Appendix A

Protein-DNA Dimerizers Targeted to a Natural HOX Site

The part of this chapter regarding the DNase I footprinting was taken in part from a manuscript co-authored with Rocco Moretti,[‡] Leslie J. Donato,[‡] Mary L. Brezinski,[‡] Helena Hoff,[‡] Jon S. Thorson,[‡] Peter B. Dervan,[†] and Aseem Z. Ansari.[‡] (Caltech[†] and University of Wisconsin, Madison[‡])

(Moretti, R., Donato, L. J., Brezinski, M. L., Stafford, R. L., Hoff, H., Thorson, J. S., Dervan, P. B., Ansari, A. Z. *provisionally titled* Chemical wedges reveal the contribution of allosteric DNA modulation in protein-DNA assembly, **2007**, *In preparation*.)

*Work on the MPE footprinting of polyamide-peptide conjugates **1**-**5** was done in collaboration with Hans-Dieter Arndt (Dervan Group; Caltech).*

A.1 Targeting an Endogenous HOX Binding Site

All previous protein-DNA dimerizers for Exd do not target natural HOX binding sites.[1-5] Compound **1** was designed to target an endogenous labial response element 5′-TGATGGATTG-3′ (Figure A.1).[6] At this composite site, Exd is expected to bind 5′-TGAT-3′ and the polyamide DNA-binding domain of **1** was designed to target the site 5′-WGGWWW-3′[7] and project the peptide to the Exd binding pocket in a similar manner to conjugates **2** and **3** which are targeted to the artificial HOX site 5′-TGATTGACCT-3′. A control compound for **1** was also synthesized in which the crucial WM amino acids were replaced with alanines (**4**). Additionally, the parent polyamide that lacks a peptide altogether serves as a negative control (**5**).

DNase I footprinting of compounds **1**, **4**, and **5** was performed to determine their DNA binding affinities and specificities for their match site (Figure A.2 and Table A.1). Footprinting of compounds **2** and **3** have been reported elsewhere.[1,4] As previously observed with conjugate **2**, the labial FYPWMKG compound **1** showed only a modest affinity for its match site 5′-TGGATT-3′ ($K_a = 2.4 \times 10^8$ M^{-1}) with no significant sequence specificity. A WM to AA "mutation" of compound **2** was shown previously to lead to a recovery of the DNA binding affinity and specificity,[1] but a similar WM to AA change to the labial compound (i.e., **1** to **4**) did not lead to any significant change in binding affinity or specificity. The polyamide **5** which contains only a propylamine appendage showed an approximately 20-fold higher DNA-binding affinity and >10-fold specificity over single base pair mismatches. The parent polyamide ImImPyPy-γ-PyPyPyPy-β-Dp has previously been reported to bind to DNA with an affinity of 5×10^8 M^{-1}, but it is hard to compare this to **5** since it was footprinted on the match sequence 5′-TGGTTA-3′.[8]

In spite of the modest DNA affinity and lack of specificity of the labial conjugate **1**, MPE footprinting experiments indicate that **1** cooperatively binds with Exd specifically at the labial binding site (Figure A.3). That is, the titration of Exd in the presence of compound **1** leads to an enlargement of the observed protection pattern exclusively at the

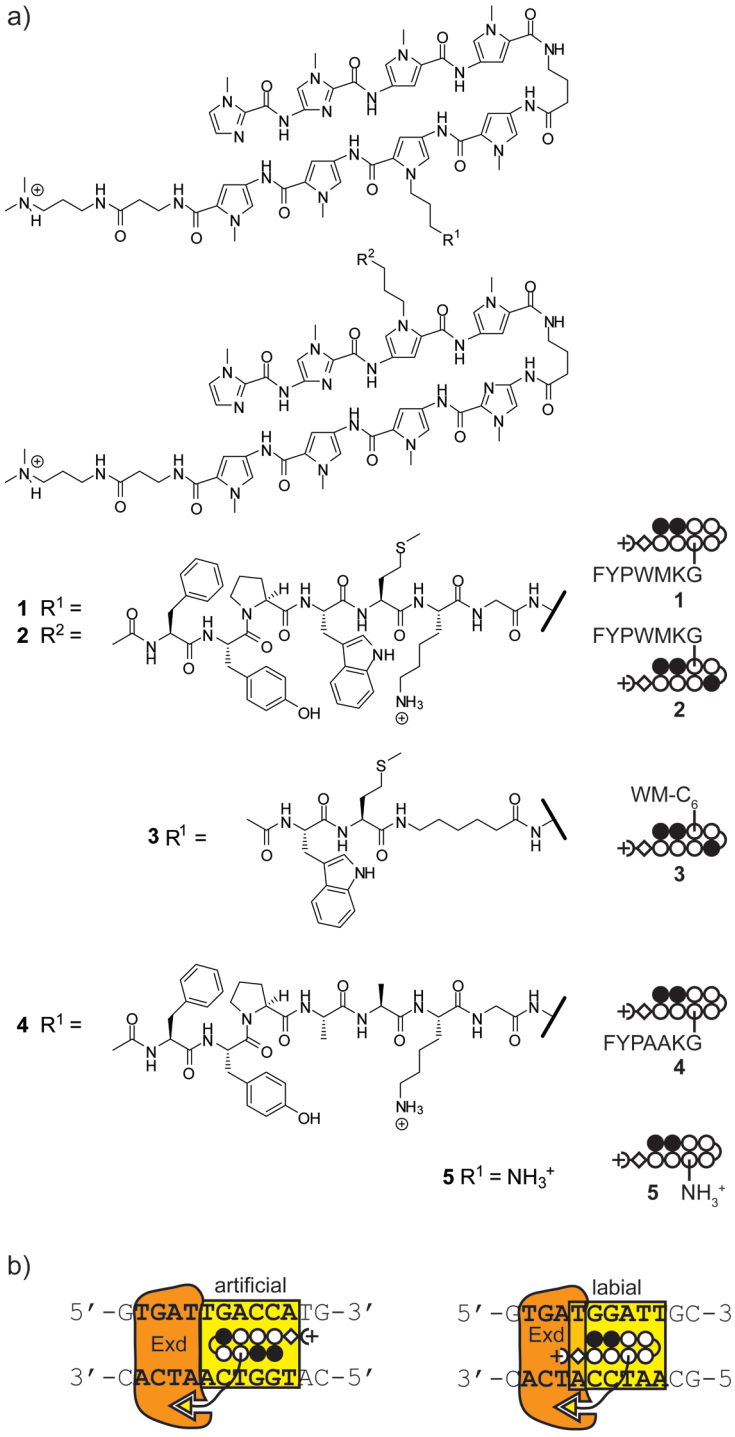


Figure A.1 Structures of polyamide-peptide conjugates targeted to artificial and natural HOX binding sites. (a) Conjugate **1** targets the natural labial HOX binding site (5′-TGA**T**G**GATT**G-3′) and **2** and **3** target the artificial site (5′-TGAT**TGACCT**-3′) where the Exd binding site is underlined and the polyamide binding site is shown in bold. Compounds **4** and **5** serve as negative controls. (b) Schematics showing the binding orientations of the artificial and natural cognate DNA sites.

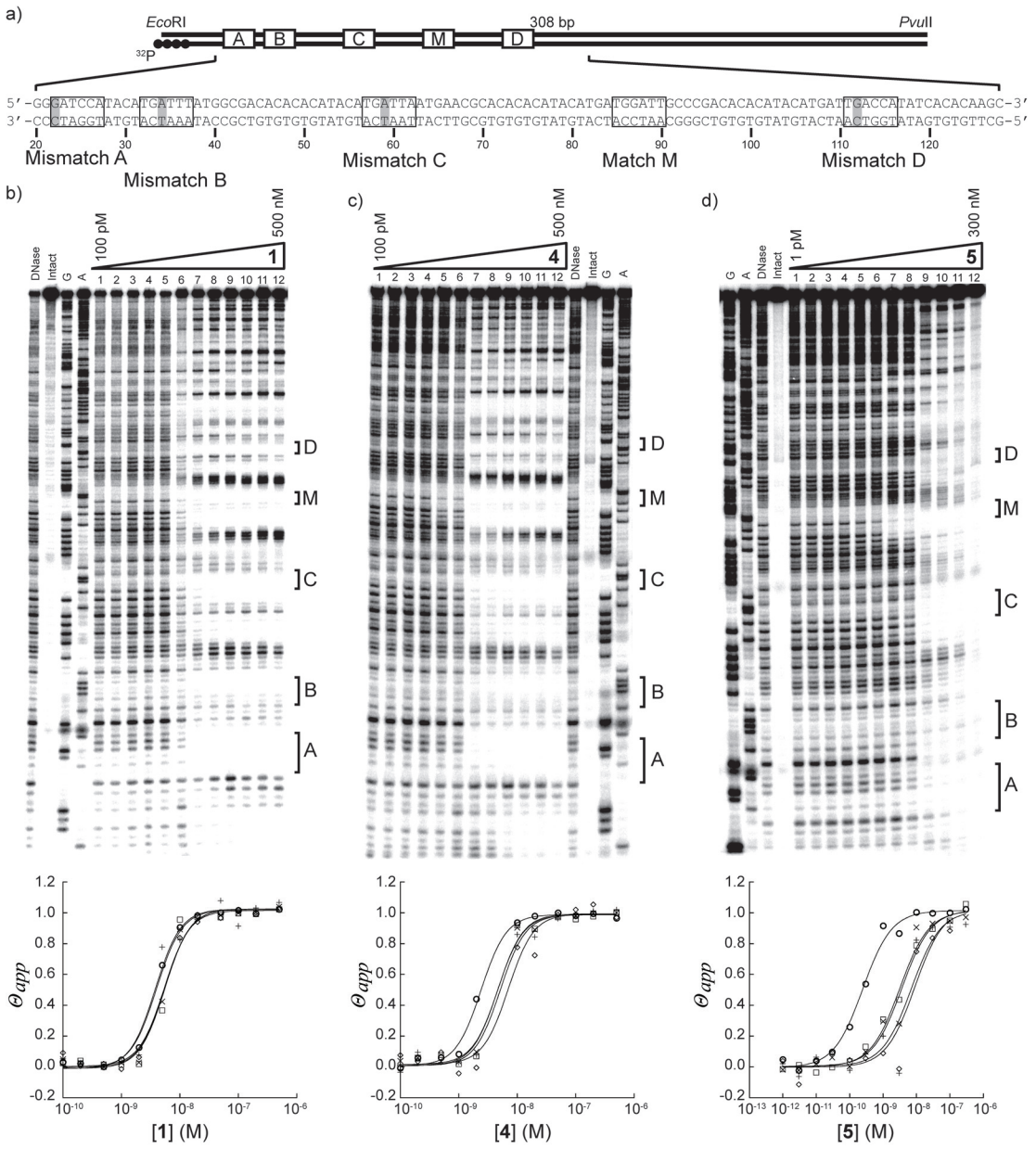


Figure A.2 Quantitative DNase footprinting titrations of compounds **1**, **4**, and **5** on a restriction fragment containing a match site M (5′-TGGATT-3′) and four single base pair mismatch sites A-D, (5′-TGGAT**C**-3′), (5′-TG**A**TTT-3′), (5′-TG**A**TTA-3′), and (5′-TGGT**C**A-3′), respectively. a) The DNA sequence of the restriction fragment with the location of the match and mismatch sites is shown. b) Titration of compound **1**. Lane 1-12: 0.1, 0.2, 0.5, 1, 2, 5, 10, 20, 50, 100, 200, 500 nM. c) Titration of compound **4** labeled as in panel b. d) Titration of compound **5**. Lane 1-12: 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100, 300 nM. DNase control, intact DNA, and G- and A-sequencing lanes are indicated for each gel. Isotherms are shown below each gel for match site M (○) and mismatch sites A (□), B (◊), C (×), and D (+).

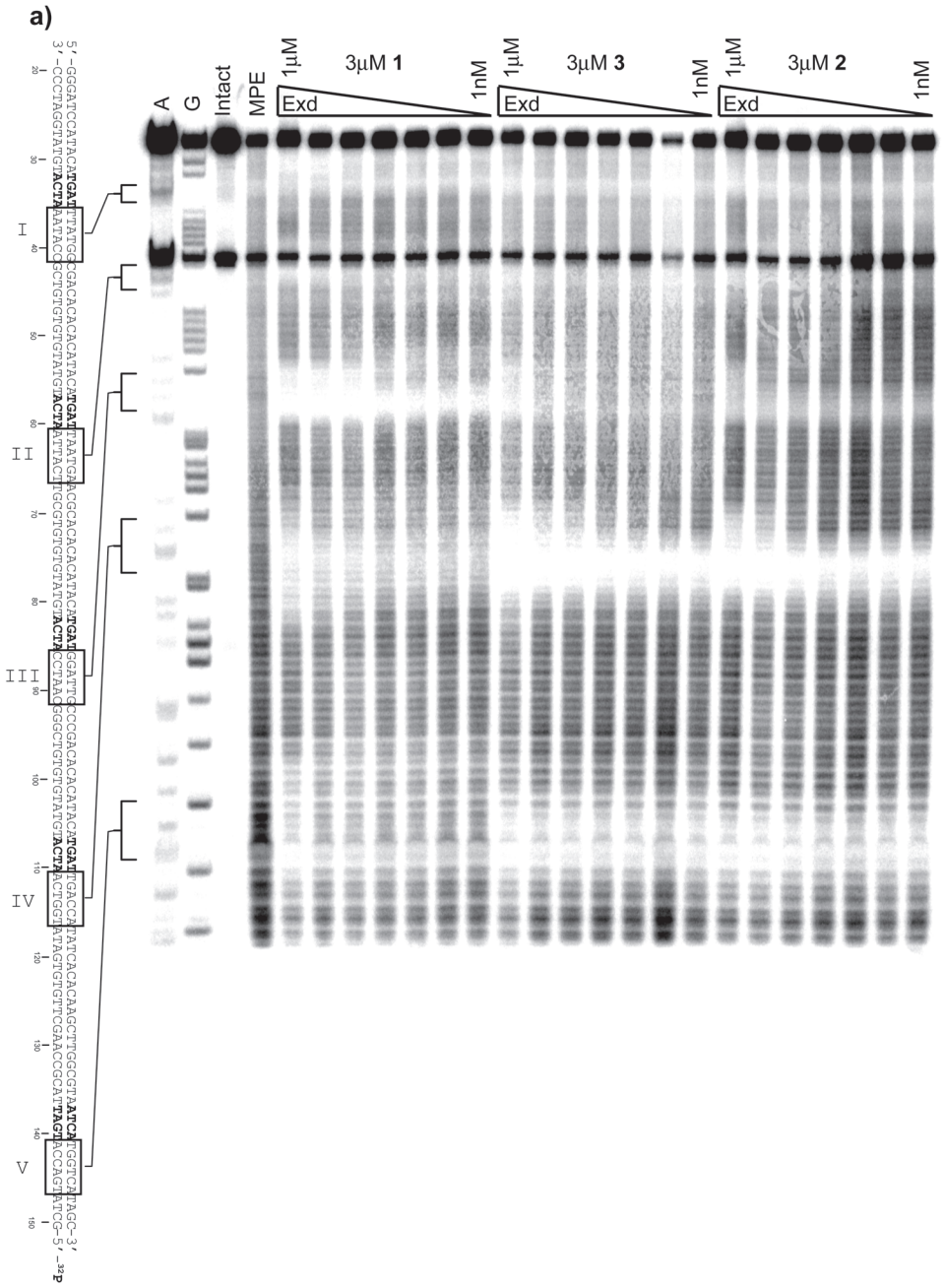


Figure A.3 Site-selective recruitment of Exd to the natural labial binding site. MPE footprinting experiments show an increase in binding site size exclusively at the labial cognate site III upon titration of Exd in the presence of conjugate **1**. In contrast, conjugates **2** and **3** which target the artificial site IV show an expansion in the binding site size exclusively at the artificial site IV.

labial match site III. Compounds **2** and **3**, which target an artificial HOX binding site, lead to increased protection exclusively at their match site IV. The control compounds **4** and **5** did not lead to increased protection at any of these sites when Exd was titrated (data not shown).

A.2 Design of a Cell-Permeable Labial Mimic

During the process of designing of a cell-permeable labial mimic, the putatively negative C-terminal β-alanine cell uptake determinant was re-evaluated. Specifically, polyamide **6** which contains a C-terminal β-alanine was synthesized by cleaving the corresponding polyamide from resin using 1,3-diaminopropane followed by coupling to fluorescein through an amide bond (FAM). Polyamide-FAM conjugate **6** was shown to yield nuclear localization in HeLa, MCF-7, PC3 cells (Figure A.4). The solubility of **6** was observed to be poor in unbuffered water likely due to the lack of a positive charge in the C-terminus. At pH 7 in phosphate buffer saline the compound was soluble, presumably due to the complete deprotonation of the carboxylic acid of the fluorescein leading to a net negative charge. One potentially useful feature of compound **6** is that it can be synthesized using pre-loaded BOC-β-Ala-PAM[9] instead of oxime resin.[10]

The parent polyamide **7** ImImPyPy-γ-PyPyPyPy-β-C_3-BOC was cleaved from PAM using mono Boc-protected 1,3-diaminopropane (Figure A.5). Following procedures previously described in Chapter 4, the desired conjugate **8** was made by coupling a peptide to the polyamide, deprotecting with TFA, and attaching FAM. Conjugate **8** showed nuclear localization in HeLa, MCF-7, and PC3 cells (Figure A.6). These compounds might be useful for performing cell culture or fruit fly experiments.

