Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 653–657

BIOORGANIC & MEDICINAL CHEMISTRY

Expanding the Recognition of the Minor Groove of DNA by Incorporation of β-Alanine in Hairpin Polyamides

Clay C. C. Wang, Ulf Ellervik and Peter B. Dervan*

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

Received 9 August 2000; accepted 16 October 2000

Abstract—In order to expand the recognition code by hairpin polyamides to include DNA sequences of the type 5′-CWWC-3′ two polyamides, PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**) and PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**) were synthesized which have in common an Py/Im pair in the terminal position for targeting C·G but differ with respect to internal placement of a β-alanine residue. The equilibrium association constants (K_a) were determined at four DNA sites which differ at a single common position, 5′-TNTACA-3′ (N = T, A, G, C). Quantitative DNase I footprint titration experiments reveal that the eight-ring hairpin PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**) binds the four binding sites with similar affinities, $K_a = 1.3–1.9 \times 10^{10}\ M^{-1}$ indicating that there is no preference for the position N. In contrast, a redesigned polyamide PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**) that places an internal flexible aliphatic β-alanine to the 5′-side of a key imidazole group bound the match site 5′-TCTACA-3′ with high affinity and good sequence discrimination (K_a(match) $= 4.9 \times 10^{10}\ M^{-1}$ and the single base pair mismatch sites with 5- to 25-fold lower affinity). These results expand the repertoire of sequences targetable by hairpins and emphasize the importance of β-alanine as a key element for minor groove recognition.

Introduction

Polyamides containing three aromatic amino acids, *N*-methylpyrrole (Py), *N*-methylimidazole (Im) and 3-hydroxy-1-methylpyrrole (Hp), bind as antiparallel pairs in the minor groove of DNA with affinities and specificities for predetermined sequences comparable to DNA-binding proteins.[1] Im/Py pair binds G·C, while a Py/Im pair recognizes C·G. A Py/Py pair binds both A·T and T·A base pairs in preference to G·C and C·G. A Hp/Py pair prefers T·A over A·T, C·G or G·C.[1]

Polyamide dimers linked by a γ turn unit creates a hairpin structure which sets the ring pairs unambiguously in register.[2] We have observed that hairpin polyamides with Im at the N-terminal position usually have higher affinities than the corresponding Py analogues.[1] The placement of an Im and Py in the first and eighth position of an eight-ring polyamide creates an Im/Py pair at the N- and C-terminal positions of the hairpin, biasing our choice of target DNA sequences beginning with a G·C base pair.[1]

We would like to explore ring pairings that would allow the DNA sequence repertoire to be broadened wherein the target sequence begins at a C·G base pair. We have shown that Py/Im pairs when placed *internal* in a hairpin distinguish C·G from G·C.[3] However, this has not been demonstrated for the terminal position in hairpin polyamides where the Im residue would necessarily be placed at the C-terminus. Replacement of a Py ring with a flexible β-alanine amino acid (β) has been demonstrated to be necessary to target GWG sequences by Im-β-Im.[4] It remains to be determined whether a flexible β-alanine linker would also be necessary to allow targeting of C·G in the terminal position. The ability to target DNA sequences such as 5′-CWWC-3′ in promoters will be an important step forward in our gene regulation studies.[1]

We describe here two hairpin polyamides with Py/Im at the N-terminal and C-terminal positions that should code for C·G according to the pairing rules (Fig. 1). The usual eight-ring hairpin PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**) is redesigned such that a β-alanine replaces a Py ring to the 5′ side of the Im residue, PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**) to test whether the flexible aliphatic β-residue resets the curvature of the hairpin pairs for optimal hydrogen bonding contact between N3 of Im and the exocyclic NH_2 of G. As a control, ImPyPyPy-(R)^{H_2N}γ-ImPyPyPy-β-Dp (**3**) was included to establish

*Corresponding author. Tel.: +1-626-356-6002; fax: +1-626-683-8753; e-mail: dervan@caltech.edu

