



Published in final edited form as:

*Bioorg Med Chem.* 2007 November 15; 15(22): 6927–6936.

## $\alpha$ -Diaminobutyric Acid-Linked Hairpin Polyamides

Michelle E. Farkas, Sherry M. Tsai, and Peter B. Dervan\*

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

### Abstract

A hairpin polyamide-chlorambucil conjugate linked by  $\alpha$ -diaminobutyric acid ( $\alpha$ -DABA) has been shown to have interesting biological properties in cellular and small animal models. Remarkably, this new class of hairpin polyamides has not been previously characterized with regard to energetics and sequence specificity. Herein we present a series of pyrrole-imidazole hairpin polyamides linked by  $\alpha$ -DABA and compare them to polyamides containing the standard  $\gamma$ -DABA turn unit. The  $\alpha$ -DABA hairpins have overall decreased binding affinities. However,  $\alpha$ -DABA polyamide-chlorambucil conjugates are sequence-specific DNA alkylators with increased specificities. Affinity cleavage studies of  $\alpha$ -DABA polyamide-EDTA conjugates confirmed their preference for binding DNA in a forward hairpin conformation. In contrast, an unsubstituted glycine-linked polyamide prefers to bind in an extended binding mode. Thus, substitution on the turn unit locks the  $\alpha$ -DABA polyamide into the forward hairpin binding motif.

### 1. Introduction

Pyrrole-imidazole polyamides are a class of synthetic ligands that can be programmed to bind a broad repertoire of DNA sequences with affinities and specificities similar to DNA-binding proteins.<sup>1, 2</sup> Eight-ring hairpin polyamides have been shown to localize to the nuclei of living cells<sup>3, 4</sup> and regulate transcription by interfering with transcription factor-DNA interfaces in the promoters of endogenous genes.<sup>5–9</sup> Conjugation of polyamides to the DNA alkylating agent chlorambucil produces molecules capable of sequence-specific covalent reaction with DNA,<sup>10</sup> previously demonstrated for targeted sequences in the HIV-1 promoter, simian virus 40, and the histone H4c gene.<sup>11–13</sup>

DNA recognition by polyamides depends on side-by-side, aromatic amino acid pairings in the minor groove. Pairing rules have been established for DNA base pair recognition: an *N*-methylpyrrole/*N*-methylimidazole (Py/Im) pair recognizes C•G, the reverse (Im/Py) specifies G•C, and a Py/Py pair targets both A•T and T•A.<sup>2</sup> For some sequences, replacement of Py with a flexible  $\beta$ -alanine ( $\beta$ ) residue can enhance polyamide binding affinity by relaxing ligand curvature, affording a more optimal fit in the minor groove of DNA.<sup>14</sup>

Although Py-Im polyamides will bind in the minor groove as anti-parallel dimers,<sup>15</sup> polyamides bind DNA optimally as a single oligomer in a hairpin structure.<sup>16</sup> Early studies demonstrated the optimal length for the turn element is three methylene units, via the use of  $\gamma$ -aminobutyric acid as the turn residue.<sup>16</sup> The polyamide hairpin binds in the ‘forward’

\*Corresponding author: Tel.: +1 626 395 6002; Fax: +1 626 683 8753; E-mail: dervan@caltech.edu.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

direction N→C with respect to the 5'→3' direction of the DNA strand.<sup>17</sup> Utilization of the substituted turn (*R*)-2,4-diaminobutyric acid ( $\gamma$ -DABA) further enhances DNA binding affinity.<sup>18</sup> Notably, employment of a linkage containing one methylene unit – provided by the use of glycine as the turn amino acid – resulted in polyamides that bound DNA in an extended 2:1 binding mode, rather than as a 1:1 hairpin complex.<sup>16, 19–21</sup> In contrast, the  $\gamma$ -aminobutyric acid linkage appears to prefer the U-shaped conformation rather than an extended dimeric conformation.<sup>20</sup>

We recently reported a new turn unit with remarkable properties in small animal models: (*R*)- $\alpha$ -diaminobutyric acid ( $\alpha$ -DABA), which is formally a substituted glycine linkage since it also includes a single methylene unit within the hairpin turn.<sup>22</sup> This initial study characterized an  $\alpha$ -DABA polyamide-chlorambucil conjugate and compared its biological activity to its  $\gamma$ -DABA analog. A dramatic difference in toxicity was observed for the two conjugates: normal mice treated with the  $\alpha$ -DABA polyamide-chlorambucil conjugate did not differ significantly from control, while the analogous  $\gamma$ -DABA conjugate was lethal.<sup>22</sup> Notably, only the  $\alpha$ -DABA conjugate has been demonstrated to arrest cancer growth in a SW620 xenograft nude mouse tumor model without apparent toxicity.<sup>13</sup>

Given the striking *in vivo* effects demonstrated by the  $\alpha$ -DABA polyamide conjugate, this new turn unit merits further study. In particular,  $\alpha$ -DABA polyamides appear to bind in a hairpin conformation and not as an extended dimer, as would be expected based on studies of glycine-linked polyamides.

We report herein the characterization of  $\alpha$ -DABA polyamides. A series of  $\alpha$ -DABA polyamides and their chlorambucil conjugates were synthesized, along with their  $\gamma$ -DABA analogs. DNA binding properties of the parent molecules were established by DNase I footprinting titration experiments, and DNA alkylation properties of the chlorambucil conjugates were determined by thermal cleavage assay experiments. Importantly, we study the binding orientation of  $\alpha$ -DABA polyamides and address their ability to bind as hairpins by comparing them to their glycine-linked analogs.