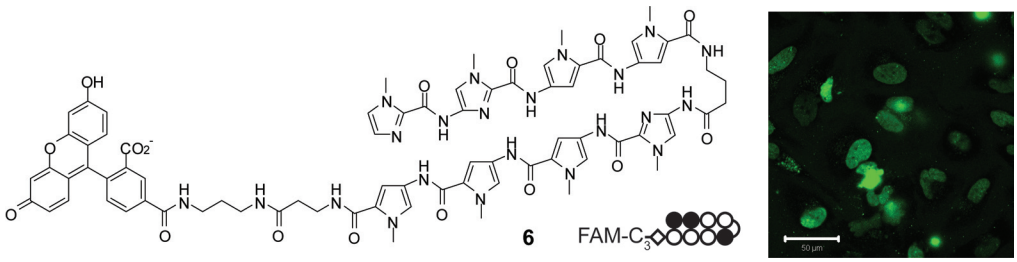


Figure A.4 Uptake results in HeLa cells for polyamide-FAM conjugate with a β-1,3-diaminopropane tail which shows primarily nuclear staining. (Scale bar = 50 µm)

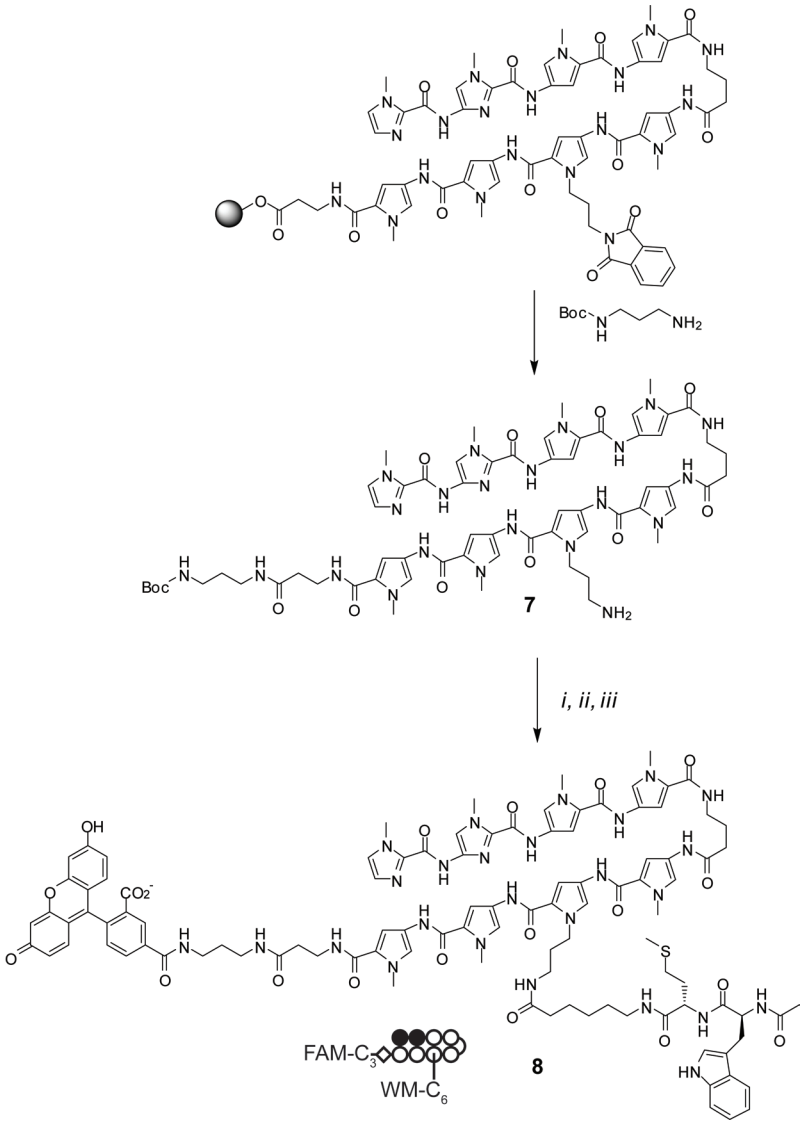
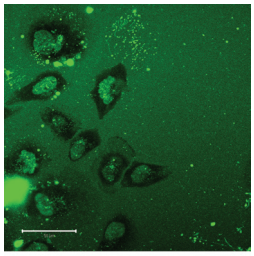


Figure A.5 Synthesis of polyamide-peptide-fluorescein conjugates targeted to the natural labial binding site: *i*) peptide, HBTU, 0.1 M DIEA in DMF, *ii*) 50% TFA/DCM, and *iii*) 5-carboxyfluorescein succinimidyl ester.



FAM-C_3

WM-C_6

8

Figure A.6 Uptake results in HeLa cells for polyamide-peptide-FAM conjugate **8** shows significant nuclear localization.

Table A.1 DNA equilibrium association constants K_a's [$\times 10^8$ M^{-1}]

Compound	5'-TGGATT-3'	5'-TGGAT**C**-3'	5'-TG**A**TTT-3'	5'-TG**A**TTA-3'	5'-TGGT**C**A-3'
1[a]	2.4 ± 0.9	1.7 ± 0.3	1.5 ± 0.3	1.7 ± 0.3	3 ± 2
4[a]	3 ± 1	1.5 ± 0.5	1.0 ± 0.4	1.5 ± 0.4	1.5 ± 0.4
5[b]	70 ± 50	6 ± 4	0.6 ± 0.3	6 ± 4	4 ± 2

[a] Data fit best using a Hill coefficient of 2.

[b] Data fit best using a Hill coefficient of 1.

A.3 Experimental Details

Synthesis

ImImPyPy-γ-PyPy$^{(Phe\text{-}Tyr\text{-}Pro\text{-}Trp\text{-}Met\text{-}Lys\text{-}Gly\text{-}propylamine)}$PyPy-β-Dp (1). This polyamide-peptide conjugate was synthesized essentially as previously described.[1] Briefly, the peptide Phe-Tyr(OtBu)-Pro-Trp(Boc)-Met-Lys(Boc)-Gly-OH (9.6 mg, 7.8 μmol) was activated with HBTU (7.5 μmol) and added to the parent polyamide **5** (2.76 μmol) in 1 mL DMF and excess DIEA. The solvent was evaporated under reduced pressure and deprotection was accomplished with 80% TFA, 5% EDT, 5% TES, and 10% DCM for 20 min at RT. The solution was added to 10 mL of ice-cold ether from which a precipitate formed and collected by centrifugation. The pellet was washed 2x with ether and dried briefly with $N_{2(g)}$. The remaining Trp carboxylate was eliminated by dissolving the pellet in 0.2 M acetic acid and incubating overnight. The final compound was purified by preparative HPLC and lyophilized to give 330 nmol of a white powder (~12% yield). (MALDI-TOF MS) [M +

H]$^+$ calcd for $C_{109}H_{138}N_{31}O_{19}S^+$ 2217.0, observed 2216.8.

ImImPy$^{(Phe\text{-}Tyr\text{-}Pro\text{-}Trp\text{-}Met\text{-}Lys\text{-}Gly\text{-}propylamine)}$Py-γ-ImPyPyPy-β-Dp (2). This polyamide-peptide conjugate was synthesized and characterized as previously described.[1]

ImImPy$^{(Trp\text{-}Met\text{-}\varepsilon Ahx\text{-}propylamine)}$Py-γ-ImPyPyPy-β-Dp (3). The polyamide-peptide conjugate was synthesized and characterized as described previously in Chapter 2.[4]

ImImPyPy-γ-PyPy$^{(Phe\text{-}Tyr\text{-}Pro\text{-}Ala\text{-}Ala\text{-}Lys\text{-}Gly\text{-}propylamine)}$PyPy-β-Dp (4). This polyamide-peptide conjugate was synthesized essentially as previously described.[1] Briefly, the peptide Phe-Tyr(OtBu)-Pro-Ala-Ala-Lys(Boc)-Gly-OH (8.0 mg, 8.4 μmol) was activated with HBTU (7.5 μmol) and added to the parent polyamide **5** (2.76 μmol) in 1 mL DMF and excess DIEA. The compound was deprotected as described for compound **1** purified by preparative HPLC and lyophilized to give 440 nmol of a white powder (~16% yield). (MALDI-TOF MS) [M + H]$^+$ calcd for $C_{99}H_{129}N_{30}O_{19}^+$ 2042.0, observed 2042.0.

ImImPyPy-γ-PyPy$^{(propylamine)}$PyPy-β-Dp (5). This polyamide was synthesized on PAM resin using standard protocols cleaving with 2 mL of Dp.[9] Total yield of 48.0 μmol of an off-white powder. (MALDI-TOF MS) [M + H]$^+$ calcd for $C_{60}H_{77}N_{22}O_{10}^+$ 1265.6, observed 1265.6.

ImImPyPy-γ-ImPyPyPy-β-C_3-FAM (6). This conjugate was synthesized by coupling 5-carboxyfluorescein succinimidyl ester (1 μmol) to the polyamide ImImPyPy-γ-ImPyPyPy-β-C_3-NH_2 (910 nmol) in 200 μL of 0.1 M DIEA in DMF for 20-30 min at RT. The final product was purified by preparative HPLC and lyophilized to give 270 nmol of a yellow-orange solid (~30% yield). (MALDI-TOF MS) [M + H]$^+$ calcd for $C_{76}H_{77}N_{22}O_{16}^+$ 1553.6, observed 1553.6.

ImImPyPy-γ-PyPy$^{(propylamine)}$PyPy-β-C_3-BOC (7). This polyamide was synthesized on PAM resin using standard protocols[9] cleaving with ~1 mL of *N*-Boc-1,3-diaminopropane. Total yield of 1.9 μmol of an off-white powder (~2% yield). (MALDI-TOF MS) [M + H]$^+$ calcd for $C_{63}H_{81}N_{22}O_{12}^+$ 1337.6, observed 1337.7.

ImImPyPy-γ-PyPy$^{(Trp\text{-}Met\text{-}\varepsilon Ahx\text{-}propylamine)}$PyPy-β-$C_3$-FAM (8). The conjugate was synthesized

in a similar manner to conjugate **1** using the peptide Ac-Trp(Boc)-Met-εAhx-OH (2.4 mg) activated with PyBOP (7.5 mg) coupled to polyamide **7** (1.5 μmol). The resulting intermediate was then coupled to 5-carboxyfluorescein succinimidyl ester (3.0 μmol). The final product was purified by preparative HPLC and lyophilized to yield 536 nmol of a yellow-orange solid (~36% yield). (ESI MS) $[M + 2H]^{2+}$ calcd for $C_{103}H_{116}N_{26}O_{20}^{2+}$ 1034.4, observed 1034.9.

***N*-Boc-1,3-diaminopropane.** Cool 14.8 g (16.7 mL, 0.2 mol) of 1,3-diaminopropane on an ice bath and add a solution of di-*tert*-butyl dicarbonate in a drop-wise manner (4.7 g, 0.021 mol in 50 mL DCM). Warm to room temperature over 1.5 hours and extract with 200 mL of 1/4 sat. $NaHCO_3$. Back extract aqueous layer 4x with 50 mL DCM. Wash combined organic layers with 50 mL saturated $NaHCO_3$ and 50 mL brine. Dry organic layer with anhydrous sodium sulfate, filter, and evaporate under reduced pressure with a small amount of toulene to azeotrope with residual water. ^{1}H NMR (DMSO, 300 MHz) δ 7.8 (br t, 1 H), δ 2.94 (q, 2 H), δ 2.49 (t, 2 H), δ 1.40 (p, 2 H), δ 1.36 (s, 9 H), 1.10 (s, 2 H); HRMS (FAB) exact mass calcd for $C_8H_{19}N_2O_2^+$ requires *m/z* 175.1447, found 175.1452.

Quantitative DNase I footprinting. Reactions were carried out in a volume of 400 μL in aqueous TKMC buffer according to published protocols[11] using a 3' ^{32}P-labeled 250 base pair *Eco*RI/*Pvu*II restriction fragment of the plasmid pHDA1. Developed gels were imaged using storage phosphor autoradiography using a Molecular Dynamics 400S Phosphorimager. Equilibrium association constants were determined as previously described.[11]

MPE footprinting with Exd. MPE footprinting was carried out in 40 μL reaction volumes according to published procedures.[11] Reactions were carried out as described in Chapter 2 in a total volume of 40 μL in aqueous TKMC buffer containing 10% glycerol using the 3' ^{32}P-labeled 308 base pair *Eco*RI/*Pvu*II restriction fragment of the plasmid pHDA1.

Cell culture and confocal microscopy experiments. All cell culture and confocal microscopy was performed essentially as described previously.[12] See Chapter 4 for a detailed explanation.

A.4 References

(1) Arndt, H.-D., Hauschild, K. E., Sullivan, D. P., Lake, K., Dervan, P. B., Ansari, A. Z. *J. Am. Chem. Soc.* **2003**, *125*, 13322-13323.

(2) Hauschild, K. E., Metzler, R. E., Arndt, H.-D., Moretti, R., Raffaelle, R., Dervan, P. B., Ansari, A. Z. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5008-5013.

(3) Warren, C. L., Kratochvil, N. C. S., Hauschild, K. E., Foister, S., Brezinski, M. L., Dervan, P. B., Phillips, G. N. Jr., Ansari, A. Z. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 867-872.

(4) Stafford, R. L., Arndt, H.-D., Brezinski, M. L., Ansari, A. Z., Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2660-2668.

(5) Stafford, R. S. and Dervan, P. B. *submitted to J. Am. Chem. Soc.* **2007**.

(6) Ryoo, H. D., Marty, T., Casares, F., Affolter, M., Mann, R. S. *Development* **1999**, *126*, 5137-5148.

(7) Dervan, P. B. and Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299.

(8) Neely, L., Trauger, J.W., Baird, E.E., Dervan, P.B., Gottesfeld, J.M. **1997** *J. Mol. Biol.* **274**, 439-445.

(9) Baird, E. E. and Dervan, P. B. *J. Am. Chem. Soc.* **1996,** *118*, 6141-6146.

(10) Belitsky, J. M., Nguyen, D. H., Wurtz, N. R. Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 2767-2774.

(11) Trauger, J. W. and Dervan, P. B. *Methods Enzymol.* **2001**, *340*, 450-466.

(12) Edelson, B. S., Best, T. P., Olenyuk, B., Nickols, N. G., Doss, R. M., Foister, S., Heckel, A., Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802-2818.

Appendix B

Toward a Small-Molecule YPWM Mimic

*Compounds **20**-**26** were synthesized by John Phillips (Dervan Group; Caltech).*

B.1 Background

Natural HOX proteins cooperatively bind to DNA with cofactor TALE proteins (Exd or Pbx) via a YPWM peptide motif.[1-3] The development of small molecules that competitively inhibit the YPWM-TALE protein-protein interaction could potentially be used as therapeutics for the treatment of pre-B cell ALL involving the Pbx-E2A oncoprotein.[4-9] A small molecule was discovered from a library of 1,4-disubstituted naphthalenes that binds to Pbx1 (IC_{50} = 65 μM) and blocks Pbx/HOX/DNA complex formation, but its effect on Pbx-E2A-mediated gene expression is unknown.[10] Such small molecules may also serve as binding domains for protein-DNA dimerizers which mimic HOX proteins.[11-15] The protein-binding domains for such protein-DNA dimerizers have been peptides which may be susceptible to endogenous proteases in cell-culture and fruit fly experiments. This appendix describes three non-peptidic mimics of the YPWM motif in the context of Exd protein-DNA dimerizers including the aforementioned 1,4-disubstituted naphthalene in search of a metabolically stable Exd protein-binding domain.

B.2 YPWM β-turn Mimetic

The YPWM motif has been highly conserved throughout evolution and is often referred to as a penta- or hexapeptide motif with an overall consensus sequence of ϕYPWM(K/R) where ϕ is a hydrophobic residue.[16] Residues flanking the hexapeptide have also been shown to contribute to YPWM-protein interactions.[17] The small buried YPWM-Exd interface (~570 Å^2)[1] when compared to mean protein-protein interfaces (~2000 Å^2)[18-19] suggested that this motif was already minimal. Thus, the initial attempt to develop a small-molecule YPWM mimetic aimed at maintaining the entire peptide structure. It was noted that the Exd/Ubx/DNA[1] and Pbx1/HoxB1/DNA[2] structures the YPWM motif adopts either a type I β-turn or the closely related type III β-turn (i.e., 3_{10} helix). Thus, a β-turn mimetic was designed based on a scaffold reported by the Burgess group.[20-23] This scaffold allowed incorporation of the absolutely conserved Trp at the *i+2* position and energy minimized

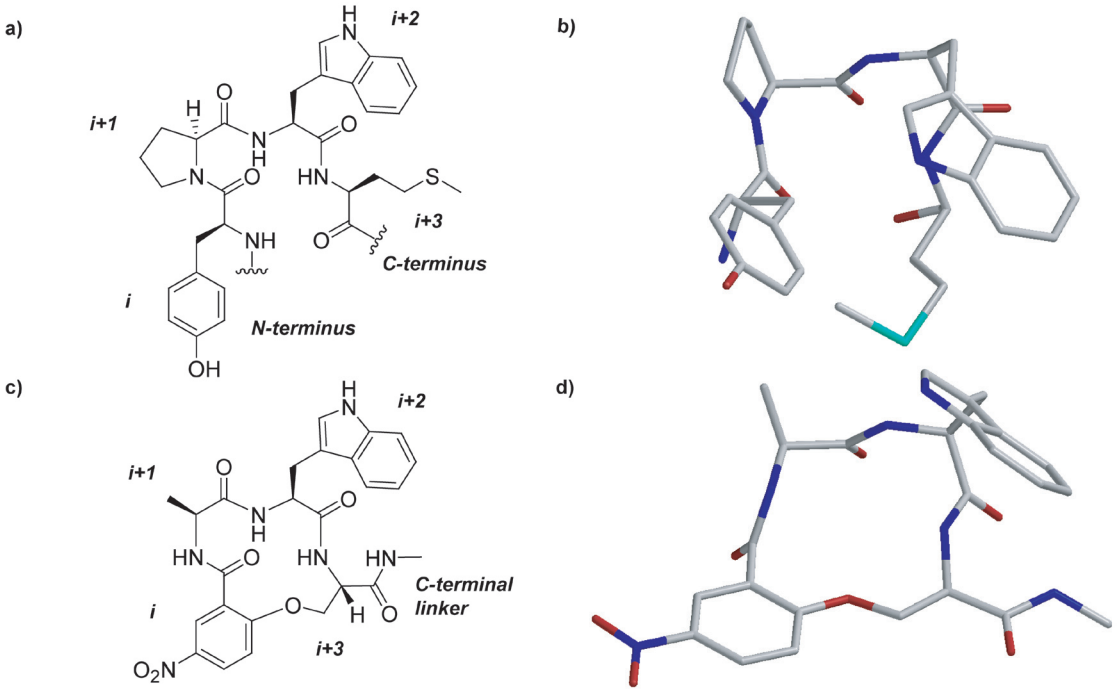


Figure B.1 Design of YPWM β-turn mimetic. a) Wild-type Ubx YPWM tetrapeptide depicted schematically in a β-turn conformation. b) Wild-type Ubx YPWM tetrapeptide from the Ubx/Exd/DNA crystal structure (PDB code 1B8I). c) Designed YPWM mimic using a β-turn scaffold from the Burgess group.[20-23] d) Energy minimized model (PM3) of the designed YPWM β-turn mimic.

models (PM3) suggested reasonable structural similarity to the native YPWM motif (Figure B.1).