0968-0896/01/$ - see front matter
PII: S0968-0896(00)00282-0

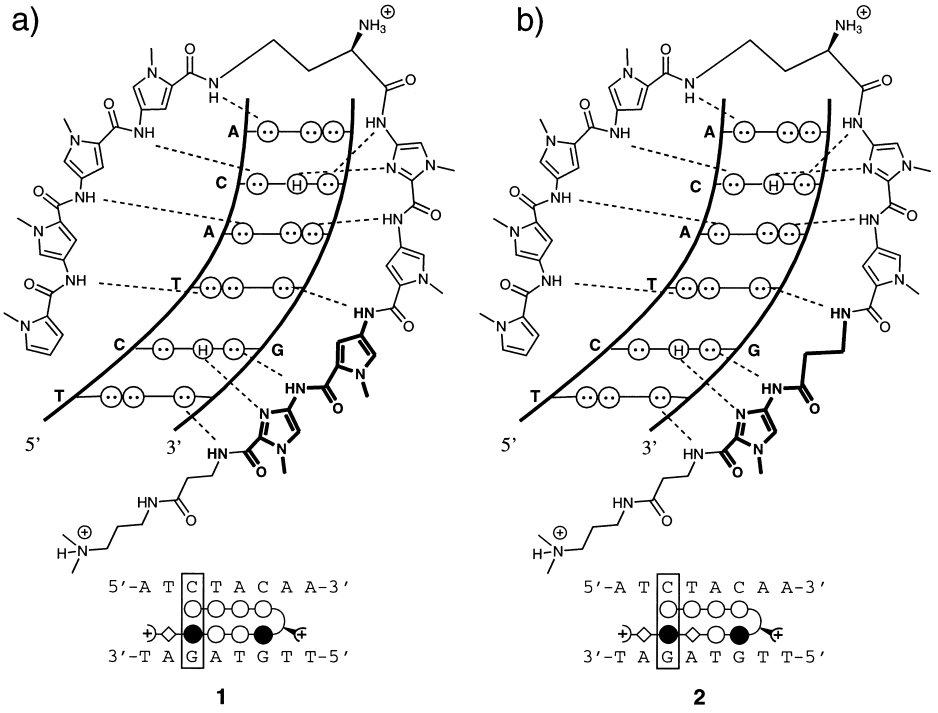

Figure 1. (a) Hydrogen bonding model of the polyamide:DNA complex between an eight-ring hairpin polyamide PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**) and a 5′-T**C**TACA-3′ site. A circle with two dots represents the lone pairs of N_3 of purines and O_2 of cytosine. Two touching circles with dots represent the two lone pairs of the O_2 of thymine. Circles containing an H represent the exocyclic amino hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. (b) Hydrogen bonding model of the polyamide:DNA complex between a seven-ring hairpin polyamide PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**) and a 5′-T**C**TACA-3′ site.

a baseline with previously published work (Fig. 2).[2b] Three hairpin polyamides (**1**–**3**) were synthesized by solid-phase methods.[6] The plasmid pCW15[5] was designed to contain the four six-base pair recognition sites 5′-T**T**TA CA-3′, 5′-T**A**TACA-3′, 5′-T**G**TACA-3′, and 5′-T**C**TA CA-3′, which differ at a single common position allowing for comparison of the affinities between different terminal pairs and the four Watson Crick base pairs in the minor groove of DNA (Fig. 3).

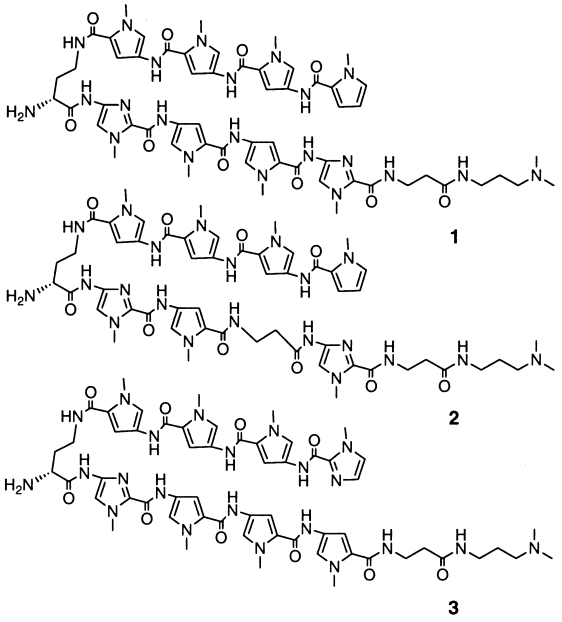

Figure 2. Structures of polyamides. PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**), PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**), ImPyPyPy-(R)^{H_2N}γ-ImPyPyPy-β-Dp (**3**).