## 2. Results and discussion

### 2.1 Synthesis of the polyamide series

Three core ring sequences were selected for characterization (Fig. 1). Two standard eight-ring hairpin sequences, as well as a hairpin containing a  $\beta$ /Py pair, were chosen. Six parent hairpin polyamides were thus prepared: three molecule pairs which code for different sequences according to the pairing rules. For each pair, one molecule was linked via the  $\alpha$ -DABA turn unit (polyamides **1**, **3**, and **5**), and the other was connected with the  $\gamma$ -DABA turn (polyamides **2**, **4**, and **6**). Polyamides were synthesized on Pam resin using standard solid phase methods.<sup>22, 23</sup> The chlorambucil conjugates of the six parent hairpin polyamides were also prepared based on previously published methods.<sup>22</sup>

### 2.2 Plasmid design

Polyamides **1** and **2** were designed to bind the sequence 5'-WWGGWGW-3' (where W = A or T). The plasmid pMFST2 was designed with an insert containing a single match site ('M' = 5'-TAGGTGT-3') and two single base pair mismatch sites ('A' = 5'-TAGGTCT-3' and 'B' = 5'-TAGCTGT-3') for polyamides **1** and **2** (Fig. 2A).<sup>22</sup>

A single plasmid, pMFST, was prepared to characterize polyamides **3-6** (Fig. 2B). Polyamides **3** and **4** were targeted to the sequence 5'-WWGGWCW-3', and polyamides **5** and **6** were designed to bind 5'-WWGWGGW-3'. The designed insert for pMFST contains match sites for each set of polyamides ('M1' = 5'-TAGTGGT-3', targeted to **5** and **6**, and 'M2' = 5'-

ATGGTCA-3', targeted to **3** and **4**), as well as two single base pair mismatch sites for each ('A' = 5'-ATATGGT-3' and 'B' = 5'-TAGTCGT-3' for **5** and **6**, and 'C' = 5'-AACGACT-3' and 'D' = 5'-TAGCACA-3' for **3** and **4**). The chlorambucil moiety is expected to alkylate at adenines proximal to the polyamide binding site.<sup>11, 22</sup> Thus, for the purposes of the thermal cleavage assays, adenines were placed next to the 3' end of the match sites and at least one of the mismatch sites.

### 2.3 DNA binding affinity and sequence specificity

Quantitative DNase I footprinting titrations were performed with the parent polyamides to measure their binding site affinities and specificities (Fig. 3 and Table 1). The  $\alpha$ -DABA polyamide **1** and its  $\gamma$ -DABA analog **2** were footprinted on the 5' <sup>32</sup>P-labeled 280 base pair PCR product of plasmid pMFST2 (Fig. 3A–3B).  $\alpha$ -DABA polyamide **1** bound the match site M with  $K_a = 3.0 \times 10^8 \text{ M}^{-1}$  and the mismatch site A with  $K_a = 8.5 \times 10^7 \text{ M}^{-1}$  (Table 1). Binding of **1** at the mismatch site B was not detected. In contrast, the  $\gamma$ -DABA polyamide **2** bound the match site M with  $K_a = 4.4 \times 10^{10} \text{ M}^{-1}$  and the mismatch sites A and B with  $K_a = 1.1 \times 10^{10} \text{ M}^{-1}$  and  $K_a = 1.1 \times 10^9 \text{ M}^{-1}$ , respectively. The  $\gamma$ -DABA polyamide **2** displays a 147-fold increased binding affinity over the  $\alpha$ -DABA polyamide **1** for their match site M.

The  $\alpha$ -DABA polyamides **3** and **5** and their  $\gamma$ -DABA analogs **4** and **6**, respectively, were footprinted on the 5' <sup>32</sup>P-labeled 284 base pair PCR product of plasmid pMFST (Fig. 3C–3F).  $\alpha$ -DABA polyamide **3** bound its match site M2 with  $K_a = 4.4 \times 10^9 \text{ M}^{-1}$  and the mismatch site C with  $K_a = 1.5 \times 10^8 \text{ M}^{-1}$ . There was no detectable binding at mismatch site D. However, the analogous  $\gamma$ -DABA polyamide **4** bound the match site M2 with  $K_a = 4.9 \times 10^{10} \text{ M}^{-1}$  and the mismatch sites C and D with  $K_a = 1.8 \times 10^{10} \text{ M}^{-1}$  and  $K_a = 3.9 \times 10^8 \text{ M}^{-1}$ , respectively. The  $\alpha$ -DABA polyamide **3** binds its match site with 11-fold decreased affinity relative to the analogous  $\gamma$ -DABA polyamide **4**.

$\alpha$ -DABA polyamide **5** bound its match site M1 with  $K_a = 8.0 \times 10^8 \text{ M}^{-1}$  and its mismatch site B with  $K_a = 6.2 \times 10^7 \text{ M}^{-1}$ . Binding was not observed at mismatch site A. In comparison, the  $\gamma$ -DABA polyamide **6** binds the match site M1 with  $K_a = 4.3 \times 10^{10} \text{ M}^{-1}$  and the mismatch sites A and B with  $K_a = 9.9 \times 10^8 \text{ M}^{-1}$  and  $K_a = 6.2 \times 10^9 \text{ M}^{-1}$ , respectively. The  $\alpha$ -DABA polyamide **5** shows a 54-fold decreased affinity for its match site compared to the analogous  $\gamma$ -DABA polyamide **6**.

These results indicate that  $\alpha$ -DABA polyamides are proficient DNA binders, with decreased affinity and comparable or improved specificity relative to  $\gamma$ -DABA molecules. The decrease in affinity is dependent on the polyamide core ring sequence. Additional comparison of molecules corroborates these conclusions (Supplemental Fig. 1). Varying the heterocyclic order of polyamides has also been shown to affect binding affinities in  $\gamma$ -DABA molecules.<sup>24</sup>

We previously reported the  $\alpha$ -DABA polyamide of sequence ImIm $\beta$ Im-(R)<sup>H2N</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp (where Dp = 3-(dimethylamino)-propylamine) binds the match site M of pMFST2 with  $K_a = 6.7 \times 10^9 \text{ M}^{-1}$ .<sup>22</sup> Polyamide **1** has an identical structure except for the replacement of a  $\beta$  residue in the core sequence with a Py, and it binds the same match site M of pMFST2 with 22-fold decreased affinity. Polyamide **5** also has the same structure except for the movement of the  $\beta$  residue from the third position in the sequence to the second, and it binds its designed match site with 8-fold decreased affinity. These results suggest the presence and position of the  $\beta$  residue is important in determining the DNA binding affinities of  $\alpha$ -DABA polyamides. We hypothesize that the shortened  $\alpha$ -DABA turn unit requires the flexibility provided by the  $\beta$  to optimize binding.