The synthesis of YPWM β-turn mimetics was performed on a low-loading TentaGel PHB resin using procedures adapted from related macrocycles (Figure B.2).[20-21] Briefly, standard Fmoc peptide-synthesis was employed to yield resin-bound intermediate **1**. Fmoc deprotection, coupling to 2-fluoro-5-nitro-benzoic acid, and selective removal of the trityl with mild acid, yielded resin-bound intermediate **2**. Macrocycle formation was accomplished under basic conditions which promoted nucleophilic aromatic substitution of the serine hydroxyl with the fluorinated benzoic acid derivative. Treatment of the resin with strong acid and HPLC purification of the resulting crude mixture yielded macrocycle **3**. Low resin loading was found to prevent the formation of undesired dimerized by-products and was important for achieving good yields. Coupling of the macrocycle carboxylic acid

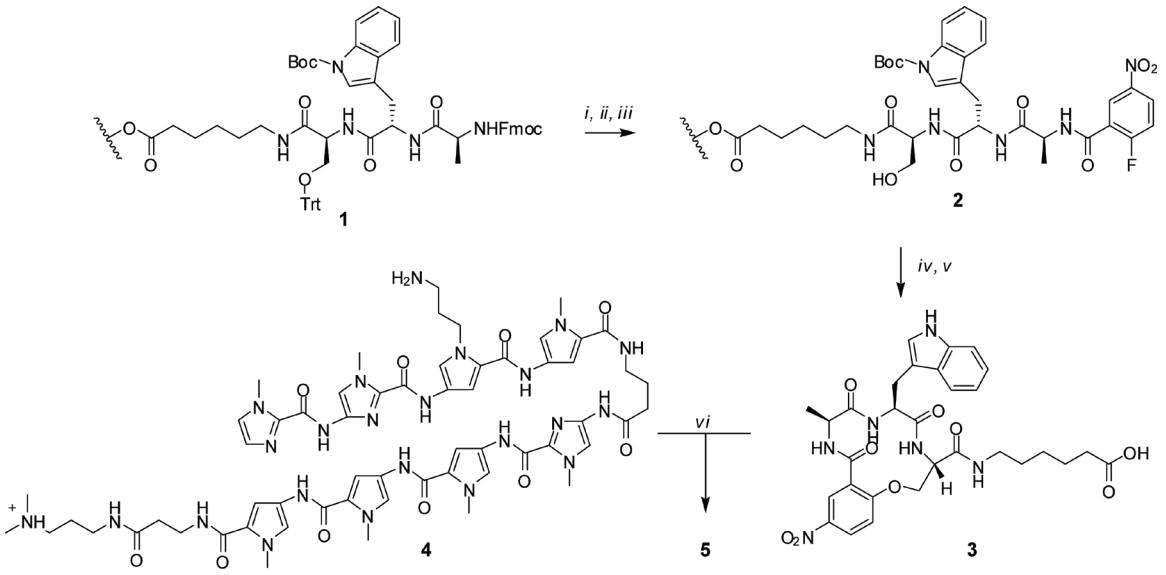


Figure B.2 Synthesis of polyamide-YPWM β-turn mimic conjugate: *i*) 20% piperidine in DMF, *ii*) 2-fluoro-5-nitrobenzoic acid, HATU, DIEA, DMF, *iii*) 1-3% TFA, 5% TIS, in DCM, *iv*) K_2CO_3 in DMF, *v*) 85% TFA, 5% TIS, 10% DCM, *vi*) HBTU, DIEA in DMF

3 to polyamide amine **4** via *in situ* HBTU activation yielded conjugate **5** (Figure B.3a). An uncyclized analogue **6** was synthesized using standard Fmoc chemistry and coupled to **4** to yield conjugate **7** to serve as a control (Figure B.3b).

An electrophoretic mobility shift assay (EMSA), or gel shift, with Exd at 4 °C was employed with conjugate **5** and **7** to assess their ability to facilitate Exd-DNA dimerization (Figure B.3c). Although the conjugate **5** enhanced Exd's DNA binding substantially, the uncyclized control compound exhibited a slightly higher Exd-DNA dimerization activity. This result suggested that the presence of a macrocycle was unnecessary for the Exd binding activity. This directly inspired the experiments that showed only the Trp and Met amino acids were necessary for Exd-DNA dimerization in the context of a polyamide-peptide conjugate.[14]

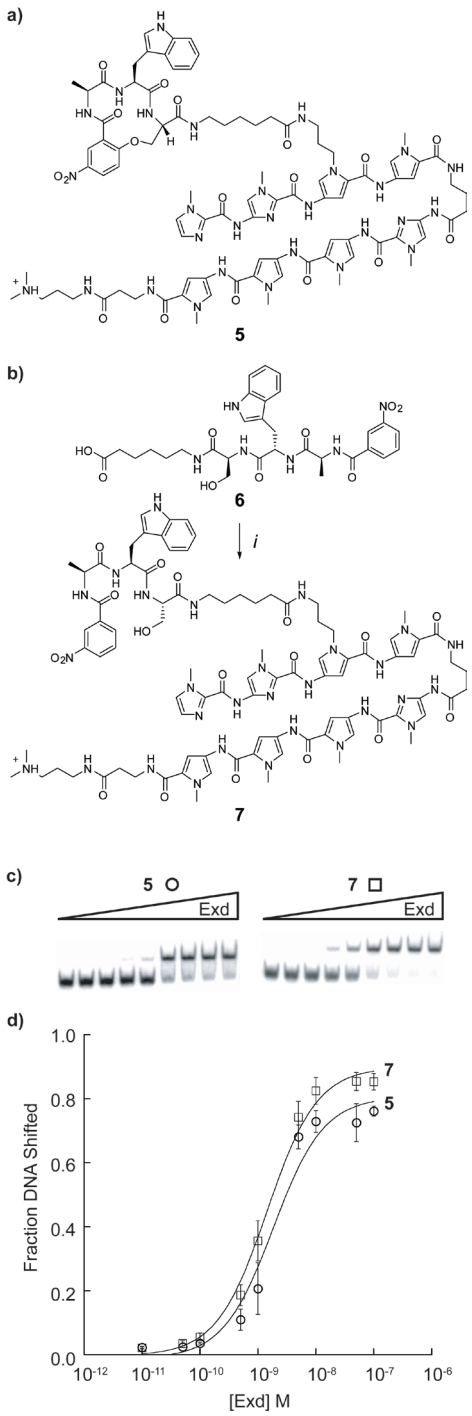


Figure B.3 Gel shifts experiments with polyamide-YPWM β-turn mimic conjugates. a) Structure of polyamide-YPWM β-turn mimic conjugate. b) Synthesis of uncyclized β-turn mimic control conjugate: *i*) **4**, HBTU, DIEA in DMF. c) Representative gel shift experiments for cyclized (**5**) and uncyclized (**7**) β-turn mimic conjugates. d) Isotherms for Exd binding in the presence of **5** and **7**. The uncyclized analogue (**7**) leads to a higher affinity complex than the cyclized compound (**5**). Gel shifts were performed at 50 nM compound titrating Exd from 10 pM to 100 nM.

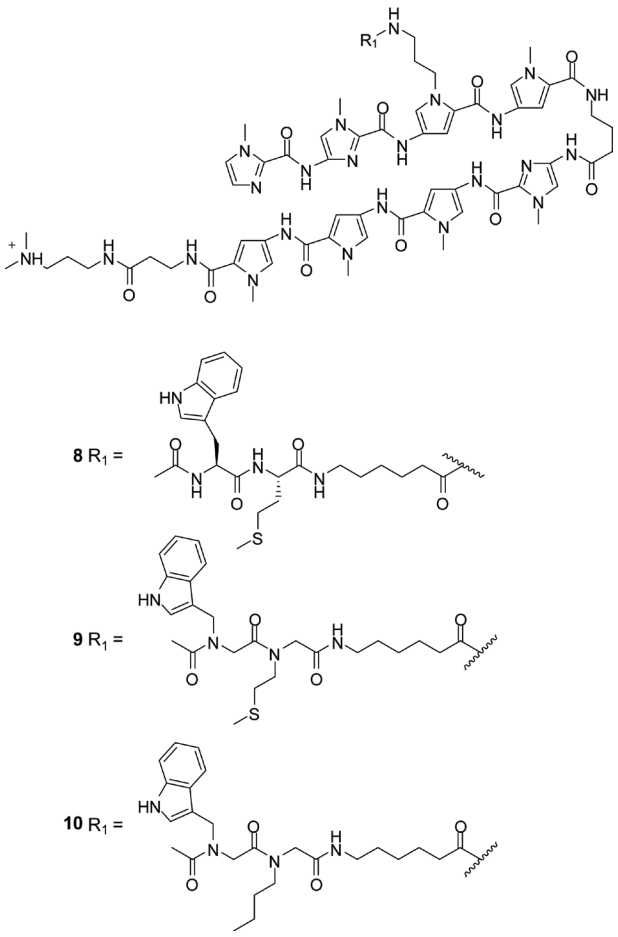


Figure B.4 Design of WM peptoid conjugates. The natural WM *peptide* conjugate **8** is shown for comparison to the WM *peptoid* conjugate **9**. A WM peptoid analogue **10** with a completely alkyl derivative of the methionine side chain is also shown.

B.3 WM Peptoid Analogue

One type of synthetic peptide analogue is a poly-*N*-substituted glycine, or peptoid.[24-26] Given that only Trp and Met residues were necessary to dimerize Exd,[14] efforts were undertaken to generate a WM peptoid (Figure B.4). Synthesis of the direct analogue of the WM peptide was attempted using the submonomer method[25-26] with 2-chlorotrityl resin, which was reported to enable the synthesis of peptoid carboxylates (Figure B.5).[27] Only the isolation of a compound consistent with **15** was observed, which was rationalized in terms of an intramolecular nucleophilic bromine displacement by the methionine thioether to give intermediate **13**, which probably resulted in the formation of the more stable diketopiperizine **14**. The alkyl analogue **17** was synthesized and coupled to the polyamide ImImPyPy-γ-ImPyPyPy-β-Dp to yield conjugate **10**. Gel shift experiments at 20 °C with **10** showed that it did not form as stable of an Exd-DNA complex as its peptide relative **8** (Figure B.6).

B.4 Naphthalene Small Molecule Inhibitor

The 1,4-disubstituted naphthalene **18** was previously reported to competitively

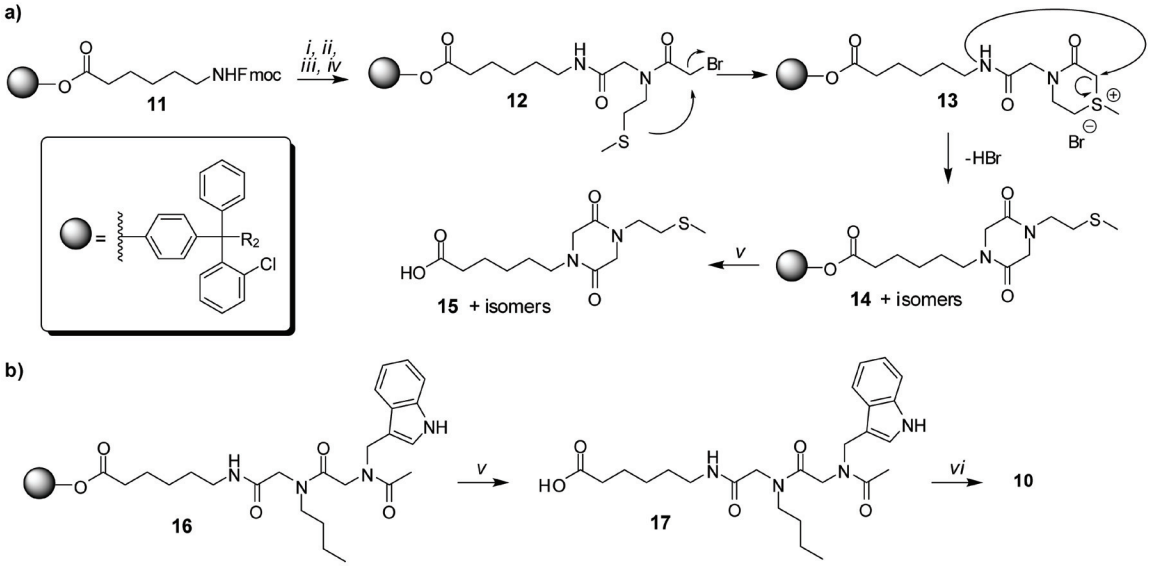


Figure B.5 Synthesis of WM peptoid analogues. a) The synthesis of WM peptoid conjugate was ultimately not successful due to the probable intramolecular bromine displacement by the methionine as shown: *i*) 20% piperidine in DMF, *ii*) α-bromoacetic acid, DIC, DIEA, DMF, *iii*) 2-(methylthio)ethylamine, DMSO, *iv*) α-bromoacetic acid, DIC, DIEA, DMF, *v*) 2:2:6 AcOH:TFE:DCM b) Standard submonomer peptoid synthesis using 2-chlorotrityl resin yielded the desired carboxylic acid **17**: *v*) 2:2:6 AcOH:TFE:DCM, *vi*) **4**, HBTU, DIEA in DMF.

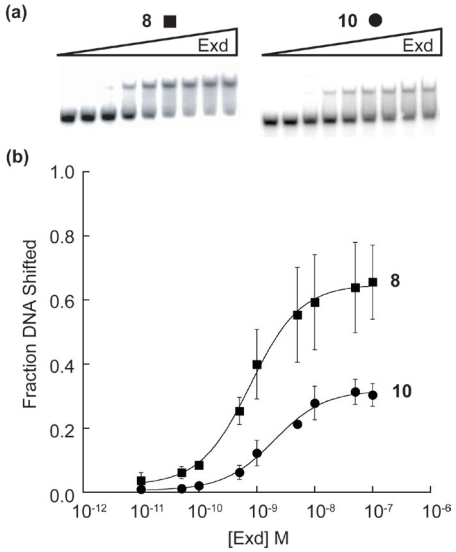
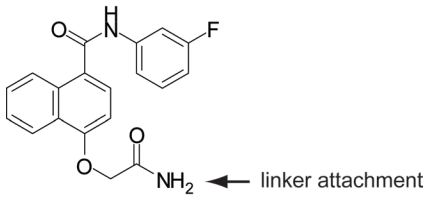


Figure B.6 Gel shifts experiments comparing polyamide-WM *peptoid* and *peptide* conjugates. a) The peptoid conjugate **10** yielded a lower affinity Exd-DNA complex than the peptide conjugate **8**. b) Isotherms for Exd bind in the presence of **8** and **10**. Gel shifts were performed with 50 nM compound titrating Exd from 10 pM to 100 nM.

inhibit Pbx/HOX/DNA complex formation presumably by competitively displacing the YPWM peptide (Figure B.7).[10] Although no rigorous structural data is available for the interaction of **18** with Pbx or Exd, published modeling studies[10] suggested the formamide moiety was a suitable attachment point for a polyamide conjugate since it was predicted to project away from the binding pocket. Accordingly, John Phillips synthesized a suitably modified version (**26**) to enable direct coupling to a polyamide. Starting from commercially available naphthyl aldehyde **19**, carboxylic acid **20** was obtained by oxidation. Coupling **20** to 3-fluoroaniline yielded amide **21** and the naphthylmethoxy was converted to aromatic hydroxyl **22** by nucleophilic displacement. The hydroxyl was condensed with α-bromomethylacetate to give methyl ester **23** which was saponified to yield carboxylic acid **24**. A four-carbon linker was then attached by amide bond formation to yield **25**, which was subsequently hydrolyzed to give the key carboxylic acid intermediate **26**. The final conjugate **27** was made via amide bond formation between a polyamide primary amine and the carboxylic acid of **26**.