Results

Polyamide synthesis

PyPyPyPy-(R)^{Fmoc}γ-ImPyPyIm-β-PAM-resin and Py PyPyPy-(R)^{Fmoc}γ-ImPy-β-Im-β-PAM-resin were synthesized in a stepwise manner from Boc-β-alanine-PAM resin (0.55 mmol/g) using solid-phase methodology[6] in 14 steps (Fig. 4). A sample of the resin was then cleaved by aminolysis with ((dimethyl)amino)propylamine (55 °C, 16 h) and purified by reversed-phase HPLC to provide PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**) and PyPyPy Py-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**) (Fig. 2).

Quantitative DNase I footprinting titrations

Quantitative DNase I footprint titrations on the 3′-^{32}P-end-labeled 284 bp pCW15 *Eco*RI/*Pvu*II restriction fragment (10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$, pH 7.0, 22 °C) (Fig. 5) reveal that Py PyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**) binds the four binding sites with similar affinities, $K_a = 1.3–1.9 \times 10^{10}$ M^{-1} indicating that there is no preference for any specific base pair (Table 1). In contrast polyamide **2** with a flexible aliphatic β-alanine next to the key imidazole group bound the 5′-T**C**TACA-3′ site with $K_a = 4.9 \times 10^{10}$ M^{-1} and the single base pair mismatched sites 5′-

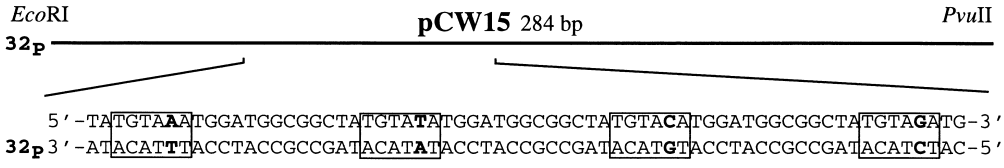

Figure 3. 284 base pair *Eco*RI/*Pvu*II restriction fragment derived from plasmid pCW15. The targeted six-base pair recognition sites are shown in boxes.