## 2.4 DNA alkylation properties

Thermal cleavage assays were performed with the polyamide-chlorambucil conjugates to determine their DNA alkylation properties (Fig. 4). Previous studies have shown conjugates to alkylate at N3 of adenines proximal to DNA binding sites.<sup>11, 13</sup> Attachment of the chlorambucil moiety to the parent polyamide does not appear to affect DNA binding properties.<sup>11</sup>

The  $\alpha$ -DABA conjugate **1-Chl** and its  $\gamma$ -DABA analog **2-Chl** were assayed on the 5' <sup>32</sup>P-labeled 280 base pair PCR product of plasmid pMFST2 (Fig. 4A). Alkylation was only observed in the designed insert of pMFST2 at the match site M for the  $\alpha$ -DABA conjugate **1-Chl**, whereas the  $\gamma$ -DABA conjugate **2-Chl** is more promiscuous, alkylating at the match site M and both of the mismatch sites A and B, as well as at additional, unexpected sites.  $\gamma$ -DABA conjugate **2-Chl** also alkylates at lower concentrations than conjugate **1-Chl** (Fig 4A, lanes 9).

Conjugates **3-Chl** and **4-Chl** were assayed on the 5' <sup>32</sup>P-labeled 284 base pair PCR product of plasmid pMFST (Fig. 4B). While parent polyamide **3** binds at site M2 and site C,  $\alpha$ -DABA conjugate **3-Chl** only alkylates at match site M2 in the designed insert. Alkylation is observed at both the match site M2 and the mismatch site C for the  $\gamma$ -DABA conjugate **4-Chl**. Although the parent polyamide **4** binds mismatch site D, alkylation would not be expected to occur there because an A•T base pair is not located adjacent to the turn unit upon conjugate binding. Again, alkylation is observed at lower concentrations for the  $\gamma$ -DABA conjugate **4-Chl** than the  $\alpha$ -DABA conjugate **3-Chl** (Fig. 4B, lanes 7).

Conjugates **5-Chl** and **6-Chl** were also assayed on the PCR product of plasmid pMFST (Fig. 4C). For the  $\alpha$ -DABA conjugate **5-Chl**, strong alkylation is only observed at match site M1 in the designed insert, with minor alkylation occurring at mismatch site B. In contrast, the  $\gamma$ -DABA conjugate **6-Chl** strongly alkylates at both match site M1 and mismatch site B, as well as at other unexpected sites. As previously, alkylation is not expected at mismatch site A since an A•T base pair is not located next to the turn unit upon conjugate binding. In addition,  $\gamma$ -DABA conjugate **6-Chl** is able to alkylate at a lower concentration than  $\alpha$ -DABA conjugate **5-Chl** (Fig. 4C, lanes 9).

These results indicate  $\alpha$ -DABA conjugates to be alkylators with increased specificity and decreased reactivity relative to  $\gamma$ -DABA conjugates. Thermal cleavage data from additional conjugates further support these conclusions (Supplemental Fig. 1).

## 2.5 Polyamide binding orientation

Previous footprinting studies indicated that glycine-linked polyamides bind DNA in extended, dimeric 2:1 complexes,<sup>16, 20</sup> and the structures were confirmed by NMR studies.<sup>19, 21</sup> Two possible binding modes were established (Fig. 5). In the “overlapped” binding mode, the two polyamides bind directly opposite one another with each heterocycle paired to another on the other strand. In the “slipped” binding mode, the polyamides integrate the 2:1 and 1:1 polyamide-DNA binding motifs at a single site.

Formally, the  $\alpha$ -DABA moiety is a substituted glycine turn, as glycine-linked polyamides also contain a single methylene group linking the heterocyclic subunits. Yet, based on their footprinting data (i.e. binding site sequence and size, and cooperativity as measured in the binding isotherms),  $\alpha$ -DABA polyamides bind in a 1:1 hairpin conformation.<sup>22</sup> In addition, although it is well-established that  $\gamma$ -DABA polyamides bind in the forward direction,<sup>17</sup> there is still ambiguity as to the preferred binding orientation of  $\alpha$ -DABA polyamides.

We hypothesized that a polyamide with the unsubstituted glycine linkage would favor the extended conformation, whereas substitution of the turn using the  $\alpha$ -DABA moiety would force

the polyamide to bind in a forward hairpin conformation. A similar phenomenon has been observed by Boger and coworkers, who studied a series of  $\alpha$ -substituted  $\beta$ -linked polyamides and found that a polyamide linked by (*R*)- $\alpha$ -methoxy- $\beta$  preferentially adopts a hairpin conformation, versus the unsubstituted molecule which binds in an extended conformation.<sup>25</sup> Interestingly, other substitutions disrupted binding affinity.

$\alpha$ -DABA polyamide **3** was chosen as the parent molecule for study. To verify its binding orientation, the EDTA conjugate **EDTA-3** was synthesized for affinity cleavage studies (Fig. 6). Polyamide **7** – the glycine-linked analog of **3** – was prepared, as well as its EDTA conjugate, **EDTA-7**.