18

Figure B.7 Structure of naphthalene compound **18** previously reported to competitively inhibit Hox/Pbx1/DNA complex formation by binding to the YPWM binding pocket of Pbx1.[10]

A gel shift experiment at 20 °C with compound **27** did not yield nearly as strong of an Exd-DNA complex as the polyamide-WM conjugate **8** (Figure B.9). The levels of complex formation and Exd affinity were only slightly better than were caused by the allosteric effects of the parent polyamide ImImPyPy-γ-ImPyPyPy-β-Dp at the same temperature.[14] It is possible that modifying the length or composition of the linker or its attachment point may increase the amount of complex formed.

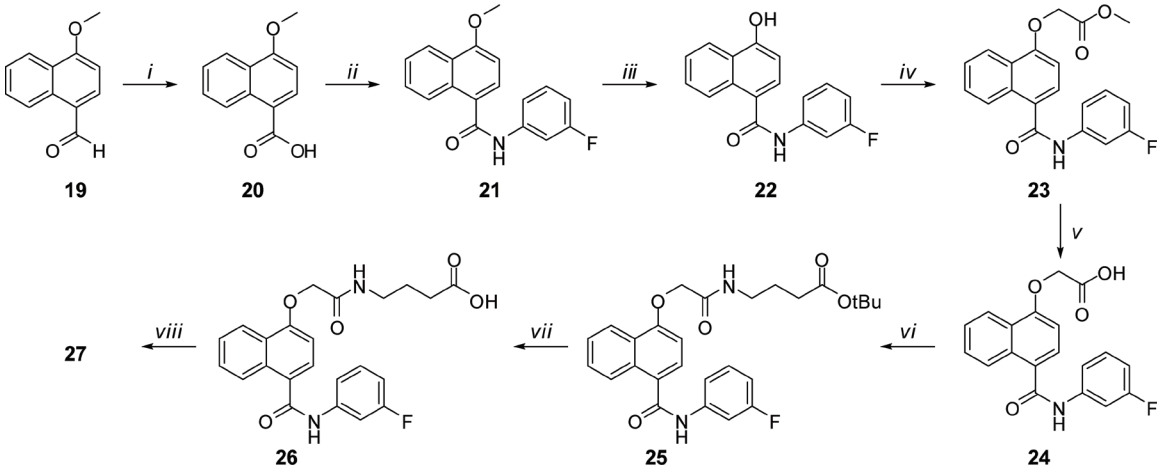


Figure B.8 Synthesis of an analogue of **18** suitable for conjugating to a polyamide: *i*) $NaClO_2$, NaH_2PO_4, H_2O, tBuOH, 2-methyl-2-butene, rt, 24 hours, *ii*) oxalyl chloride, 3-fluoroaniline, DIEA, THF, reflux, *iii*) ethanethiol, $AlCl_3$, DCM, 0 to 20 °C, *iv*) methyl bromoacetate, K_2CO_3, THF, reflux, 1 hour, *v*) LiOH, MeOH, H_2O, reflux, 2 hours, *vi*) γ-aminobutyric acid *tert*-butyl ester, PyBOP, DIEA, DMF, *vii*) 1:2 TFA:DCM, *viii*) **4**, HBTU, DIEA, DMF.

B.5 Experimental Details

Macrocycle (3). The synthesis of macrocycles was accomplished using a modified version of previously published methods.[20] Synthesis was performed using TentaGel S PHB resin (Fluka, loading ~0.24 meq/g). Standard Fmoc SPPS was employed initially to obtain peptide intermediate **1** which was deprotected with 20% piperidine in DMF for 20 min. Coupling of 2-fluoro-5-nitrobenzoic acid (2.5 equivalents) was performed by *in situ* activation of the free acid using HATU to give the activated ester which was then added to the deprotected amine of the previous residue. The trityl protecting group on the serine residue was selectively removed by approximately 5 treatments of 1 to 3%TFA and 5% TIS in DCM over roughly 2-3 hours until the yellow color dissipated to yield resin-bound intermediate **2**. Analytical cleavage of from the resin using the same method described for cleavage from Wang resin allowed characterization of this intermediate by mass spectrometry. After standard rinsing DMF (3x) and DCM (3x), cyclization was initiated with 5 equivalents of finely ground K_2CO_3 in DMF and allowed to proceed at room temperature for >48 hours.

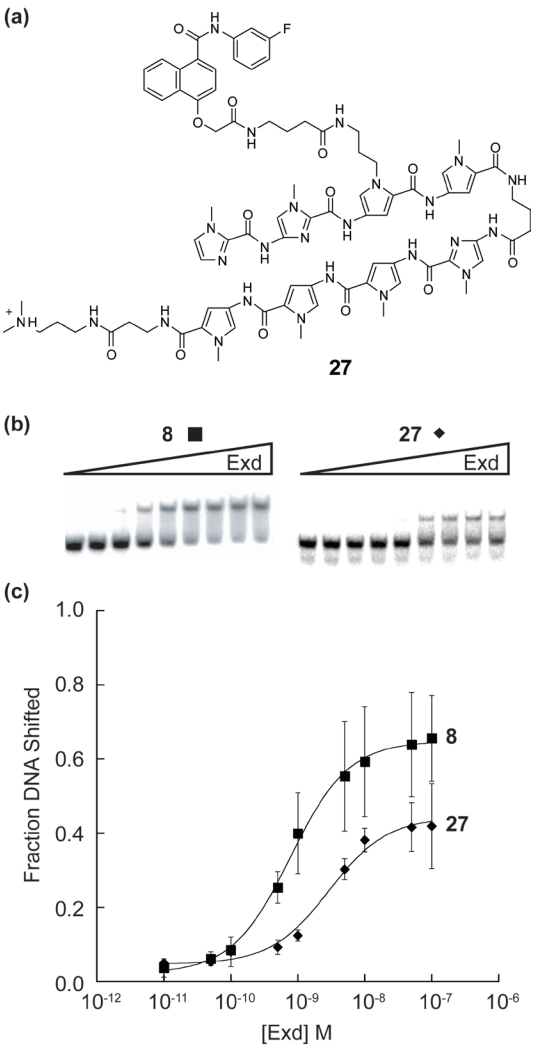


Figure B.9 Gel shift experiments comparing polyamide-WM conjugate **8** to polyamide-naphthalene conjugate **27**. a) Structure of polyamide-naphthalene conjugate. b) Representative gel shifts in the presence of **8** and **27**. c) Isotherms for Exd binding in the presence of **8** and **27**. Gel shifts were performed with 50 nM compound titrating Exd from 10 pM to 100 nM.

Cleavage from resin was performed using 85%TFA, 5%TIS, and 10%DCM (~4x, 2.5 mL) followed by removal of most of the solvent under reduced pressure. The resulting solid was dissolved in a small volume of DMF (~200 μL) and enough 20% CH_3CN/0.1% TFA in water was added to give a total volume of 10mL which was purified by preparative HPLC. Upon lyophilization of the appropriate fractions yielded 4.1 mg an off-white solid (3.7% yield). (ESI MS) average mass calculated for $C_{30}H_{35}N_6O_9^+$ $[M+H]^+$: 623.6, found 623.8 and calculated for $C_{30}H_{33}N_6O_9^-$ $[M-H]^-$: 621.6, found 621.8.

ImImPy$^{(propylamine)}$Py-γ-ImPyPyPy-β-Dp (4). The polyamide was synthesized according to standard protocols on BOC-β-Ala-PAM resin.[28] 265 mg of resin cleaved with 2 mL of neat *N,N*-dimethylaminopropylamine (Dp) at 37 °C for 48 hours, purified by preparative HPLC and lyophilized to give an off-white powder (10.9 μmol, ~17% yield). (MALDI-TOF MS) calculated for $C_{59}H_{76}N_{23}O_{10}^+$ 1266.6, observed 1266.7.

ImImPy$^{(macrocycle-propylamine)}$Py-γ-ImPyPyPy-β-Dp (5). Excess polyamide **4** (3.55 μmol, 2.2 eq) was added to macrocycle **3** (1 mg, 1.6 μmol, 1 eq) pre-activated (~5 min, RT) with

HBTU (1 eq) in DMF (~500 μL) and DIEA (8 equivalents). The resulting compound was purified by prepartive HPLC and lyophilzed to give a solid. (MALDI-TOF MS) $[M+H]^+$ base peak calculated for $C_{89}H_{108}N_{29}O_{18}^+$ 1871.8, found 1871.8.

Noncycle (6). This control compound was synthesized on Wang resin (0.305 g, 0.82 mmol/g, 0.250 mmol scale) using standard Fmoc peptide chemistry. The 3-nitrobenzoic acid moiety (Aldrich, 99%, 3.5 eq) was pre-activated with HATU (3.4 eq) in DMF with DIEA (4.5 eq) and coupled for 2 hours at RT. Cleavage from the resin was performed with 95% TFA/5% TIS, worked up as usual, and purified by preparative HPLC and lyophilized to give a solid. (ESI MS) average mass calculated for $C_{30}H_{37}N_6O_9^+$ $[M+H]^+$: 625.3, found 625.4 and calculated for $C_{30}H_{35}N_6O_9^-$ $[M-H]^-$: 623.3, found 623.4.

ImImPy$^{(noncycle\text{-}propylamine)}$Py-γ-ImPyPyPy-β-Dp (7). Polyamide **4** (0.97 μmol, 1 eq) was added to noncycle **6** (4.2 mg, 6.8 μmol, ~7 eq) pre-activated (~5 min, RT) with HBTU (3 eq) in DMF (~500 μL) and DIEA (20 eq). The resulting compound was purified by prepartive HPLC and lyophilzed to give a solid. (MALDI-TOF MS) $[M+H]^+$ calculated for $C_{89}H_{110}N_{29}O_{18}^+$ 1872.9, found 1873.0.

ImImPy$^{(Trp\text{-}Met\text{-}\varepsilon Ahx\text{-}propylamine)}$Py-γ-ImPyPyPy-β-Dp (8). This compound was synthesized and characterized as described in Chapter 2.

ImImPy$^{(N\text{-}Trp\text{-}N\text{-}But\text{-}\varepsilon Ahx\text{-}propylamine)}$Py-γ-ImPyPyPy-β-Dp (10). Polyamide **4** (1.5 μmol, 1 eq) was added to peptoid **17** pre-activated (~5 min, RT) with excess PyBOP and DIEA in DMF. The resulting compound was purified by prepartive HPLC and lyophilzed to give a solid (660 nmol, ~44 % yield). (MALDI-TOF MS) $[M+Na]^+$ calculated for $C_{84}H_{109}N_{27}O_{14}Na^+$ 1742.9, found 1742.9.

(15). This compound was synthesized using standard peptoid submonomer procedures[25,26] on 2-cholortrityl resin[27] employing 2-(methylthio)-ethylamine (Aldrich) and 3-(aminomethyl)indole (97%, Carbocore). (ESI-MS) Calculated for $C_{13}H_{23}N_2O_4S^+$ 303.1, observed 303.1.

N-Trp-N-Met-εAhx-OH (17). This compound was synthesized using standard peptoid

submonomer procedures[25,26] on 2-cholortrityl resin[27] employed butylamine (Aldrich) and 3-(aminomethyl)indole (97%, Carbocore). The peptoid was cleaved using 6:2:2 DCM:TFE: AcOH (1 mL, 3x), rinsed with DCM (1 mL), added to hexanes (5 mL, 2x), and rotavapped to dryness. The resulting film was dissolved in 500 μL DMF and diluted in 0.1% TFA in water and acetonitrile and purified by preparative HPLC. Fractions with consistent mass spectra by MALDI-TOF MS were pooled and lyophilized to give a white powder. (ESI-MS) $[M+Na]^+$ Calcd for $C_{25}H_{36}N_4O_5Na^+$ 495.3, observed 495.3. The MS/MS (35% collision energy) fragmentation pattern of the isolated *m/z* of 495.3 peak was consistent with the desired compound.

4-methoxy-1-naphthoic acid (20). This reaction was performed based on a previously reported procedure.[29] The starting 4-methoxy-1-naphthaldehyde **19** (5 g, 26.85 mmol) was dissolved in tBuOH (40 mL) and 2-methyl-2-butene (20 mL) to give a yellow solution. $NaClO_2$ (3.95 g in 40 mL ddH_2O with 4.80 g $NaH_2PO_4 \cdot H_2O$) was added in a drop-wise manner leading to the formation of a white precipitate. The reaction was allowed to proceed for 24 hours before filtering to collect the solid which was washed with hexanes (3x 50mL). The solid was dissolved in 0.1 N NaOH (60 mL) and extracted with hexanes and EtOAc (1x hexanes 30 mL, 1x EtOAc 50 mL, 1x hexanes 50 mL). The aqueous phase was acidified (pH 2) and the white precipitate was extracted with EtOAc (2 x 60 mL) and diethyl ether (1 x 60 mL). The combined organic layer was washed with brine (1x), dried with anhydrous magnesium sulfate, filtered and concentrated to yield 2.4 g of a white crystalline solid (45% yield). ^{1}H NMR ($CDCl_3$, 300 MHz) δ 4.102 (s, 3H), δ 6.87 (d, J = 8.7 Hz, 1H), δ 7.58 (t, 1H), δ 7.68 (t, 1H), δ 8.36 (d, J = 8.1 Hz, 1H), δ 8.44 (d, J = 9 Hz, 1H), δ 9.16 (d, J = 8.7 Hz, 1H); HRMS (FAB) exact mass calcd for $C_{12}H_{11}O_3$ requires 203.0708 *m/z*, found 203.0712.

***N*-(3-fluorophenyl)-4-methoxy-1-naphthamide (21).** The following procedure was adapted from a previously described reaction.[30] Into a dry, ventilated flask under argon the carboxylic acid **20** was added (2.104 g, 10.4 mmol) and dissolved in anhydrous DCM (20

mL) to which oxalyl chloride (20 mL) was added in a drop-wise manner. The vigorous reaction was allowed to proceed for 0.5 hours and the solvent and excess oxalyl chloride was removed under reduced pressure in a fume hood to give a yellow solid. The solid was dissolved in anhydrous THF (20 mL) with DIEA (7.25 mL, 41.6 mmol) and 3-fluoroaniline (1.20 mL, 12.48 mmol). The solution was refluxed for 30 min and the reaction was quenched with aqueous NH_4OH (20 mL, 30%) and the solvent was removed under reduced pressure. The dark brown residue was dissolved in EtOAc (100 mL) and extracted with HCl (1 x 100 mL, 5%), NaOH (1 x 100 mL, 0.1N), and brine (1 x 100 mL). The combined organic phase was removed under reduced pressure and the residue dissolved in minimal amount of EtOAc, precipitated from hexanes (200 mL), and collected by filtration as a white solid (1.8g, 59% yield). ^{1}H NMR (CD_3OD, 300 MHz) δ 4.07 (s, 3H), δ 6.91 (t, 1H), δ 6.98 (d, 1H), δ 7.3-7.6 (m, 4H), δ 7.73 (t, 2H), δ 8.30 (t, 2H); HRMS (FAB) exact mass calcd for $C_{18}H_{15}FNO_2$ requires 296.1087 *m/z*, found 296.1087.

***N*-(3-fluorophenyl)-4-hydroxy-1-naphthamide (22).** The following procedure was adapted from a previously described reaction.[31] In a dry flask, $AlCl_3$ (2.03 g, 15.2 mmol, 3 eq) was added to a solution ethanethiol (5 mL) and DCM (5 mL) cooled at 0 °C. After allowing the solution to reach RT, compound **21** (1.5 g, 5.08 mmol) was added under argon. The reaction was allowed to proceed overnight and the $AlCl_3$ was quenched by *very slowly* adding ddH_2O. The resulting precipitate was dissolved in EtOAc (40 mL) and the combined organic and aqueous phases were acidified to pH 2 with 1 N HCl and extracted with EtOAc (40 mL, 2x). The combined organic layers were washed with brine (1x 20 mL), dried with anhydrous magnesium sulfate, and concentrated under reduced pressure. The crude residue was purified over silica gel (10-50% EtOac/Hex gradient) and the product was isolated as an off-white amorphous solid (1.15 g, 80%). ^{1}H NMR (CD_3OD, 300 MHz) δ 6.85-6.88 (m, 2H), δ 7.33-7.60 (m, 4H), δ 7.65 (m, 2H), δ 8.26-8.29 (m, 2H); HRMS (FAB) exact mass calcd for $C_{17}H_{13}FNO_2$ requires 282.0930 *m/z*, found 282.0942.

methyl 2-(4-(3-fluorophenylcarbamoyl)naphthalen-1-yloxy)acetate (23). The following

procedure was based on a previously described reaction.[32] A solution of hydroxyl **22** (50 mg, 178 μmol) was made in THF (1 mL) with K_2CO_3 (37 mg, 196 μmol, 1.5 eq). Methylbromoacetate was added (18 μL, 588 μmol, 3.3 eq) and the resulting solution was refluxed for more than 3 hours. The solvent was removed under reduced pressure to give a brown residue and the remaining methylbromoacetate was removed under a high vacuum for 2 hours. The residue was dissolved in EtOAc (10 mL), extracted with water (10 mL, 2x) and brine (10 mL, 1x), dried over anhydrous magnesium sulfate, filtered, and concentrated under reduced pressure. The yellow crude material was purified over silica gel (10-30% hexanes/EtOAc gradient) to yield white solid with a yellow impurity (60 mg, 95% yield). ^{1}H NMR (CD_3OD, 300 MHz) δ 3.8 (s, 3H), δ 5.0 (s, 2H), δ 6.8-6.95 (m, 2H), δ 7.3-7.45 (m, 2H), δ 7.55-7.65 (m, 2H), δ 7.7 (d, 2H), δ 8.25 (d, 1H), δ 8.4 (d, 1H); HRMS (FAB) exact mass calcd for $C_{20}H_{17}FNO_4$ requires 354.1142 *m/z*, found 354.1150.