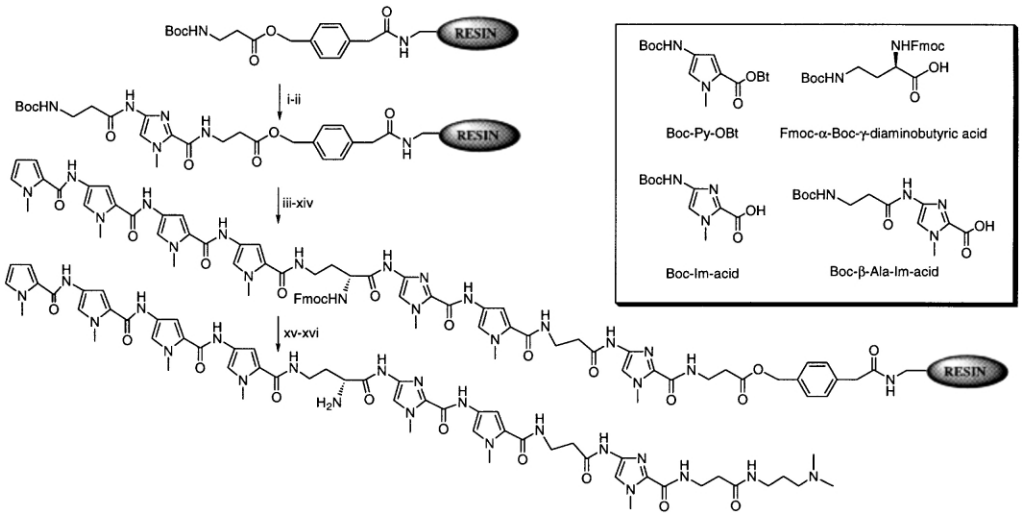

Figure 4. Solid-phase synthetic scheme for PyPyPyPy-(R)^{H_2N}γImPy-β-Im-β-Dp (**2**) starting from commercially available Boc-β-PAM-resin: (i) 80% TFA/DCM, 0.4 M PhSH; (ii) Boc-β-Ala-Im-acid, HBTU, DIEA, DMF; (iii) 80% TFA/DCM, 0.4 M PhSH; (iv) Boc-Py-OBt, DIEA, DMF; (iii) 80% TFA/DCM, 0.4 M PhSH; (iv) Boc-Im-acid, DCC, HOBt, DIEA, DMF; (v) 80% TFA/DCM, 0.4 M PhSH; (vi) Fmoc-α-Boc-γ-diaminobutyric acid, HBTU, DIEA DMF; (vii) 80% TFA/DCM, 0.4 M PhSH; (viii) Boc-Py-OBt, DIEA, DMF; (ix) 80% TFA/DCM, 0.4 M PhSH; (x) Boc-Py-OBt, DIEA, DMF; (xi) 80% TFA/DCM, 0.4 M PhSH; (xii) Boc-Py-OBt, DIEA, DMF; (xiii) 80% TFA/DCM, 0.4 M PhSH; (xiv) Py-acid, HBTU, DIEA, DMF; (xv) Piperidine: DMF 3:1; (xvi) (Dimethylamino)propylamine, 55 °C, 16 h.

T**G**TACA-3′, 5′-T**T**TACA-3′, and 5′-T**A**TACA-3′ with lower affinities ($K_a = 1.0\times10^{10}M^{-1}$, $K_a = 8.0\times10^{9}\,M^{-1}$ and $K_a = 2.0\times10^{9}\,M^{-1}$, respectively). The control compound ImPyPyPy-(R)^{H_2N}γ-ImPyPyPy-β-Dp (**3**) bound the 5′-T**G**TACA-3′ site with $K_a = 2.9\times10^{11}\,M^{-1}$ and the three single base pair mismatches at lower affinities ($K_a = 1.1$–$2.9\times10^{10}\,M^{-1}$).[5]

Discussion

The classic eight-ring hairpin PyPyPyPy-(R)^{H_2N}γ-ImPy PyIm-β-Dp (**1**) showed no preference for its designed matched binding site 5′-T**C**TACA-3′ over the three single base pair mismatches. The lack of discrimination is likely due to the mispositioning of the Im residue located at the C-terminal end of the four-ring polyamide subunit. Polyamide PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**) instead showed a 5- to 25-fold selectivity for the 5′-T**C**TACA-3′ site over the three single base pair mismatched sites. For comparison, the selectivity of the control polyamide **3** for its match site 5′-TGTACA-3′ versus single mismatch sites is 10- to 26-fold. The replacement of a pyrrole ring with a flexible β-alanine likely resets the imidazole ring to provide the optimal hydrogen bond formation between the imidazole N3 and the exocyclic amine of guanine. This study opens the opportunity to target sequences starting at a C·G base pair and validates the use of flexible β-alanine residue to reset the register of adjacent Im residues within the hairpin motif.[4] These results broaden the sequence repertoire for hairpin polyamides to 5′-CWWC-3′ and gives us an opportunity to target new biologically important DNA sequences in gene regulation studies.

Experimental

Boc-β-alanine-(4-carboxamidomethyl)-benzyl-ester-copoly (styrene-divinylbenzene) resin (Boc-β-PAM-resin), *N*,*N*′-dicyclohexylcarbodiimide (DCC), *N*-hydroxybenzotriazole (HOBt), 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexa-fluorophosphate (HBTU), and Boc-β-alanine were purchased from Peptides International. *N*,*N*-diisopropylethylamine (DIEA) and *N*,*N*-dimethylformamide (DMF), were purchased from Applied Biosystems. DMF was distilled under reduced pressure prior to synthesis. (*R*)-2-Fmoc-4-Boc-diaminobutyric acid was from Bachem, dichloromethane (DCM) was reagent grade from EM, thiophenol (PhSH) and dimethylaminopropylamine from Aldrich, and trifluoroacetic acid (TFA) from Halocarbon. ^{1}H NMR were recorded on a Varian Mercury 300 instrument. Chemical shifts are reported in ppm downfield from the signal for Me_4Si, with reference to the solvent residual signal. UV spectra were measured on a Hewlett-Packard Model 8452A diode array spectrophotometer. Matrix-assisted, laser desorption/ionization time of flight mass spectrometry was