Four binding orientations are possible for polyamides **3** and **7**. The expected match site for the molecules binding as forward hairpins is 5'-WWGGWCW-3', whereas for reverse-binding hairpins the expected match site is 5'-WCWGGWW-3'. In the overlapped binding mode, the expected target sequence is 5'-WWGGWCWGWCCWW-3'. It should be noted that the overlapped match site is inherently ambiguous because it contains two side-by-side copies of the forward hairpin match site. Binding affinity in the slipped mode is sensitive to the particular sequence of the 1:1 binding portion, and there are multiple slipped binding modes for a given polyamide core ring sequence.<sup>20</sup> Since an Im/Im pair is disfavored,<sup>26</sup> we studied the slipped orientation for **3** and **7** believed to be preferred for binding. We expect this binding model to give an 18 base pair match site with sequence 5'-WWWWWNWGGCCWNWWWW-3' where N = A, T, G, or C.<sup>20, 27</sup>

Plasmid pMFST5 was prepared to study the binding preferences of the selected molecules (Fig. 7A). The designed insert contains four match sites for each of the possible binding modes for polyamides **3** and **7**: slipped extended binding ('S' = 5'-ATTAACAGGCCACAATTA-3'), overlapped extended binding ('O' = 5'-ATGGTCAGACCAT-3'), reverse hairpin binding ('R' = 5'-ACTGGTA-3'), and forward hairpin binding ('M' = 5'-ATGGTCA-3').

Quantitative DNase I footprinting titrations were performed with  $\alpha$ -DABA polyamide **3** and glycine polyamide **7** (Fig. 7B and Table 2), and affinity cleavage studies were performed with  $\alpha$ -DABA conjugate **EDTA-3** and glycine conjugate **EDTA-7** (Fig. 7C–7D) on the 5' <sup>32</sup>P-labeled 301 base pair PCR product of plasmid pMFST5. A dramatic difference between the two molecules was observed. The  $\alpha$ -DABA polyamide **3** binds the forward hairpin binding site M with  $K_a = 9.8 \times 10^9 \text{ M}^{-1}$ , while no binding was detected at the slipped binding site S or reverse hairpin site R. Cleavage by **EDTA-3** was observed at match site M at the locus in agreement with forward binding and is 3'-shifted, consistent with minor groove occupancy.<sup>28</sup>

Additional binding of polyamide **3** was observed at the overlapped site O with  $K_a = 4.7 \times 10^9 \text{ M}^{-1}$ . Unequal cleavage intensities by **EDTA-3** were observed at the 3' and 5' ends of the match site. A fully anti-parallel overlapped motif would be expected to afford equal intensities. Therefore it is likely that polyamide **3** is binding as a hairpin in either of the adjacent hairpin match sites embedded in the 13 base-pair sequence O (Fig. 7D, top).

In contrast, the glycine polyamide **7** binds the slipped extended site S with  $K_a = 1.3 \times 10^9 \text{ M}^{-1}$ , and it binds the forward hairpin site M and overlapped site O with ~ 2-fold decreased affinities ( $K_a = 7.0 \times 10^8 \text{ M}^{-1}$  and  $K_a = 5.9 \times 10^8 \text{ M}^{-1}$ , respectively). No binding was observed at the reverse hairpin site R. Strong cleavage by **EDTA-7** was only observed at site S, probably due to the weaker binding affinities at the other sites. The cleavage pattern was 3' shifted, verifying minor groove occupancy of the ligand, and the observed pattern was consistent with dimeric, antiparallel binding in the slipped mode. There is again ambiguity regarding the binding mode at the overlapped site O.

It is interesting to note that though polyamide **7** is a relatively small ligand (MW = 1196), it displays specificity for an 18 base pair sequence. It is likely that the binding affinity of **7** in the slipped binding mode could be enhanced by modifying the sequence of the W tracts in the 1:1 binding portions of the molecules to a DNA sequence that favors a narrower minor groove.<sup>20</sup>

These data show a clear distinction in preferred binding modes between  $\alpha$ -DABA polyamide **3** and glycine polyamide **7**. The  $\alpha$ -DABA polyamide appears to bind exclusively in a forward hairpin conformation, while the glycine polyamide seems to exhibit mixed binding motifs, preferring the extended binding mode but also appearing to be capable of binding as a forward hairpin. Thus, substitution of the glycine turn 'locks' the polyamide into a hairpin conformation. Affinity cleavage studies performed on additional  $\alpha$ -DABA polyamides confirm their general binding in the forward hairpin mode (Supplemental Fig. 2 and 3).

### 3. Conclusion

Herein we present a characterization of  $\alpha$ -DABA polyamides. We have confirmed their ability to bind DNA and established their binding affinities to be decreased 2-fold to 147-fold relative to  $\gamma$ -DABA polyamides. The magnitude of this reduced affinity appears to be sensitive to the presence and positioning of an internal  $\beta$  residue in the polyamide structure. Remarkably, despite the reduced affinity,  $\alpha$ -DABA polyamide-chlorambucil conjugates alkylate DNA with greater specificity and lower reactivity than  $\gamma$ -DABA conjugates. Importantly, we have demonstrated that polyamides linked by the (*R*)- $\alpha$ -diaminobutyric acid moiety bind DNA in a forward hairpin binding mode. Polyamides linked by a glycine amino acid prefer to bind DNA in an extended binding mode, indicating that it is the substitution of the glycine turn unit that locks the  $\alpha$ -DABA polyamide into the hairpin conformation. While it has been previously shown that  $\alpha$ -DABA polyamide-chlorambucil conjugates are efficacious in cellular and small animal models,<sup>13, 22, 29</sup> cellular uptake experiments of polyamide-fluorescein conjugates displayed limited nuclear uptake properties (Supplemental Fig. 4, Table 1), possibly limiting the general use of  $\alpha$ -DABA polyamides *in vivo*.

### 4. Experimental

#### 4.1 General

Boc- $\beta$ -Ala-Pam resin was from Peptides International. Trifluoroacetic acid (TFA) was purchased from Halocarbon. Dichloromethane (DCM) was from Fisher Scientific. (Benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate (PyBOP) was from Novabiochem. Chlorambucil, *N,N*-diisopropylethylamine, 3-(dimethylamino)-propylamine, 3,3'-diamino-*N*-methyl-dipropylamine, *N,N*-dimethylformamide, and 1-methyl-2-pyrrolidone (NMP) were purchased from Aldrich. Oligonucleotides were synthesized by Integrated DNA Technologies. Enzymes were purchased from Roche. RNase free water was from USB and used for all footprinting and cleavage reactions.