2-(4-(3-fluorophenylcarbamoyl)naphthalen-1-yloxy)acetic acid (24). The following procedure was based on a previously described reaction.[10] To a solution of ester **23** (50 mg, 142 μmol) in MeOH (2 mL) LiOH was added (213 μL of a 1M solution). The solution was refluxed for 2 hours and the solvent was removed under reduced pressure to yield a crude yellow residue which was dissolved in 10 mL of water which was extracted with hexanes (10 mL, 1x), EtOac (10 mL, 1x), and hexanes (10 mL, 1x). The aqueous layer was acidified to pH 2 with 1 N HCl to form a white precipitate which was extracted with EtOAc (20 mL, 3x). The combined organic layer were dried over anhydrous magnesium sulfate, filtered, and concentrated to dryness under reduced pressure to give an amorphous off-white solid (45 mg, 94% yield). ^{1}H NMR (d6-DMSO, 300 MHz) δ 4.97 (s, 2H), δ 6.89-6.99 (m, 2H), δ 7.33-7.80 (m, 6H), δ 8.23-8.32 (m, 2H), δ 10.63 (s, 1H); HRMS (FAB) exact mass calcd for $C_{19}H_{15}FNO_4$ requires 340.0985 *m/z*, found 340.0995.

4-(2-(4-(3-fluorophenylcarbamoyl)naphthalen-1-yloxy)acetamido)butanoic acid (26). The following procedure was based on a previously described reaction sequence.[33] To a solution of carboxylic acid **24** (45 mg, 133 μmol) in 1 mL anhydrous DMF and DIEA (23

μL, 1 eq) PyBOP was added (83 mg, 159 μmol, 1.2 eq) followed by γ-aminobutyric acid *tert*-butyl ester (29 mg, 146 μmol, 1.1 eq) and more DIEA (69 μL, 3 eq). Reaction was allowed to proceed for 18 hours and the solvent was removed under reduced pressure and the residue was dissolved in DCM (10 mL) and extracted with water, saturated $NaHCO_3$, 1 N HCl, and brine (10 mL, 1x each). The organic layer was evaporated to leave a yellow residue. The ester was hydrolyzed with 2:1 DCM:TFA (1.5 mL) and the solvent was removed under reduced pressure to yield a crude yellow-white residue which was dissolved in a minimum volume of DCM and purified over silica gel (9:1 EtOAc:hexanes + 1% AcOH) to give an off-white solid (41.2 mg, 73% yield). ^{1}H NMR (CD_3OD, 300 MHz) δ 1.98 (p, 2H), δ 2.33 (t, 2H), δ 3.37 (t, 2H), δ 4.80 (s, 2H), δ 6.85-6.99 (m, 2H), δ 7.32-7.45 (m, 2H), δ 7.56-7.74 (m, 4H), δ 8.29 (d, 1H), δ 8.51 (d, 1H); HRMS (FAB) exact mass calcd for $C_{23}H_{22}FN_2O_4$ requires 425.1513 *m/z*, found 425.1491; A_{max} ~230 nm (ε at 230 nm ≈ 2×10^4 $M^{-1}\cdot cm^{-1}$ and ε at 310 nm = 10^4 $M^{-1}\cdot cm^{-1}$).

ImImPy$^{(1,4\text{-napthalene-propylamine})}$Py-γ-ImPyPyPy-β-Dp (27). Polyamide **4** (2 μmol, 1 eq) was added to carboxylic acid **26** (10 mg, 23.6 μmol) pre-activated (~5 min, RT) with HBTU (8 μmol, 4 eq) in DMF (~280 μL) and DIEA (20 μmol). The resulting compound was purified by prepartive HPLC and lyophilzed to give a solid. (MALDI-TOF MS) $[M+H]^+$ calculated for $C_{82}H_{95}FN_{25}O_{14}^+$ 1672.8, found 1672.8. The conjugate's extinction coefficient was assumed to be the sum of the parent polyamide **4** and the small molecule **26** components (conjugate ε at 310 nm ≈ 79,500 $M^{-1}\cdot cm^{-1}$).

Electrophoretic Mobility Shift Assays (EMSAs). All EMSAs, or gel shifts, were performed at 4 or 20 °C as described in Chapter 2.[14]

B.6 References

(1) Passner, J. M., Ryoo, H. D., Shen, L. Y., Mann, R. S., Aggarwal, A. K. *Nature* **1999**, *397*, 714-719.

(2) Piper, D. E., Batchelor, A. H., Chang, C. P., Clearly, M. L., Wolberger, C. *Cell* **1999**, *96*, 587-597.

(3) LaRonde-LeBlanc, N. A. and Wolberger, C. *Genes Dev.* **2003**, *17*, 2060-2072.

(4) Pui, C., Relling, M. V., Downing, J. R. *N. Eng. J. Med.* **2004**, *350*, 1535-1548.

(5) Lu, Q., Wright, D. D., Kamps, M. P. *Mol. Cell. Biol.* **1994**, *14*, 3938-3948.

(6) Knoepfler, P. S. and Kamps, M. P. *Oncogene* **1997**, *14*, 2521-2531.

(7) McWhirter, J. R., Neuteboom, S. T. C., Wancewicz, E. V., Monia, B. P., Downing, J. R., Murre, C. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 11464-11469.

(8) Fu, X., McGrath, S., Pasillas, M., Nakazawa, S., Kamps, M. P. *Oncogene* **1999**, *18*, 4920-4929.

(9) Casagrande, G., Kronnie, G., Basso, G. *Haematologica/The Hematology Journal* **2006**, *91*, 765-771.

(10) Ji, T., Lee, M., Pruitt, S. C., Hangauer, D. G. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3875-3879.

(11) Arndt, H.-D., Hauschild, K. E., Sullivan, D. P., Lake, K., Dervan, P. B., Ansari, A. Z. *J. Am. Chem. Soc.* **2003**, *125*, 13322-13323.

(12) Hauschild, K. E., Metzler, R. E., Arndt, H.-D., Moretti, R., Raffaelle, R., Dervan, P. B., Ansari, A. Z. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 5008-5013.

(13) Warren, C. L., Kartochvil, N. C. S., Hauschild, K. E., Foister, S., Brezinski, M. L., Dervan, P. B., Phillips, G. N., Ansari, A. Z. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 867-872.

(14) Stafford, R. L., Arndt, H.-D., Brezinski, M. L., Ansari, A. Z., Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2660-2668.

(15) Stafford, R. L. and Dervan, P. B. *submitted to J. Am. Chem. Soc. 2007.*

(16) Duboule, D. editor in *Guidebook to the Homeobox Genes*; Oxford University Press: Oxford, 1994.

(17) Shanmugam, K., Featherstone, M. S., Saragovi, H. U. *J. Biol. Chem.* **1997**, *272*, 19081-19087.

(18) Chakrabartim, P. and Janin, J. *Proteins* **2002**, *47*, 334-343.

(19) Lo Conte, L., Chothia, C., Janin, J. *J. Mol. Biol.* **1999**, *285*, 2177-2198.

(20) Feng, Y. and Burgess, K. *Chem. Eur. J.* **1999**, *5*, 3261-3272.

(21) Wang, Z., Jin, S., Feng, Y., Burgess, K. *Chem. Eur. J.* **1999**, *5*, 3273-3278.

(22) Burgess, K. *Acc. Chem. Res.* **2001**, *34*, 826-835.

(23) Zhang, A. J., Khare, S., Gokulan, K., Linthicum, D. S., Burgess, K. *Bioorg. Med. Chem. Lett.* **2001**, *11*, 207-210.

(24) Simon, R. J., Kania, R. S., Zuckerman, R. N., Huebner, V. D., Jewll, D. A., Banville, S., Ng, S., Wang, L., Rosenberg, S., Marlowe, C. K., Spellmeyer, D. C., Tan, R., Frankel, A. D., Santi, D. V., Cohen, F. E., Bartlett, P. A. *Proc. Natl. Acad. Sci.* **1992**, *89*, 9367-9371.

(25) Zuckerman, R. N., Kerr, J. M., Kent, S. B. H., Moos, W. H. *J. Am. Chem. Soc.* **1992**, *114*, 10646-10647.

(26) Figliozii, G. M., Goldsmith, R., Ng, S. C., Banville, S. C., Zuckerman, R. N. *Methods Enzym.* **1996**, *267*, 437-447.

(27) Anne, C., Fournie-Zaluski, M-C., Roques, B. P., Cornille, F. *Tet. Lett.* **1998**, *39*, 8973-8974.

(28) Baird, E. E. and Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146.

(29) Hillis, L. R. and Ronald, R. C. *J. Org. Chem.* **1985**, *50*, 470-473.

(30) Spencer, J. R., McGee, D., Allen, D., Katz, B. A., Luong, C., Sendzik, M., Squires, N., Mackman, R. L. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2023-2026.

(31) Node, M., Nishide, K., Fuji, K., Fujita, E. *J. Org. Chem.* **1980**, *45*, 4275-4277.

(32) Kramer, B., Fröhlich, R., Bergander, K., Waldvogel, S. R. *Synthesis* **2003**, *1*, 91-96.

(33) Haberhauer, G., Somogyi, L., Rebek, J., *Tet. Lett.* **2000**, *41*, 5013-5016.

Appendix C

Exd Expression Protocols

Exd was expressed by the Protein Expression Center (Beckman Institute, Caltech).

Mass spectrometry analysis of the Exd protein fragments was performed by the Peptide/ Protein MicroAnalytical Lab (PPMAL, Beckman Institute, Caltech).

C.1 First Exd Preparation Experimental Details

A plasmid "Exd-C" was received from the Ansari lab (University of Wisconsin, Madison) which was originally obtained from the Aggarwal lab (Mount Sinai School of Medicine) (Figure C.1). Protein expression and purification was performed by Dr. Peter Snow of the Protein Expression Center (Beckman Institute, Caltech). The plasmid was transformed into *E. coli* BL21-DE3 cells and overexpressed by treating with 1 mM IPTG. Cells were lysed by sonication and crude purification was performed by taking the 50% (or 90%) ammonium sulfate cut. The crude preparation was dialyzed against buffer A (150 mM potassium glutamate, 50 mM HEPES, 1 mM DTT, pH 7.0, 5% glycerol, 100 mM NaCl). The dialyzed protein was purified by FPLC (Figure 2a; flow rate = 0.5 mL/min, eluant A = buffer A, eluant B = buffer A with 1 M NaCl, 0.5 mL fractions). An SDS-PAGE gel stained with Coomassie blue showed pure fractions (29-33) with the predicted mass (~10 kDa) for the desired Exd fragment (Figure C.2b). Band excision (fraction 32) followed by proteolytic digestion (Lys-C achromobacter) and MALDI-TOF MS analysis was consistent with the predicted spectra using Sherpa 3.3.1 for Exd (residues 238-314) (Figure C.3c-d). Mascot (Matrix Science) search results from LC-MS/MS data for the digested protein identified Dpbx (i.e., *Drosophila* Pbx or Exd) as the sole match protein (score = 163). Fractions 29-33 were combined and dialyzed against buffer A. A sizing column (G30) was used to remove aggregates. MALDI-TOF MS analysis of the combined fractions of the undigested protein from Caltech was consistent with the presence of Exd residues 238-320 (calculated $[M+H]^+$ 9495.7 *m/Z*, observed $[M+H]^+$ 9496.0) (Figure C.3a). An undigested "crude" Exd preparation from the Ansari lab yielded a similar mass spectrum (Figure C.3b). The stock Exd concentration was determined to be 432 ± 22 μM using the method of Stoscheck[1] and a calculated extinction coefficient[2] of 12,600 $M^{-1}cm^{-1}$ for Exd residues 238-320 (Figure C.4). Total yield was 4.1 mg.

```
    1          11         21         31         41         51
    |          |          |          |          |          |
  1 MEDPNRMLAH TGGMMAPQGY GLSGQDDGQN AGSENEVRKQ KDIGEILQQI MSISEQSLDE

 61 AQARKHTLNC HRMKPALFSV LCEIKEKTVL SIRNTQEEEP PDPQLMRLDN MLIAEGVAGP

121 EKGGGGAAAA SAAAASQGGS LSIDGADNAI EHSDYRAKLA QIRQIYHQEL EKYEQACNEF

181 TTHVMNLLRE QSRTRPITPK EIERMVQIIH KKFSSIQMQL KQSTCEAVMI LRSRFLDARR

241 KRRNFSKQAS EILNEYFYSH LSNPYPSEEA KEELARKCGI TVSQVSNWFG NKRIRYKKNI

301 GKAQEEANLY AAKKAAGASP YSMAGPPSGT TTPMMSPAPP QDSMGYPMGS GGYDQQQPYD

361 NSMGGYDPNL HQDLSP
```

Figure C.1 Amino acid sequence of Exd isoform C of *D. Melanogaster*. Bold-faced type indicates the region of Exd expressed (238-320) and italicized type indicates the region of Exd used in the Ubx/Exd/DNA crystal structure (PDB code 1B8I). Both regions contain Exd's DNA-binding homeodomain.

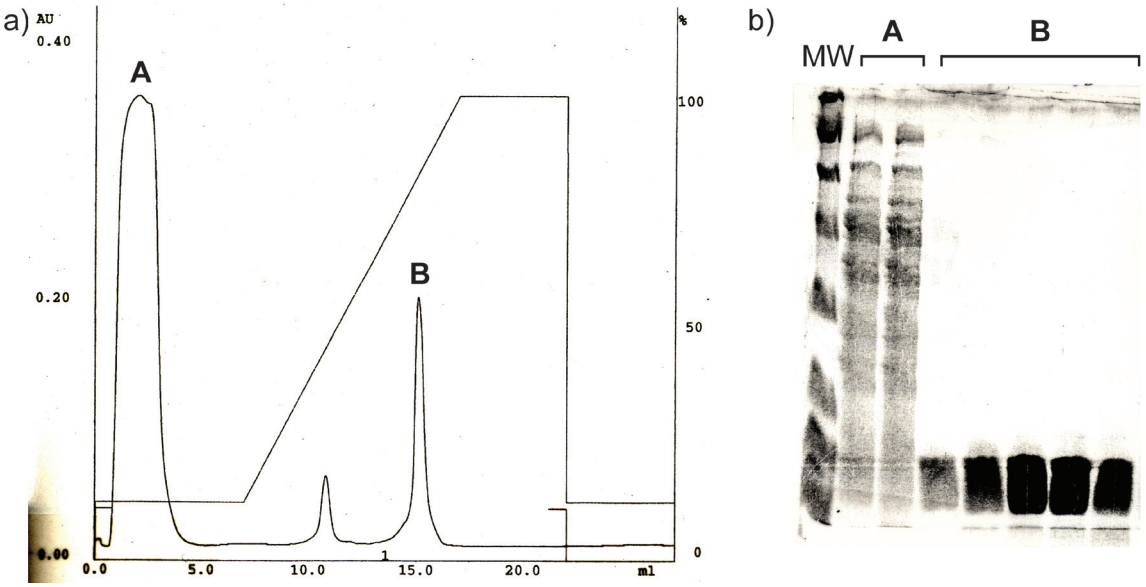


Figure C.2 Purification and analysis of the first Exd preparation. a) FPLC trace with major peaks A and B labeled. See text for details of purification. b) SDS-PAGE gel stained with Coomassie blue shows the content of peaks A (combined fractions 2-5) and B (fractions 29-33).