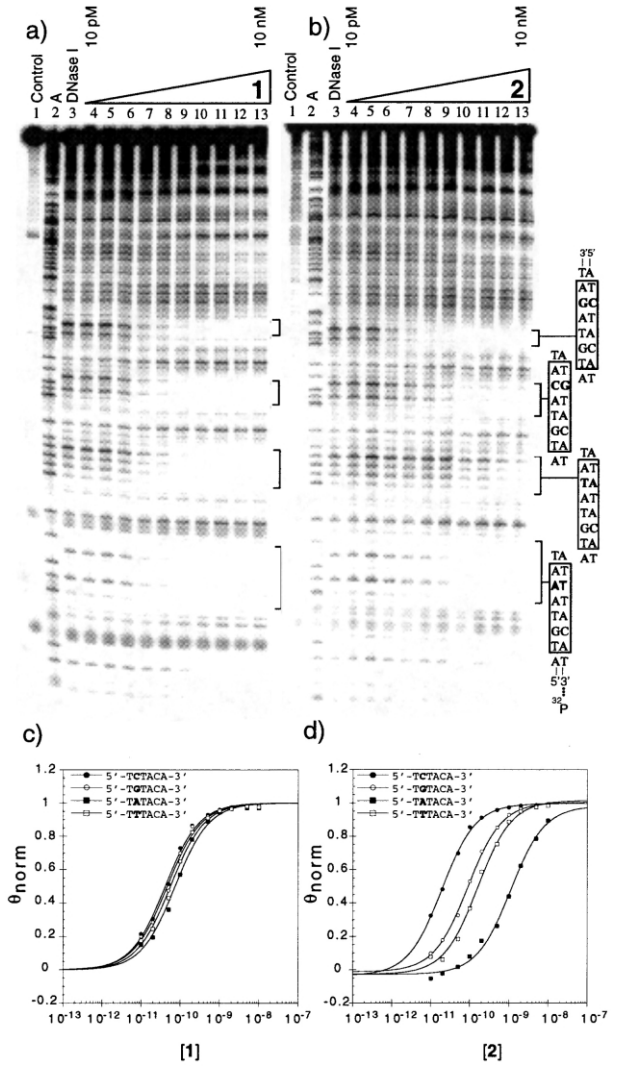

Figure 5. (a) Quantative DNase I footprint titration experiment PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**) on the 284 bp *Eco*RI/*Pvu*II restriction fragment from plasmid pCW15: lane 1, intact DNA; lane 2, A specific reaction; lane 3, DNase I standard; lanes 4–13, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (**1**); The four sites 5′-T**T**TACA-3′, 5′-T**A**TACA-3′, 5′-T**G**TACA-3′ and 5′-T**C**TACA-3′ sites that were analyzed are shown on the right side of the gel; (b) Quantative DNase I footprint titration PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**) on the 284 bp *Eco*RI/*Pvu*II restriction fragment from plasmid pCW15: lane 1, intact DNA; lane 2, A specific reaction lane 3, DNase I standard; lanes 4–13, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Im-β-Dp (**2**); The four sites 5′-T**T**TACA-3′, 5′-T**A**TACA-3′, 5′-T**G**TACA-3′ and 5′-T**C**TACA-3′ sites that were analyzed are shown on the right side of the gel; (c) Binding isotherms derived from the DNase I footprinting gels of polyamide **1** and the four different sites 5′-T**T**TACA-3′, 5′-T**A**TACA-3′, 5′-T**G**TACA-3′ and 5′-T**C**TACA-3′; (d) Binding isotherms derived from the DNase I footprinting gels of polyamide **2** and the four different sites 5′-T**T**TACA-3′, 5′-T**A**TACA-3′, 5′-T**G**TACA-3′ and 5′-T**C**TACA-3′.

carried out at the Protein and Peptide Microanalytical Facility at the California Institute of Technology. HPLC analysis was performed on a Beckman Gold system using a RAINEN C_{18}, Microsorb MV, 5 μm, 300×4.6 mm reversed-phase column in 0.1% (w/v) TFA with acetonitrile as eluent and a flow rate of 1.0 mL/min, gradient elution 1.25% acetonitrile/min. Preparatory HPLC was carried out on a Beckman HPLC using a Waters Delta-Pak 25×100 mm, 100 μm C_{18} column, 0.1% (w/v) TFA, 0.25% acetonitrile/min. 18 MΩ water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered. Reagent-grade chemicals were used unless otherwise stated.