#### 4.2 Polyamide synthesis

Polyamides were synthesized on Boc- $\beta$ -Ala-Pam resin following manual solid-phase methods using Boc-protected amino acids as previously reported.<sup>23</sup> Boc-Py-OBt (OBt = benzotriazol-1-yloxy), Boc- $\beta$ -Im-OH, Im-CCl<sub>3</sub>, Boc-Py-Py-OBt, Boc-Py-Im-OH, Boc-Im-OH, and Im-Im-OH were used as building blocks. The  $\alpha$ -DABA polyamides **1**, **3**, **5**, **1-Chl**, **3-Chl**, **5-Chl**, and **EDTA-3** were synthesized using Boc-D-Dab(Fmoc)-OH as the turn unit. The  $\gamma$ -DABA polyamides **2**, **4**, **6**, **2-Chl**, **4-Chl**, and **6-Chl** were synthesized using Fmoc-D-Dab(Boc)-OH as the turn unit. The glycine-linked polyamides **7** and **EDTA-7** were synthesized using Boc-Gly-OH as the turn unit.

Boc deprotections were conducted at room temperature for 25 min using 80% TFA/DCM. Carboxylic acids (3 eq) were activated with DIEA (5 eq) and PyBOP (2.5 eq) for 30–45 min at 37 °C prior to coupling. Couplings were allowed to proceed in NMP for 2 h at room temperature. Polyamides were cleaved from resin with either 3-(dimethylamino)-propylamine (Dp) or 3,3'-diamino-N-methyl-dipropylamine (Da) neat at 55 °C overnight. Products were purified by preparatory reverse-phase high performance liquid chromatography (prep HPLC) and characterized by analytical HPLC, UV-visible spectroscopy, and matrix-assisted laser desorption ionization-time-of-flight (MALDI-TOF) mass spectrometry.

Polyamide-chlorambucil conjugates were generated by activating the chlorambucil carboxylic acid (4 eq) with PyBOP (4 eq) and DIEA (12 eq) for 30 min at room temperature followed by addition of the polyamide. The reaction was allowed to proceed at room temperature for 1–2 h. Products were purified by reverse-phase prep HPLC, immediately lyophilized, and characterized as above.

Polyamide-EDTA conjugates were synthesized by slow addition of a 50% DMF/DIEA polyamide solution to a solution of EDTA dianhydride (20 eq) dissolved in a 1 DMSO:1 DMF: 2 DIEA mixture, vigorously shaking at 55 °C. The reaction was allowed to proceed an additional 30 min, followed by hydrolysis of the remaining EDTA dianhydride with 0.1 N NaOH for 10 min. The reaction was diluted and purified by reverse-phase prep HPLC. Subsequent Boc deprotection was accomplished by dissolving the dried EDTA conjugate in a minimum volume of neat TFA doped with triethylsilane. The reaction was allowed to proceed for 10–30 min at room temperature, then immediately diluted with water and purified by reverse-phase prep HPLC.

**1: ImImPyIm-(R)<sup>H2N</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1238.58; found: 1238.82.

**2: ImImPyIm-(R)<sup>H2N</sup> $\gamma$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1238.58; found: 1238.59.

**3: ImImPyPy-(R)<sup>H2N</sup> $\alpha$ -ImPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1238.58; found: 1238.90.

**4: ImImPyPy-(R)<sup>H2N</sup> $\gamma$ -ImPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>57</sub>H<sub>72</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1238.58; found: 1238.65.

**5: Im $\beta$ ImIm-(R)<sup>H2N</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>54</sub>H<sub>71</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1188.28; found: 1188.5.

**6: Im $\beta$ ImIm-(R)<sup>H2N</sup> $\gamma$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>54</sub>H<sub>71</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1188.28; found: 1188.86.

**7: ImImPyPy-Gly-ImPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>55</sub>H<sub>67</sub>N<sub>22</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1195.54; found: 1195.43.

**1-Chl: ImImPyIm-(R)<sup>Chl</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>71</sub>H<sub>89</sub>Cl<sub>2</sub>N<sub>24</sub>O<sub>11</sub><sup>+</sup> [M+H]<sup>+</sup>: 1523.65; found: 1523.66.

**2-Chl: ImImPyIm-(R)<sup>Chl</sup> $\gamma$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>71</sub>H<sub>89</sub>Cl<sub>2</sub>N<sub>24</sub>O<sub>11</sub><sup>+</sup> [M+H]<sup>+</sup>: 1523.65; found: 1523.70.

**3-Chl: ImImPyPy-(R)<sup>Chl</sup> $\alpha$ -ImPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>71</sub>H<sub>89</sub>Cl<sub>2</sub>N<sub>24</sub>O<sub>11</sub><sup>+</sup> [M+H]<sup>+</sup>: 1523.65; found: 1523.88.

**4-Chl: ImImPyPy-(R)<sup>Chl</sup> $\gamma$ -ImPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>71</sub>H<sub>89</sub>Cl<sub>2</sub>N<sub>24</sub>O<sub>11</sub><sup>+</sup> [M+H]<sup>+</sup>: 1523.65; found: 1523.57.

**5-Chl: Im $\beta$ ImIm-(R)<sup>Chl</sup> $\alpha$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>68</sub>H<sub>88</sub>Cl<sub>2</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1474.48, found: 1474.95.

**6-Chl: Im $\beta$ ImIm-(R)<sup>Chl</sup> $\gamma$ -PyPyPyPy- $\beta$ Dp** MALDI-TOF MS calcd for C<sub>68</sub>H<sub>88</sub>Cl<sub>2</sub>N<sub>23</sub>O<sub>10</sub><sup>+</sup> [M+H]<sup>+</sup>: 1474.48, found: 1474.26.

