382**, 559-561.
16. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. and Dervan, P.B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *Nature*, **391**, 468-471.
17. Kielkopf, C.L., Baird, E.E., Dervan, P.D. and Rees, D.C. (1998) Structural basis for G center dot C recognition in the DNA minor groove. *Nat. Struct. Biol.*, **5**, 104-109.
18. Kielkopf, C.L., White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E., Dervan, P.B. and Rees, D.C. (1998) A structural basis for recognition of A center dot T and T center dot A base pairs in the minor groove of B-DNA. *Science*, **282**, 111-115.
19. Mrksich, M., Parks, M.E. and Dervan, P.B. (1994) Hairpin Peptide Motif - a New Class of Oligopeptides for Sequence-Specific Recognition in the Minor-Groove of Double-Helical DNA. *J. Am. Chem. Soc.*, **116**, 7983-7988.
20. Swalley, S.E., Baird, E.E. and Dervan, P.B. (1999) Effects of gamma-turn and beta-tail amino acids on sequence-specific recognition of DNA by hairpin polyamides. *J. Am. Chem. Soc.*, **121**, 1113-1120.

21. Herman, D.M., Baird, E.E. and Dervan, P.B. (1998) Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.*, **120**, 1382-1391.
22. Baird, E.E. and Dervan, P.B. (1996) Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.*, **118**, 6141-6146.
23. Belitsky, J.M., Nguyen, D.H., Wurtz, N.R. and Dervan, P.B. (2002) Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.*, **10**, 2767-2774.
24. Trauger, J.W. and Dervan, P.B. (2001) Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. *Methods Enzymol.*, **340**, 450-466.
25. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
26. Maxam, A.M. and Gilbert, W. (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. *Methods Enzymol.*, **65**, 499-560.
27. Iverson, B.L. and Dervan, P.B. (1987) Adenine Specific DNA Chemical Sequencing Reaction. *Nucleic Acids Res.*, **15**, 7823-7830.