PyPyPyPy-(R)^{Fmoc}γ-ImPyPyIm-β-PAM-resin and PyPyPyPy-(R)^{Fmoc}γ-ImPy-β-Im-β-PAM-resin. PyPyPyPy-(R)^{Fmoc}γ-ImPyPyIm-β-PAM-resin (0.333 mmol/g) and PyPyPyPy-(R)^{Fmoc}γ-ImPy-β-Im-β-PAM-resin (0.338 mmol/g) were synthesized in a stepwise fashion from 0.55 mmol/g Boc-β-PAM-resin by manual solid-phase methods.[6] (*R*)-2-Fmoc-4-Boc-diaminobutyric acid was incorporated as previously described for Boc-γ-aminobutyric acid.[2c,6]

Procedure for cleavage from the resin. After the coupling was completed the resin was filtered off the reaction mixture and washed with DMF (2×30 s). DMF (1 mL) and piperidine (3 mL) were added and the mixture shaken for 30 min at 22 °C. The resin was filtered off and washed with DMF (2×30 s), DCM (3×30 s), MeOH (1×30 s), Et_2O (1×30 s) and dried in vacuo. The resin was treated with (dimethylamino)propylamine (1 mL) with periodic agitation at 55 °C for 16 h. The reaction mixture was then filtered to remove the resin, TFA (0.1% (w/v), 7 mL) added and the resulting solution purified by reversed-phase HPLC. The pure compounds were recovered as white powders upon lyophilization of the appropriate fractions.

PyPyPyPy-(R)^{H_2N}γ-ImPyPyIm-β-Dp (1). PyPyPyPy-(R)^{Fmoc}γ-ImPyPyPy-β-PAM-resin (100 mg) was cleaved and purified by preparative HPLC to give **1** (13 mg, 31% isolated yield). UV λ_{max} (H_2O) 240, 312 (69500); ^{1}H NMR (DMSO-d_6) δ 10.97, 10.26, 10.07 (s, 1H each), 9.84–9.92 (m, 3H), 9.78 (s, 1H), 8.20–8.30 (m, 3H), 8.12 (t, 1H, J = 5.8 Hz), 8.03 (t, 2H, J = 6.3 Hz), 7.91 (t, 1H, J = 5.8 Hz), 7.48 (s, 1H), 7.44 (s, 1H), 7.20 (d, 1H, J = 1.2 Hz), 7.19 (d, 1H, J = 1.8 Hz), 7.16 (t, 2H, J = 1.5 Hz), 7.12 (d, 2H, J = 1.8 Hz), 7.06 (d, 1H, J = 1.8 Hz), 7.02 (d, 1H, J = 1.8 Hz), 6.99 (d, 1H, J = 1.8 Hz), 6.90 (d, 1H, J = 1.8 Hz), 6.89 (t, 1H, J = 2.1 Hz), 6.86 (dd, 1H, J = 2.1, 4.2 Hz), 5.99 (dd, 1H, J = 2.7, 4.2 Hz), 3.92, 3.87, 3.81, 3.80, 3.79, 3.78, 3.77, 3. 75 (s, 3H each), 3.04 (q, 2H, J = 6.3 Hz), 2.90–3.00 (m, 2H), 2.67 (d, 6H, J = 4.8 Hz), 2.30 (t, 2H,

Table 1. Equilibration association constants (M^{-1})[a]

Polyamide	5′-T**T**TACA-3′	5′-T**A**TACA-3′	5′-T**G**TACA-3′	5′-T**C**TACA-3′
1	1.7 (±0.3)×10^{10}	1.3 (±0.4)×10^{10}	1.9 (±0.5)×10^{10}	1.9 (±0.5)×10^{10}
2	8.0 (±2.0)×10^{9}	2.0 (±1.0)×10^{9}	1.0 (±0.4)×10^{10}	4.9 (±0.9)×10^{10}
3	2.0 (±0.5)×10^{10}	1.1 (±0.2)×10^{10}	2.9 (±0.5)×10^{11}	2.9 (±0.5)×10^{10}

[a]Values reported are mean values from at least three DNase I footprint titration experiments, with the standard deviation for each data set indicated in parenthesis. The assays were performed at 22 °C at pH 7.0 in the presence of 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

$J = 6.0$ Hz), 1.90–2.00 (m, 2H), 1.66 (bp, 2 H, $J =$ 7.8 Hz); MALDI-TOF-MS calcd for $C_{58}H_{73}N_{22}O_{10}$ (M + H): 1237.6; found: 1237.7.