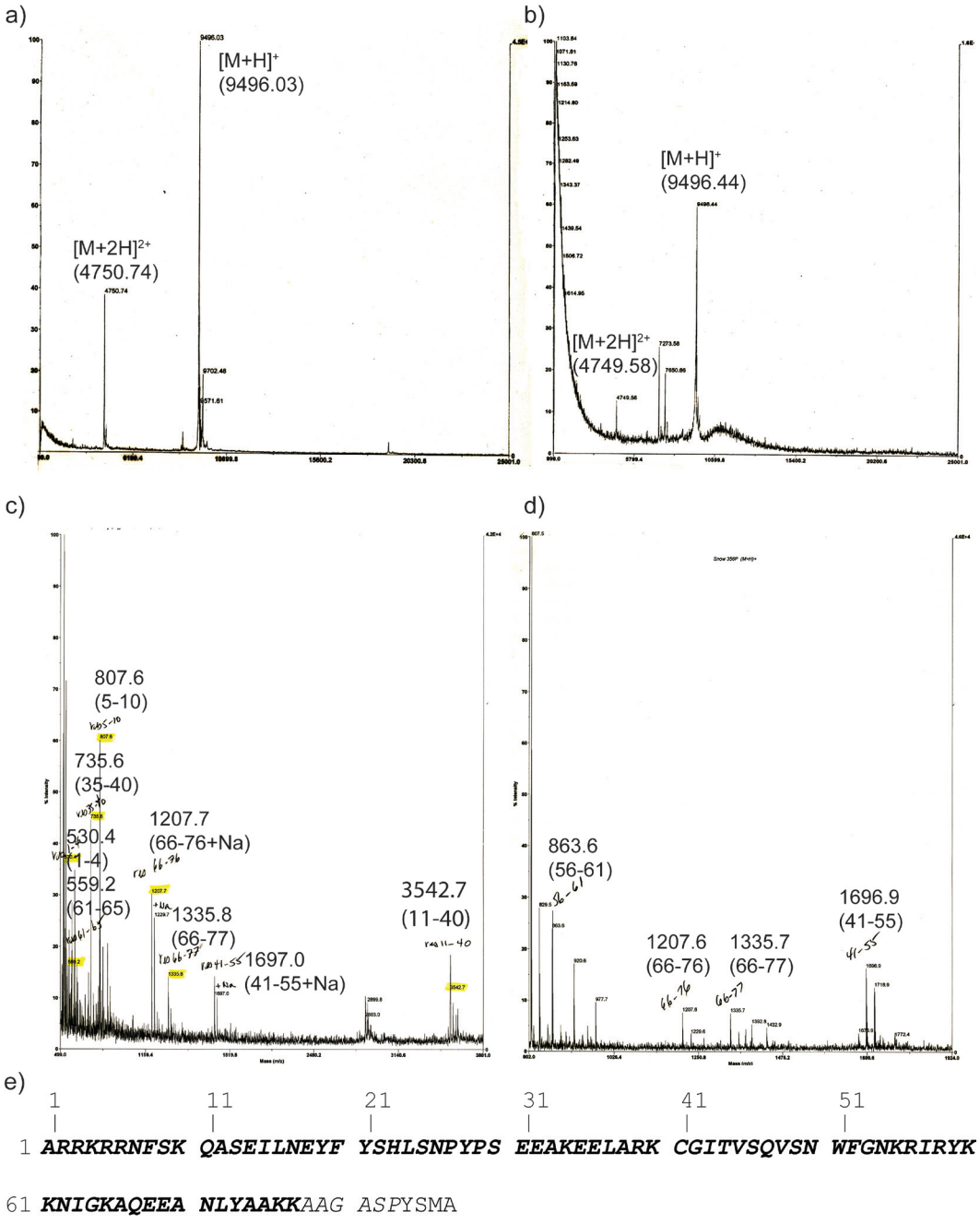


Figure C.3 MALDI-TOF MS analysis of different Exd preparations. a) Mass spectrum of Exd expressed at the Caltech Protein Expression Center by Peter Snow. b) Mass spectrum of "crude" Exd expressed by the Ansari lab and the University of Wisconsin, Madison. c) Mass spectrum #1 of proteolytic peptide fragments of Exd as determined at the Peptide/ Protein MicroAnalytical Lab (PPMAL) at Caltech by Gary Hathaway. d) Mass spectrum #2 of proteolytic peptide fragments of Exd determined by the PPMAL. e) Sequence of Exd reported in the original Arndt *et al* manuscript.[3] Bold-faced type indicates the region of Exd identified by proteolytical/MS analysis and italicized type indicates the region observed by undigested protein samples from Caltech and the University of Wisconsin, Madison.

C.2 Second Exd Preparation Experimental Details

The Exd-C plasmid was transformed in BL21-DE3 cells as above by Jost Vielmetter and Gilberto Salvo of the Caltech Protein Expression Center. Cells were grown to an OD of 0.6 at 37 °C and induced with 1 mM IPTG. Protein over-expression was allowed to proceed for 4 hours at 30 °C and cells were collected by centrifugation at 3000g for 30 min. For storage the cell pellet was flash frozen and stored at -80 °C. Cells were lysed in buffer B (1X PBS, 1 mM EDTA, 5 mM benzamidine, 1x PMSF, 1x complete protease inhibitor, pH 7.4) by sonication (Sonic Dismembrator 550 with Microtip at level 4 in pulsar/timer mode: 2 min duty cycle time, 0.5 s pulses interupted with 0.5 s pauses on ice). Cells were spun 2x at 13,000 rpm, 4 °C, 30 min with a Sorvall SS-34 rotor decanting the lysate into a fresh tube. A total of 5 mL lysate was obtained and stored at -80 °C. The diluted crude lysate (15 mL) was loaded onto a HiTrap Heparin HP column in injection buffer (25 mM TrisHCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 10% glycerol) and eluted with elution buffer (25 mM TrisHCl, pH 7.5, 1 M NaCl, 1 mM EDTA, 10% glycerol) with a flow rate of 1 mL/min collecting 1 mL fractions. Fractions 44-77 were pooled and concentrated down to 1 mL. No ammonium sulfate fractionation was employed. The pooled fractions were purified on a MonoS column by FPLC with a flow rate of 0.5 mL/min collecting 1 mL fractions (A buffer = injection buffer, B buffer = elution buffer) (Figure C.5). The major peaks were analyzed by an SDS-PAGE gel stained with Coomassie blue. Fractions 15 and 16 from the monoS column were pooled and 200 μL was loaded onto a sephadex-75 sizing column purified by an isochratic methods using the injection buffer (Figure C.6). Fractions 17-20 were pooled and concentrated down to ~200 μL. The concentration was estimated to be ~2 mg/mL (~30.2 μM) by Jost Vielmetter from the UV absorbance at 280 nm using a nanodrop spectrometer. Total yield was ~0.4 mg. An analytical sample showed a similar DNA binding affinity (within 2-fold) in the presence of a polyamide-FYPWMKG peptide conjugate at 20 °C as the preparation described in section C.1.

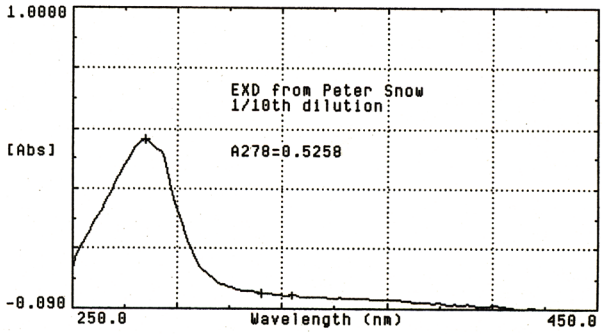


Figure C.4 UV-Vis spectrum of a 1/10th dilution of the Exd stock expressed by Peter Snow in 6.0 M guanadinium HCl, 0.01 M KH_2PO_4, 0.01 M K_2HPO_4 (pH 6.5) used for concentration determination.

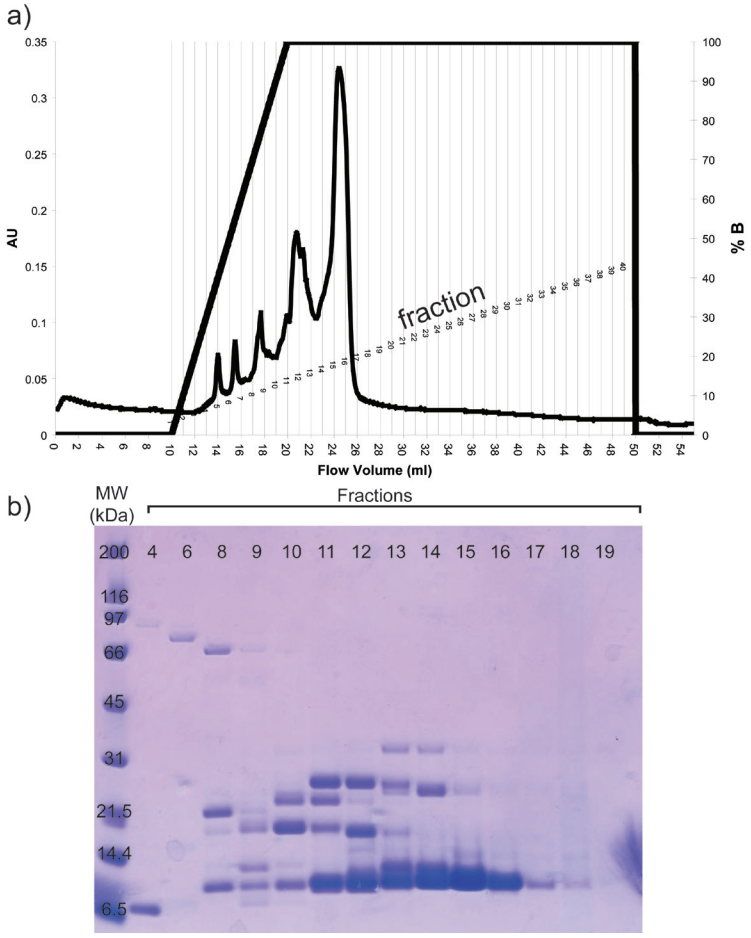


Figure C.5 Primary purification and analysis of the second Exd preparation. a) FPLC trace. See text for details of purification. b) SDS-PAGE gel stained with Coomassie blue shows the content of fractions 4, 6, and 8-19.

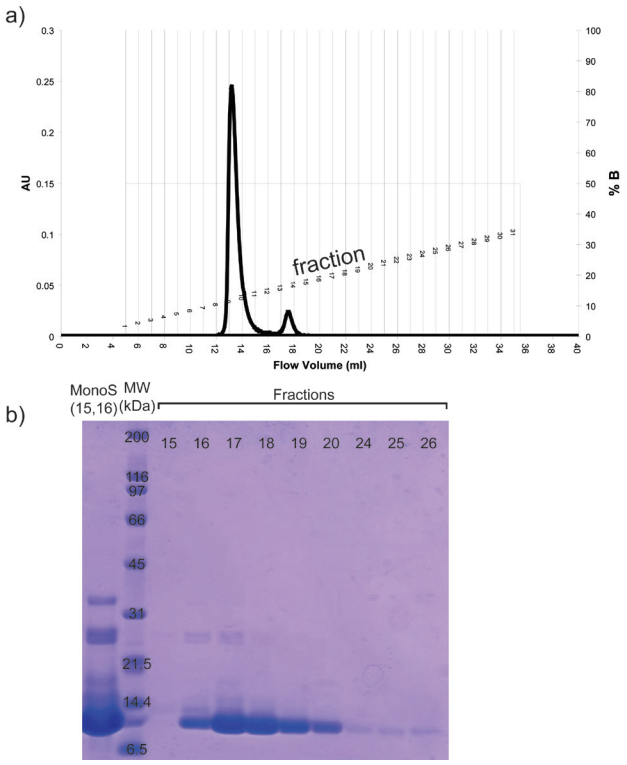


Figure C.6 Secondary purification and analysis of the second Exd preparation. a) FPLC trace. See text for details of purification. b) SDS-PAGE gel stained with Coomassie blue shows the content of fractions 15-20 and 24-26.

C.3 References

1. Stoscheck, C. M. *Methods in Enzym.* **1990**, *182*, 50-69.

2. Gill, S. C., von Hippel, P. H. *Anal. Biochem.* **1989**, *182*, 319-326.

3. Arndt, H.-D., Hauschild, K. E., Sullivan, D. P., Lake, K., Dervan, P. B., Ansari, A. Z. *J. Am. Chem. Soc.* **2003**, *125*, 13322-13323.

Appendix D

Solid-Phase Synthesis of Polyamides Using Marshall-Liener Resin

This work was done in collaboration with Rachel Wang and Mike Brochu (Dervan Group; Caltech).

D.1 A Brief Review of Solid-Phase Polyamide Synthesis

The solid-phase synthesis of DNA-binding pyrrole-imidazole polyamides was first described in the mid-1990s by the Dervan group.[1] The commercially available Boc-β-Ala-PAM resin solid support allows rapid synthesis of polyamides using Boc-protected monomers of *N*-methylpyrrole (pre-activated Boc-Py-OBt), *N*-methylimidazole (Boc-Im-OH), and γ-aminobutyric acid (Boc-γ-OH). Dimers of *N*-methylimidazole (Boc-Py-Im-OH and Boc-γ-Im-OH) are prepared by solution-phase methods due to poor solid-phase imidazole coupling to pyrrole and γ-aminobutyric acids. Nucleophilic cleavage of Boc-β-Ala-PAM resin with *N,N*-dimethylaminpropylamine (Dp) affords polyamides with -β-Dp tails which possesses the occasionally undesirable W over S selectivity[2] (where W = A or T and S = G or C) and which appears to often be a negative determinant for cell uptake.[3] The "β-less" Boc-Py-PAM resin has been reported to exhibit poor susceptibility to nucleophilic cleavage in the presence of amines such as Dp,[4] although cleavage to the corresponding carboxylic acid is possible using palladium.[5] The synthesis of polyamides using Kaiser oxime resin allows covalently attachment of Py monomers directly to the solid support through an oxime functional group. The oxime ester shows facile cleavage in the presence of a variety of nucleophiles to yield polyamides without a β-alanine spacer.[6] Care must be taken to minimize the amount of TFA during BOC deprotection (20% for Py and 50% for Im compared to 80% for PAM). Although the yield and purity of polyamides derived from oxime resin are often comparable to those from PAM resin, sometimes synthesis using oxime resin can be problematic and low-yielding.

A number of additional methods have been reported for the solid-phase synthesis of polyamides in efforts to obtain higher yields at minimal cost while still enabling flexibility of synthesis. Specifically, Fmoc-based solid-phase polyamide synthesis has also been developed using Wang resin[7] and MBHA resin[8] and both produce as good or better yields and purities as Boc-based chemistry. The Sugiyama group frequently employs Fmoc-based synthesis,[9] but the Boc-based chemistry is still frequently used by the Dervan group.

Safety-catch resins have also been reported employing both Fmoc and Boc protection strategies to produce comparable yields to PAM and oxime resins,[10-11] but have not garnered widespread use. One report has suggested improvements for Boc-based polyamide synthesis using alternative coupling reagents (i.e., HATU) and scavengers (phenol and water),[12] but one disadvantage is that the coupling reagent is currently more expensive than traditional ones (i.e., HBTU and DCC). More convergent Boc-based polyamide synthesis using dimers and trimers for coupling have also been reported to give improved yields,[13] but this often requires the time-consuming preparation of unique oligomers prior to the solid-phase synthesis of the target polyamide. Similarly, scattered reports of complete solution-phase synthesis have shown high yields (~1 g) for polyamides are obtainable,[14-15] but the approach is not as rapid as solid-phase methodology. A putative liquid-solid phase system has also been reported, but only small oligomers have been synthesized thus far using this approach.[16] This appendix describes the use of 4-hydroxyphenylsulfanylmethyl polystyrene resin, originally reported by Marshall and Liener in the 1970s for peptide synthesis,[17] as an alternative solid-phase support for rapid Boc-based polyamide synthesis. Similar to oxime resin, polyamides can be synthesized without a C-terminal β-alanine, but deprotection with higher amounts of TFA is possible.

D.2 Experimental Protocols

Resin loading. A total of 300 mg 4-hydroxyphenylsulfanylmethyl polystyrene resin (i.e Marshall-Liener resin, NovaBiochem;1.70 mmol/g, ~0.51 mmol scale) was treated with 2 equivalents of Boc-Py-OBt in 5 mL DMF and 3 equivalents of DIEA for 48 hours at RT. Successful loading has been accomplished in as short as 24 hours.

Coupling. Couplings were performed as described previously for Boc-β-Ala-PAM resin.[1]

Deprotection. All deprotection reactions were performed with 50% TFA/DCM for 25 min.

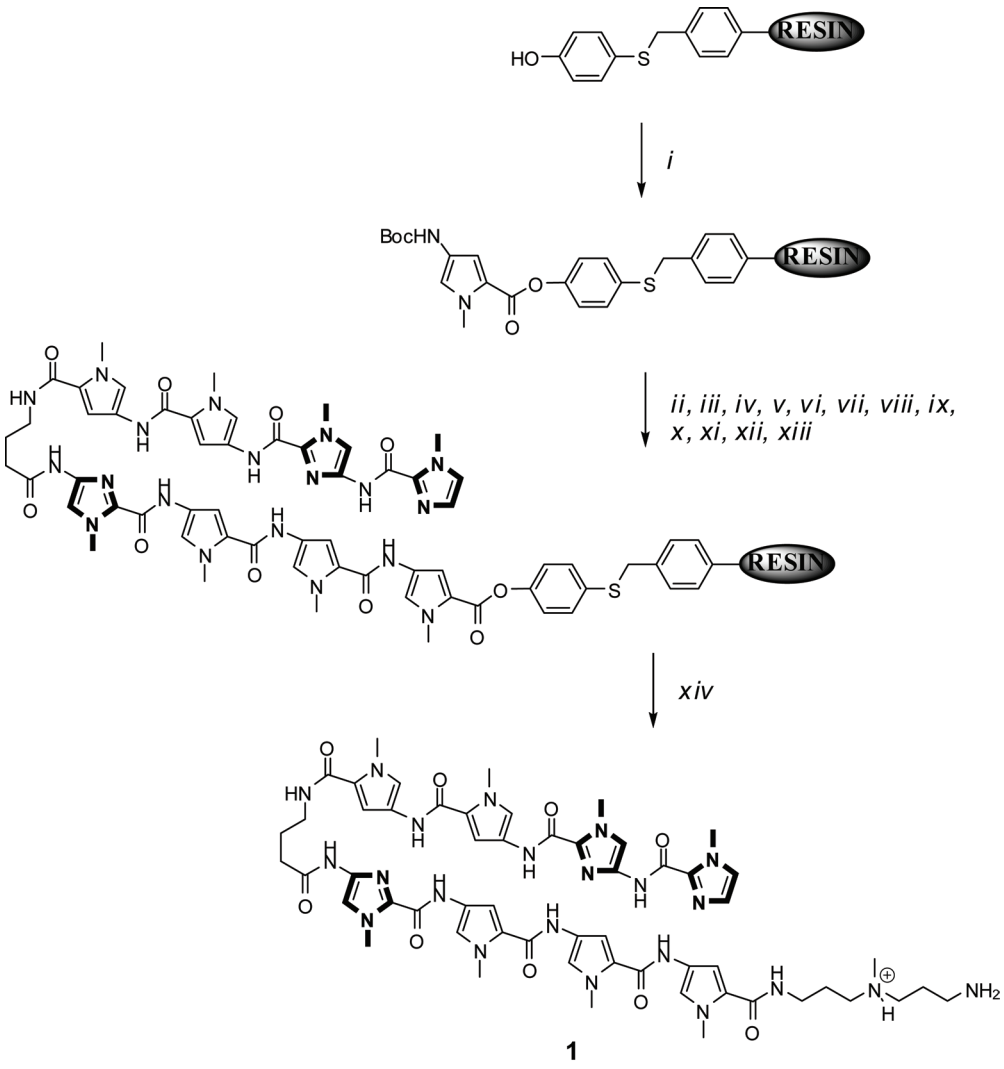


Figure D.1 Solid phase synthetic scheme for ImImPyPy-γ-ImPyPyPy-β-(+)-NH_2 from commercially available 4-hydroxyphenylsulfanylmethyl polystyrene resin, i.e., Marshall-Liener resin: *i*) 2 equivalents Boc-Py-OBt, 3 equivalents DIEA, DMF, 24 h *ii*) 50% TFA/DCM, 30 min *iii*) 2 equivalents Boc-Py-OBt, 3 equivalents DIEA, DMF, 2 h, *iv*) 50% TFA/DCM, 30 min, *v*) 2 equivalents Boc-Py-OBt, 3 equivalents DIEA, DMF, 2 h *vi*) 50% TFA/DCM, 30 min, *vii*) 2 equivalents Boc-γ-Im-OH, 1.9 equivalents HBTU, 3 equivalents DIEA, DMF, 2 h, *viii*) 50% TFA/DCM, 30 min, *ix*) 2 equivalents Boc-Py-OBt, 3 equivalents DIEA, DMF, 2 h, *x*) 50% TFA/DCM, 30 min, *xi*) 2 equivalents Boc-Py-OBt, 3 equivalents DIEA, DMF, 2 h, *xii*) 50% TFA/DCM 30 min, *xiii*) 2 equivalents ImIm-OH, 1.9 equivalents HBTU, 3 equivalents DIEA, DMF, 2 h, *xiv*) 3,3′-diamino-*N*-methyl-dipropylamine, 50 °C, 24 h.