**EDTA-3: ImImPyPy-(R)<sup>H2N</sup> $\alpha$ -ImPyPyPy- $\beta$ Da-EDTA** MALDI-TOF MS calcd for C<sub>69</sub>H<sub>91</sub>N<sub>26</sub>O<sub>17</sub><sup>+</sup> [M+H]<sup>+</sup>: 1555.71; found: 1555.73.

**EDTA-7: ImImPyPy-Gly-ImPyPyPy- $\beta$ Da-EDTA** MALDI-TOF MS calcd for C<sub>67</sub>H<sub>86</sub>N<sub>25</sub>O<sub>17</sub><sup>+</sup> [M+H]<sup>+</sup>: 1512.66; found: 1512.54.

## 4.2 Construction of plasmids

Plasmids were constructed by annealing oligonucleotide pairs as follows: for pMFST – 5'-AGCTGTACAATCATTAGTGGTTACAATCATATGGTCATACAATC-ATTAGTCGTTACAATCATTAGCACACACAATCATC-3' and 5'-GATCGATGATTGTGTGTGCTAATGATTGTAACGACTAATGATTGTATGACCATATGATTGTAACCACTAATGATTGTAC-3'; for pMFST5 – 5'-AGCTGCGGCTATAATTAACAGGCCACAATTATACGATATATGCTAATGGTCAGACCATTACGATACATACTGGTATCTACTAATGGTCATACGGC-3' and 5'-GATCGCCGTATGACCATTAGTAGATTACCAGTATGTATCGTAATGGTCTGACCATTAGCATATATCGTATAATTGTGCCCTGTTAATTATAGCCGC-3'.

Annealed nucleotides were ligated into the *Bam*HI/*Hind*III restriction fragment of pUC19 using T4 DNA ligase, and the plasmid was transformed into *Escherichia coli* JM109 competent cells. Ampicillin-resistant white colonies were selected from 25 mL Luria-Bertani agar plates containing 50 mg/mL ampicillin treated with XGAL and isopropyl- $\beta$ -D-thiogalactopyranoside (IPTG) solutions and grown overnight at 37 °C. Cells were harvested the following day, and purification of the plasmids were performed with a Wizard Plus Midiprep DNA purification kit (Promega). DNA sequencing of the plasmid insert was performed by the sequence analysis facility at the California Institute of Technology.

## 4.3 Preparation of 5' <sup>32</sup>P-end-labeled DNA

The primer 5'-GAATTCGAGCTCGGTACCCGGG-3' was <sup>32</sup>P-labeled at the 5' end and used with the primer 3'-CAGCCCTTTGGACAGCACGGTC-5' to PCR amplify plasmids as described previously.<sup>28</sup>

## 4.4 DNase footprinting titrations

Polyamide equilibrations and DNase I footprinting titrations were conducted on the 5' end-labeled PCR product of the appropriate plasmid according to standard protocols.<sup>28</sup> DNA was incubated with polyamide or water (controls) for 12–16 h at room temperature prior to reaction. Determination of equilibrium association constants were performed as previously described.<sup>28</sup> Data points were fitted to a Langmuir binding isotherm by using the modified Hill equation with  $n = 1$ , except for the quantitation of glycine polyamide **7**, where  $n$  was optimized to the generated curve for the slipped site S and the overlapped site O.

#### 4.5 Thermal cleavage assays

Thermal cleavage assay experiments were conducted on the 5' end-labeled PCR products of the designated plasmids as described previously.<sup>11</sup> DNA was incubated with polyamide conjugates or water (control) for 24 h at 37 °C prior to work-up.

#### 4.6 Affinity cleavage assays

Affinity cleavage experiments were conducted on the 5' end-labeled PCR product of plasmid pMFST5 as previously described.<sup>28</sup> DNA was incubated with polyamide conjugates or water (control) for 13 h at room temperature prior to addition of ferrous ammonium sulfate.

### Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

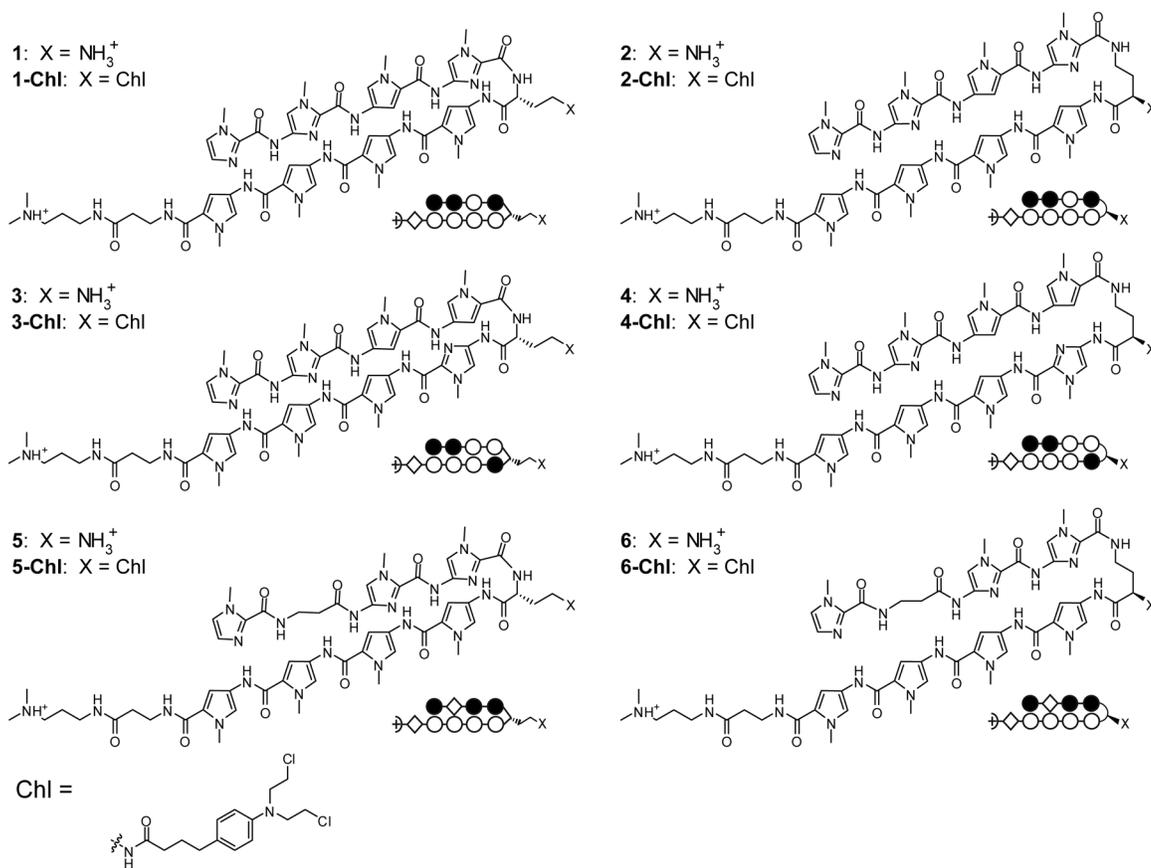
#### Acknowledgements

The authors thank the National Institutes of Health for research support and for a predoctoral NRSA training grant to M.E.F. Mass spectrometry analyses were performed in the Mass Spectrometry Laboratory of the Division of Chemistry and Chemical Engineering at the California Institute of Technology.