PyPyPyPy-(R)^{H_2N}γ-ImPy-β-Py-β-Dp (2). PyPyPyPy-(R)^{Fmoc}γ-ImPy-β-Im-β-PAM-resin (100 mg) was cleaved and purified by preparative HPLC to give **2** (8 mg, 18% isolated yield). UV λ_{max} (H_2O) 240, 312 (60830); ^{1}H NMR (DMSO-d_6) δ 11.10, 10.44, 10.13, 10.03, 10.01, 9.92 (s, 1H each), 8.40, (bs, 2H), 8.26 (t, 1H, $J = 6.0$ Hz), 8.16 (t, 2H, $J = 6.0$ Hz), 8.04 (t, 1H, $J = 5.7$ Hz), 7.60 (s, 1H), 7.48 (s, 1H), 7.30 (t, 2H, $J = 1.5$ Hz), 7.29 (d, 1H, $J = 1.8$ Hz), 7.25 (d, 1H, $J = 1.5$ Hz), 7.15 (d, 1H, $J = 1.5$ Hz), 7.13 (d, 1H, $J = 1.5$ Hz), 7.04 (d, 1H, 1.5 Hz), 7.00–7.15 (m, 2H), 6.99 (dd, 1H, $J = 2.1$, 4.2 Hz), 6.13 (dd, 1H, $J = 2.4$, 3.9 Hz), 4.04, 3.98, 3.95, 3.93, 3.92, 3.88, 3.87 (s, 3H each), 3.30–3.40 (m, 2H), 3.17 (q, 2H, $J = 6.0$ Hz), 3.07 (p, 2H, $J =$ 4.8 Hz), 2.81 (d, 6H, $J = 5.1$ Hz), 2.63 (d, 2H, $J = 6.9$ Hz), 2.43 (t, 2H, $J = 6.9$ Hz), 2.00–2.10 (m, 2H), 1.80 (p, 2H, $J = 7.2$ Hz); MALDI-TOF-MS calcd for $C_{55}H_{72}N_{21}O_{10}$ (M + H): 1186.6; found: 1186.7.

Preparation of 3′- and 5′-end-labeled restriction fragments. The plasmid pCW15[5] was linearized with *Eco*RI and *Pvu*II and then treated with Sequenase (version 2.0 from United States Biochemical), deoxyadenosine 5′-[α–^{32}P] triphosphate and thymidine 5′-[α–^{32}P]triphosphate for 3′ labeling. The 3′ labeled fragment was loaded onto a 7% non-denaturing polyacrylamide gel, and the desired 284 base pair band was visualized by autoradiography and isolated. Chemical sequencing reactions were performed according to published methods.[7]

Quantitative DNase I footprinting. DNase I footprinting reactions were carried out as previously described.[2b] Photostimulable storage phosphorimaging plates (Kodak Storage Phosphor Screen S0230 obtained from Molecular Dynamics) were pressed flat against gel samples and exposed in the dark at 22 °C for 12–16 h. A Molecular Dynamics Typhoon PhosphorImager was used to obtain all data from the storage screens. The data were analyzed by performing volume integration of all bands using ImageQuant v. 3.2 software. All DNA manipulations were performed according to standard protocols.[8]

Acknowledgements

We are grateful to the National Institutes of Health (GM-27681) for research support, the Knut and Alice Wallenberg Foundation for a post-doctoral fellowship to U.E. and the National Institutes of Health for a research traineeship award (GM 08501) to C.C.C.W. We thank G. M. Hathaway for MALDI-TOF mass spectrometry.

References and Notes

1. Dervan, P. B.; Bürli, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688.
2. (a) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983. (b) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559. (c) Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 1382.
3. White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569.
4. Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 6219.
5. Ellervik, U.; Wang C. C. C.; Dervan P. B. *J. Am. Chem. Soc.* **2000**, *122*, 9354.
6. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.
7. (a) Iverson, B. L.; Dervan, P. B. *Nucl. Acids Res.* **1987**, *15*, 7823. (b) Maxam, A. M.; Gilbert, W. *Methods Enzymol.* **1980**, *65*, 499.
8. Sambrook, J.; Fritsch, E. F.; Maniatis, T. *Molecular Cloning*; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1989.