Washing. Following each deprotection the resin was drained and rinsed 2x DCM, 1x 4:1 DMF:DIEA, and 2x DMF. Following each coupling the resin was drained and rinsed 2x DMF and 2x DCM.

Analytical cleavages. To monitor the progress of synthesis a small sample of the resin was treated with a small volume of *N,N*-dimethylaminopropylamine for 15 min at 95 °C, filtered, diluted in 0.1% TFA in H_2O, and analyzed by HPLC.

Preparative cleavage. Similar to a couple relatively recent reports,[18-19] cleavage from Marshall-Liener resin did not require sulfur oxidation with hydrogen peroxide as originally suggested.[17]

ImImPyPy-γ-ImPyPyPy-(+)-NH_2 (1). 105 mg of resin containing the polyamide was treated with 1 mL of neat 3,3′-diamino-*N*-methyl-dipropylamine and incubated at 50 °C for 16.5 hours. The resin filtered and the filtrate was diluted with 0.1% TFA in H_2O and acetonitrile and purified 2x by preparative HPLC. Fractions with product identified by MALDI-TOF MS were flash frozen and lyophilized to yield a solid (11.0 μmol, 14.6% yield, 80-90% pure). (MALDI-TOF MS) $[M+H]^+$ calcd for $C_{56}H_{71}N_{22}O_9^+$ 1195.6, observed 1195.7.

ImImPy$^{(propylamine)}$Py-γ-ImPyPyPy-(+)-NHBoc (2). 50 mg of resin containing the polyamide was treated with 300 mg of *N*-Boc-3,3′-diamino-*N*-methyl-dipropylamine in 500 μL DMF at 50 °C for 24 hours. The resin was filtered and purified as described for above (2.3 μmol, 8.5% yield, ~90% pure). (MALDI-TOF MS) $[M+H]^+$ calcd for $C_{63}H_{84}N_{23}O_{11}^+$ 1338.7, observed 1338.7.

D.3 References

1. Baird, E. E., Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141-6146.

2. Swalley, S. E., Baird, E. E., Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 1113-1120.

3. Edelson, B. S., Best, T. P., Olenyuk, B., Nickols, N. G., Doss, R. M., Foister, S., Heckel, A., Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802-2818.

4. Barid, E. E. Ph.D. Thesis, California Institute of Technology, 1999.

5. Herman, D. M., Turner, J. M., Baird, E. E., Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 1121-1129.

6. Belitsky, J. M., Nguyen, D. H., Wurtz, N. R., Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 2767-2774.

7. Wurtz, N. R., Turner, J. M., Baird, E. E., Dervan, P. B. *Org. Lett.* **2001**, *3*, 1201-1203.

8. Vázquez, E., Caamaño, A. M., Castedo, L., Mascareñas, J. L. *Tet. Lett.* **1999**, *40*, 3621-3624.

9. Ayama, H., Saito, T., Bando, T., Fukuda, N., Sugiyama, H. *Nucleic Acids Res. Supp.* **2003**, *3*, 67-68.

10. Fattori, D., Kinzel, O., Ingallinella, P., Bianchi, E., Pessi, A. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 1143-1147.

11. Marquez, M. A. Ph. D. Thesis, California Institute of Technology, 2005.

12. Krutzik, P. O., Chamberlain, A. R. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 2129-2132.

13. Sinyakov, A. N., Feshchenko, M. V., Ryabinin, V. A. *Russ. J. Bioorg. Chem.* **2003**, *4*, 402-404.

14. Sinyakov, A. N., Feshchenko, M. V., Ryabinin, V. A. *Russ. J. Bioorg. Chem.* **2004**, *30*, 98-99.

15. Yuan, G., Tang, F. *ARKIVOC* **2003**, *ii*, 32-37.

16. König, B., Rödel, M. *Chem. Commun.* **1998**, 605-606.

17. Marshall, D. L., Liener, I. E. *J. Org. Chem.* **1970**, *35*, 867-868.

18. Fantauzzi, P. P., Yager, K. M. *Tet. Lett.* **1998**, *39*, 1291-1294.

19. Breitenbucher, J. G., Johnson, C. R., Haight, M., Phelan, J. C. *Tet. Lett.* **1998**, *39*, 1295-1298.

Appendix E

Dimerizer-Exd-DNA Crystallization Trials

E.1 Experimental Details

Crystallization of a polyamide-peptide/Exd/DNA complex was inspired by the three related HOX/PBC/DNA ternary complex crystal structures (Figure E.1).[1-3] Initial crystallization experiments were attempted using the originally reported conjugate **1** (HPLC purified, >99% pure) (Figure E.2).[4] Crude DNA inspired from the Exd/Ubx/DNA structure[1] was obtained from Integrated DNA Technologies and purified by non-denaturing PAGE utilizing UV shadowing and band extraction. DNA was eluted from the gel with 2M NaCl and filtered. DNA was precipitated with iPrOH (1x) and 75% EtOH in H_2O (2x) and lyophilized before being dissolved in 18 MΩ water. DNA purity was assessed by analytical HPLC using a Clarity oligo-RP column (150 mm × 4.60 mm, 5 μm, Phenomenex, A buffer = 0.05 M TEAA (pH 7, Fluka), B buffer = acetonitrile). DNA was annealed prior to addition of conjugate **1** or Exd by heating to 95 °C followed by slowly cooling to RT (~20 °C). Exd expressed by Peter Snow (see section C.1) was concentrated to ~19.4 mg/mL (~2.04 mM) at 4 °C using a Microcon YM-3 (MWCO of 3 kDa) and diluted down to ~10 mg/mL in crystallization buffer (10 mM HEPES, 1 mM EDTA, 1 mM DTT, pH 8.0) similar to a previously reported procedure for Pbx.[2] A total of 60 nmol of each component was added in the following order: 24 μL 2.5 mM dsDNA, 46 μL H_2O (18 MΩ), 50 μL 1.2 mM conjugate **1** (pre-sonicated to ensure homogeneity), and 60 μL Exd, which was incubated for ~3 hours at 20 °C. Crystallization was performed using the hanging drop method, i.e., 1:1 mixture of precipitant to complex solution. Crystallization screens from Hampton (Natrix and Crystal Screen Cryo) and Emerald Biosystems (Cryo I/II) and follow-up grid screens (precipitate concentration in 1% increments and pH in 0.1 unit increments) yielded several crystals (Figure E.3). All crystals diffracted poorly (best resolution ~6-8 Å) at room temperature and many cryoprotectant solutions destroyed crystals (30% PEG 400, PFO-X125, 25% glycerol, 25% ethylene glycol, 25% isopropanol, 5% (2R,3R)-(+)-2,3-butanediol, 25% DMSO, 1.4 M LiCl, and 0.68 M $Mg(OAc)_2$). Initial screens with poorly

formed crystals showed that some appeared to tolerate carbohydrate-based cryoprotectants better (D-(+)-raffinose, D-(+)-trehelose, and D-(+)-glucose).

E.2 References

(1) Passner, J. M.; Ryoo, H. D.; Shen, L. Y.; Mann, R. S.; Aggarwal, A. K. *Nature* **1999**, *397*, 714-719.

(2) Piper, D. E.; Batchelor, A. H.; Chang, C. P.; Clearly, M. L.; Wolberger, C. *Cell* **1999**, *96*, 587-597.

(3) LaRonde-LeBlanc, N. A.; Wolberger, C. *Genes Dev.* **2003**, *17*, 2060-2072.

(4) Stafford, R. L., Arndt, H.-D., Brezinski, M. L., Ansari, A. Z., Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2660-2668.

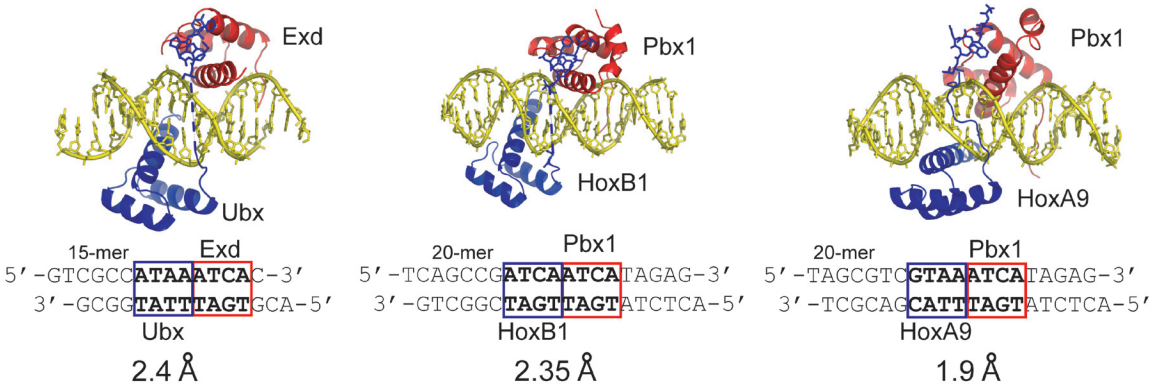


Figure E.1 Summary of HOX/PBC/DNA ternary complex crystal structures.

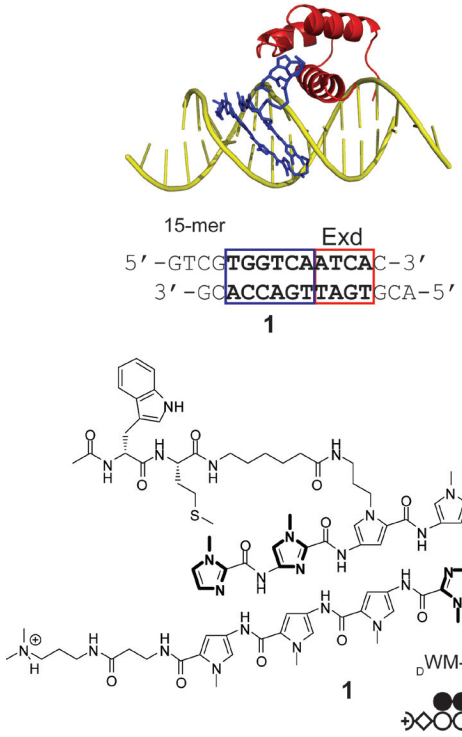


Figure E.2 Design of polyamide-peptide/Exd/DNA complex crystallization experiments. A model of the presumed complex is shown above. The sequence of the DNA duplex and the structure of the polyamide-peptide conjugate **1** are shown below.

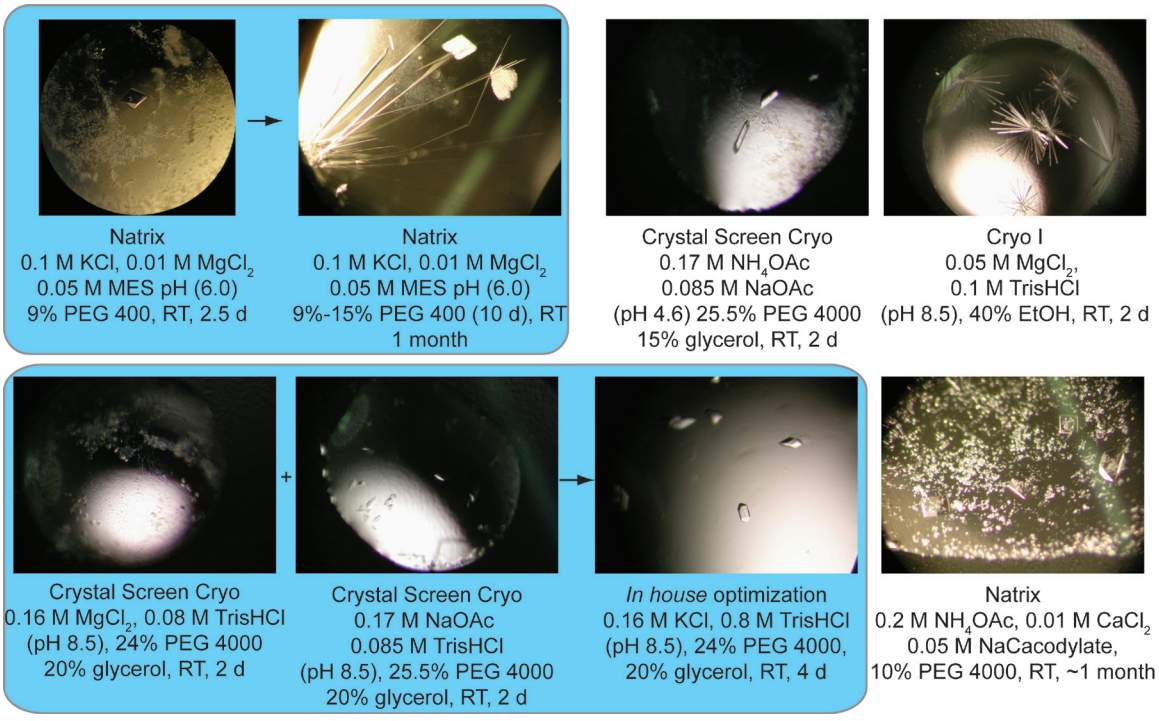


Figure E.3 Pictures of select crystals and the conditions used for crystallization.

Appendix F

Supplemental Figures

The figures in this section were taken in part from a manuscript coauthored with Peter B. Dervan (Caltech).

(Stafford, R. L, Dervan, P. B. "The Reach of Linear Protein-DNA Dimerizers" *submitted to J. Am. Chem. Soc.* **2007**)

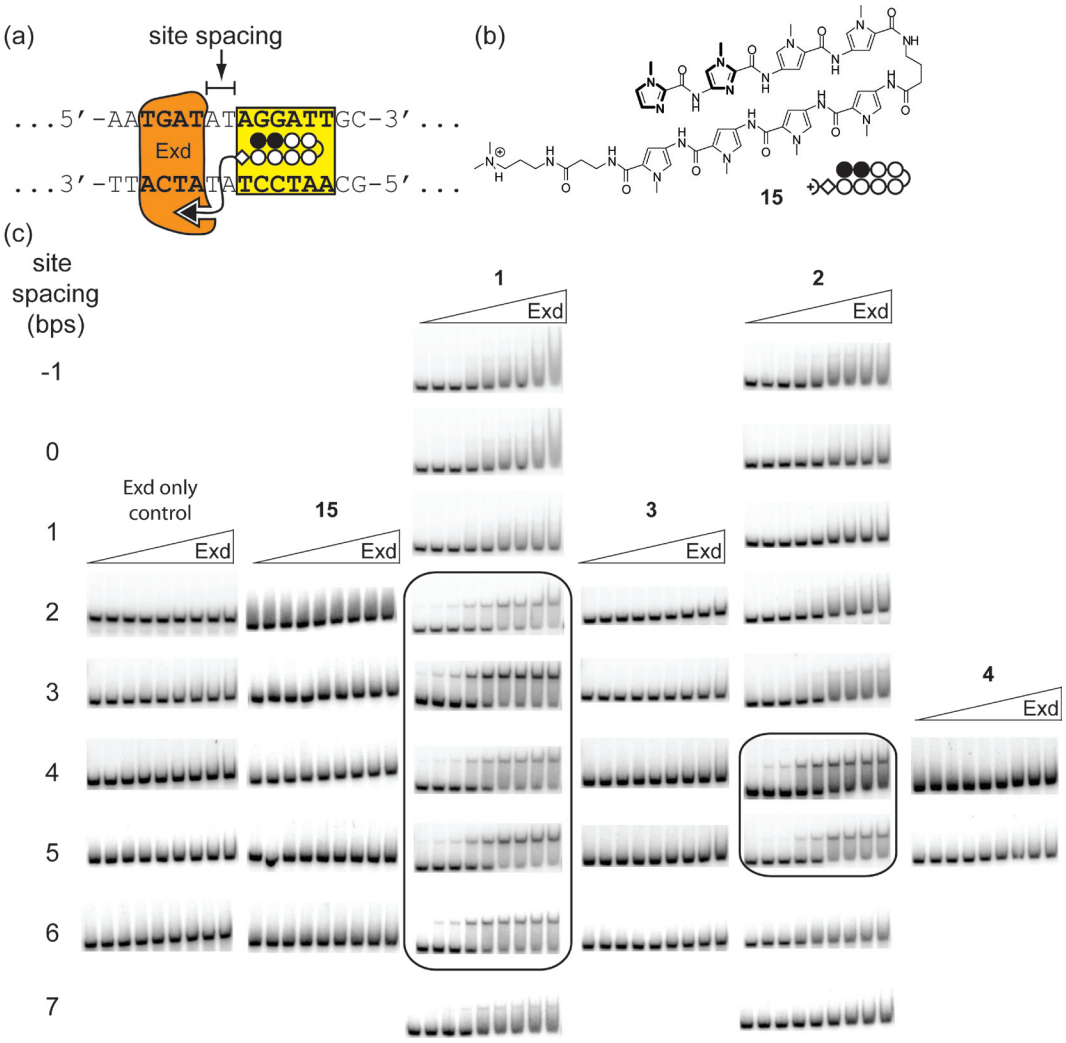


Figure F.1 Representative gel shift results for Exd in the presence of linear protein-DNA dimerizers with differently sized linkers between the two proximally oriented DNA sites *with control compounds*. (a) Schematic of the positions of the Exd and polyamide DNA-binding sites showing the A,T spacer region which was varied. (b) Structure of parent polyamide ImImPyPy-γ-PyPyPyPy-β-Dp (**15**). (c) Gel shift experiments show that **1** forms a stable complex with Exd and DNA with site spacings between 2-6 bps, whereas **2** only forms stable complexes from 4-5 bps. It is also shown that Exd by itself, in the presence of the parent polyamide **15**, or double-AA mutants **3** and **4**, does not yield a stable complex with DNA at these sites. A concentration of 50 nM of the indicated compound was added to all lanes and Exd was titrated from 10 pM to 100 nM.

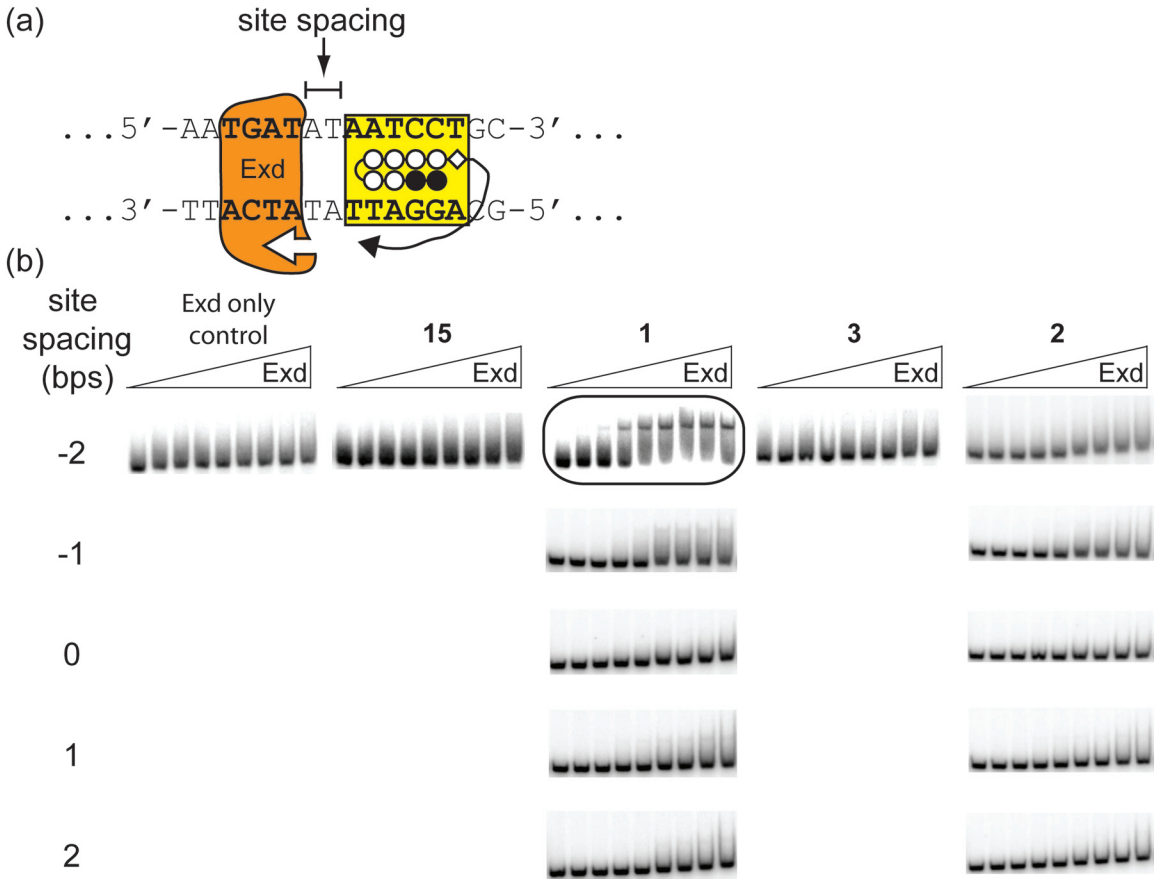


Figure F.2 Representative gel shift results for Exd in the presence of linear protein-DNA dimerizers with differently sized linkers between the two distally oriented DNA sites *with control compounds*. (a) Schematic of the positions of the Exd and polyamide DNA-binding sites showing the A,T spacer region which was varied. (b) Gel shift experiments show that **1** forms a stable complex with Exd and DNA when there is a 2 bp overlap, i.e., a site spacing of -2, between the two sites, whereas **2** does not form a complex with Exd at any of these sites. It is also shown that Exd by itself, in the presence of parent polyamide **15**, or double-AA mutant **3**, does not yield a stable complex with DNA at this site. A concentration of 50 nM of the indicated compound was added to all lanes and Exd was titrated from 10 pM to 100 nM.

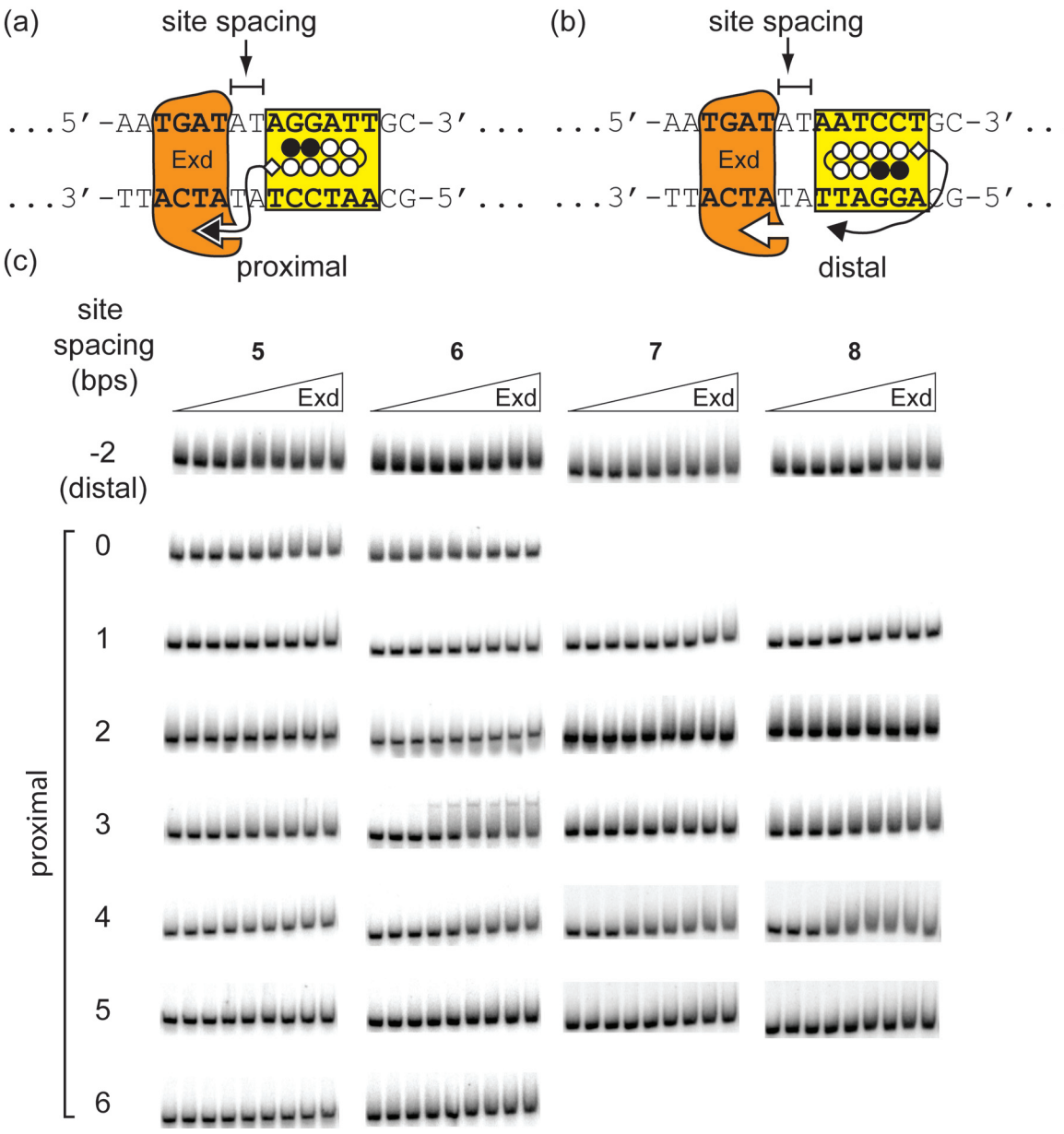


Figure F.3 Representative gel shift results for Exd in the presence of WM dipeptide linear protein-DNA dimerizers with differently sized spacers between the two DNA sites. (a) Schematic of the positions of the Exd and polyamide DNA-binding sites showing the poly-AT spacer region which was varied in the proximal orientation. (b) Schematic of the positions of the Exd and polyamide DNA-binding sites showing the A,T spacer region which was varied in the distal orientation. (c) Gel shift experiments show that none of the site spacings lead to stable complex formation for compounds **5**-**8**. A weak complex with compound **6** is observed at a spacing of 3 bp in the proximal orientation ($\Theta_{app} = 0.3 \pm 0.1$). A concentration of 50 nM of the indicated compound was added to all lanes and Exd was titrated from 10 pM to 100 nM.

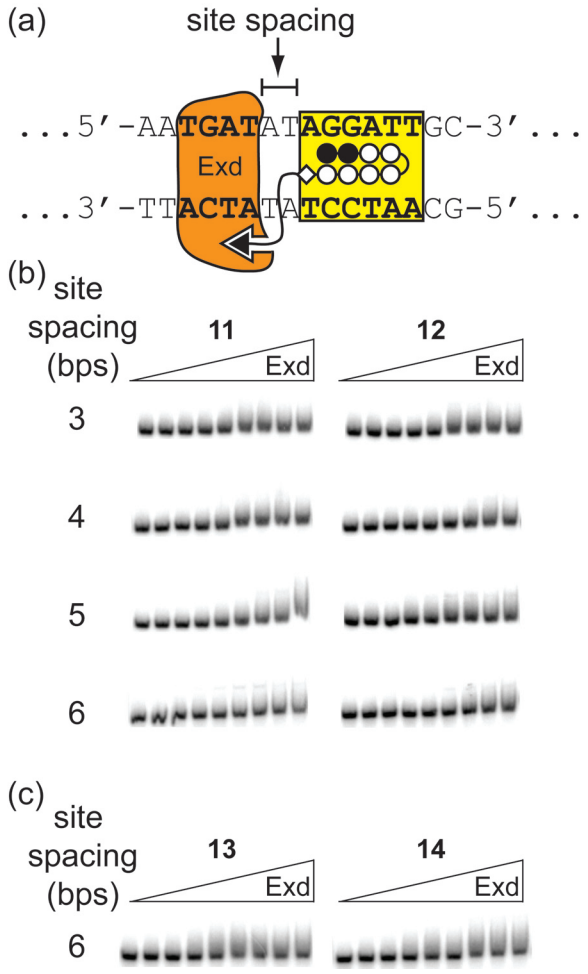


Figure F.4 Representative gel shift results for Exd in the presence of WMK tripeptide linear protein-DNA dimerizers with differently sized spacers between the two proximally oriented DNA sites. (a) Schematic of the positions of the Exd and polyamide DNA-binding sites showing the A,T spacer region which was varied in the proximal orientation. (b) Gel shift experiments show that none of the site spacings lead to stable complex formation for compounds **11** or **12**. (c) Gel shift experiments show that acetylated WMK tripeptide conjugates **13** and **14** do not lead to a stable complex at the proximal 6 bp spacing site. A concentration of 50 nM of the indicated compound was added to all lanes and Exd was titrated from 10 pM to 100 nM.

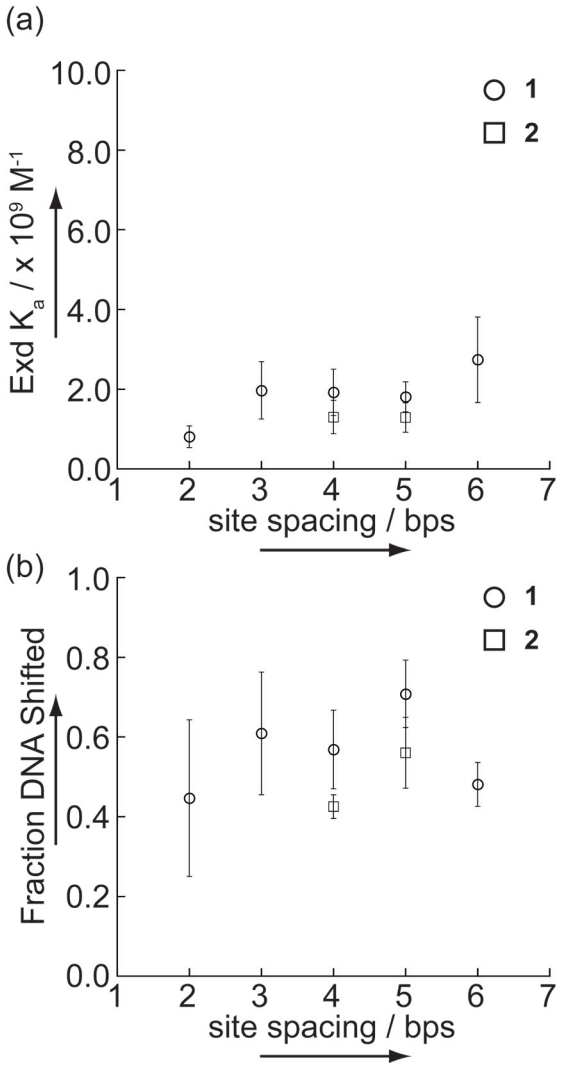


Figure F.5 The dependence of binding affinity (Ka) and the apparent fraction of DNA shifted with respect to the site spacing is shown. (a) The binding affinity of Exd does not change by more than 3-fold as the site spacing is changed in the presence of compounds **1** and **2** when a complex is observed. (b) The amount of complex formed also does not change substantially as the site spacing is changed.

Proximal -1 duplex (47 bps):
5′-GATCTCCCGGCGAA**TGATGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTACCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 0 duplex (48 bps):
5′-GATCTCCCGGCGAA**TGATTGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTAACCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 1 duplex (49 bps):
5′-GATCTCCCGGCGAA**TGAT**A**TGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**T**ACCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 2 duplex (50 bps):
5′-GATCTCCCGGCGAA**TGAT**AT**AGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**TA**TCCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 3 duplex (51 bps):
5′-GATCTCCCGGCGAA**TGAT**ATA**AGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**TAT**TCCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 4 duplex (52 bps):
5′-GATCTCCCGGCGAA**TGAT**ATAT**AGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**TATA**TCCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 5 duplex (53 bps):
5′-GATCTCCCGGCGAA**TGAT**ATATA**TGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**TATAT**ACCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 6 duplex (54 bps):
5′-GATCTCCCGGCGAA**TGAT**ATATAT**AGGATT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**TATATA**TCCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Proximal 7 duplex (55 bps):
5′-GATCTCCCGGCGAA**TGAT**ATATATA**TGGATT**GCGTCGGCGCCACTGTCACCCGGA-3
3′-CTAGAGGGCCGCTT**ACTA**TATATAT**ACCTAA**CGCAGCCGCGGTGACAGTGGGCCT-5

Figure F.6 List of the proximal series of DNA duplexes used in gel shift experiments.

Distal -2 duplex (46 bps):
5′-GATCTCCCGGCGAA**TGATTCCT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTAAGGA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Distal -1 duplex (47 bps):
5′-GATCTCCCGGCGAA**TGATATCCT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTATAGGA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Distal 0 duplex (48 bps):
5′-GATCTCCCGGCGAA**TGATTATCCT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTAATAGGA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Distal 1 duplex (49 bps):
5′-GATCTCCCGGCGAA**TGAT**A**TATCCT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**T**ATAGGA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Distal 2 duplex (50 bps):
5′-GATCTCCCGGCGAA**TGAT**AT**AATCCT**GCGTCGGCGCCACTGTCACCCGGA-3′
3′-CTAGAGGGCCGCTT**ACTA**TA**TTAGGA**CGCAGCCGCGGTGACAGTGGGCCT-5′

Figure F.7 List of the distal series of DNA duplexes used in gel shift experiments.