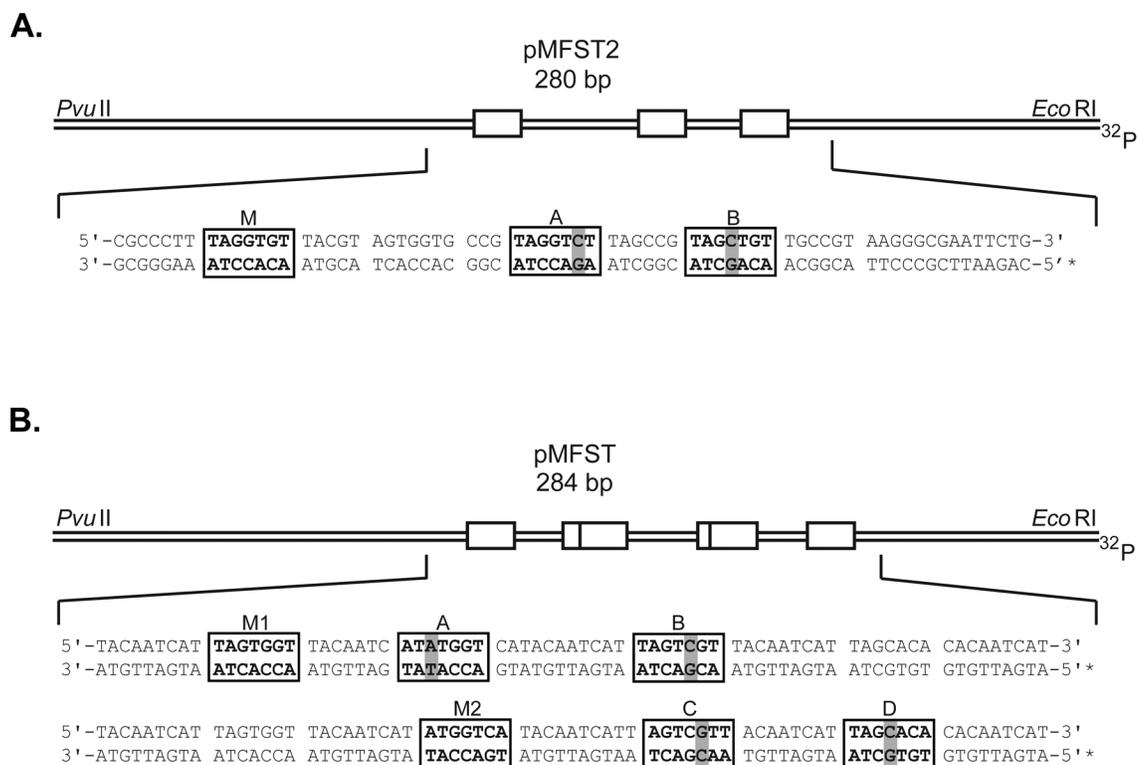
### References

1. Wemmer DE, Dervan PB. *Curr Opin Struct Biol* 1997;7:355–361. [PubMed: 9204277]
2. Dervan PB, Poulin-Kerstien AT, Fechter EJ, Edelson BS. *Top Curr Chem* 2005;253:1–31.
3. Best TP, Edelson BS, Nickols NG, Dervan PB. *Proc Natl Acad Sci U S A* 2003;100:12063–12068. [PubMed: 14519850]
4. Edelson BS, Best TP, Olenyuk B, Nickols NG, Doss RM, Foister S, Heckel A, Dervan PB. *Nucleic Acids Res* 2004;32:2802–2818. [PubMed: 15155849]
5. Olenyuk BZ, Zhang GJ, Klco JM, Nickols NG, Kaelin WG, Dervan PB. *Proc Natl Acad Sci U S A* 2004;101:16768–16773. [PubMed: 15556999]
6. Burnett R, Melander C, Puckett JW, Son LS, Wells RD, Dervan PB, Gottesfeld JM. *Proc Natl Acad Sci U S A* 2006;103:11497–11502. [PubMed: 16857735]
7. Kageyama Y, Sugiyama H, Ayame H, Iwai A, Fujii Y, Huang LE, Kizaka-Kondoh S, Hiraoka M, Kihara K. *Acta Oncol* 2006;45:317–324. [PubMed: 16644575]
8. Nickols NG, Jacobs CS, Farkas ME, Dervan PB. *Nucleic Acids Res* 2007;35:363–370. [PubMed: 17175539]
9. Nickols NG, Dervan PB. *Proc Natl Acad Sci U S A* 2007;104:10418–10423. [PubMed: 17566103]
10. Bando T, Sugiyama H. *Accounts Chem Res* 2006;39:935–944.
11. Wurtz NR, Dervan PB. *Chem Biol* 2000;7:153–161. [PubMed: 10712931]
12. Wang YD, Dziegielewska J, Wurtz NR, Dziegielewska B, Dervan PB, Beerman TA. *Nucleic Acids Res* 2003;31:1208–1215. [PubMed: 12582240]
13. Dickinson LA, Burnett R, Melander C, Edelson BS, Arora PS, Dervan PB, Gottesfeld JM. *Chem Biol* 2004;11:1583–1594. [PubMed: 15556009] *Chem. Biol.* 2006. 13, 339
14. Turner JM, Swalley SE, Baird EE, Dervan PB. *J Am Chem Soc* 1998;120:6219–6226.
15. Mrksich M, Wade WS, Dwyer TJ, Geierstanger BH, Wemmer DE, Dervan PB. *Proc Natl Acad Sci U S A* 1992;89:7586–7590. [PubMed: 1323845]
16. Mrksich M, Parks ME, Dervan PB. *J Am Chem Soc* 1994;116:7983–7988.
17. White S, Baird EE, Dervan PB. *J Am Chem Soc* 1997;119:8756–8765.
18. Herman DM, Baird EE, Dervan PB. *J Am Chem Soc* 1998;120:1382–1391.
19. Geierstanger BH, Mrksich M, Dervan PB, Wemmer DE. *Nat Struct Biol* 1996;3:321–324. [PubMed: 8599756]
20. Trauger JW, Baird EE, Mrksich M, Dervan PB. *J Am Chem Soc* 1996;118:6160–6166.

21. de Clairac RPL, Seel CJ, Geierstanger BH, Mrksich M, Baird EE, Dervan PB, Wemmer DE. *J Am Chem Soc* 1999;121:2956–2964.
22. Tsai SM, Farkas ME, Chou CJ, Gottesfeld JM, Dervan PB. *Nucleic Acids Res* 2007;35:307–316. [PubMed: 17170006]
23. Baird EE, Dervan PB. *J Am Chem Soc* 1996;118:6141–6146.
24. Hsu CF, Phillips JW, Trauger JW, Farkas ME, Belitsky JM, Heckel A, Olenyuk B, Puckett JW, Wang CC, Dervan PB. *Tetrahedron* 2007;63:6146–6151.
25. Woods CR, Ishii T, Wu B, Bair KW, Boger DL. *J Am Chem Soc* 2002;124:2148–2152. [PubMed: 11878968]
26. White S, Baird EE, Dervan PB. *Chem Biol* 1997;4:569–578. [PubMed: 9281524]
27. Urbach AR, Dervan PB. *Proc Natl Acad Sci U S A* 2001;98:4343–4348. [PubMed: 11296283]
28. Trauger JW, Dervan PB. *Meth Enzymol* 2001;340:450–466. [PubMed: 11494863]
29. Alvarez D, Chou CJ, Latella L, Zeitlin SG, Ku S, Puri PL, Dervan PB, Gottesfeld JM. *Cell Cycle* 2006;5:1537–1548. [PubMed: 16861886]

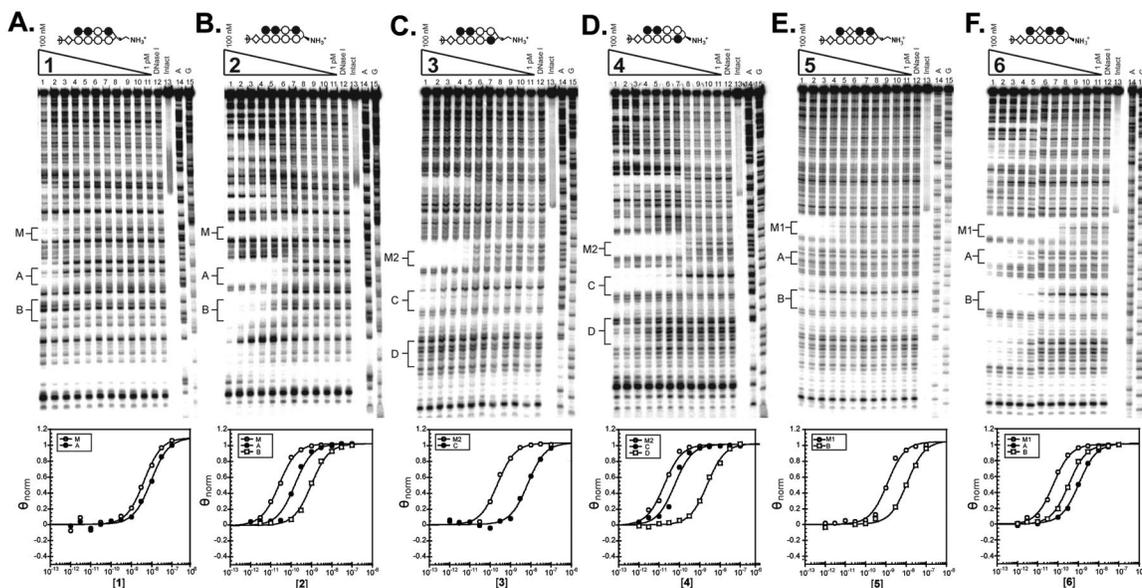


**Figure 1.** Chemical and ball-and-stick structures of polyamides **1-6** and their **Chl** conjugates. (Left)  $\alpha$ -diaminobutyric acid-linked hairpin polyamides. (Right)  $\gamma$ -diaminobutyric acid-linked hairpin polyamides. The ball and stick symbols are defined as follows: an open circle denotes a pyrrole ring, a filled circle denotes an imidazole, and a diamond shape denotes  $\beta$ -alanine.



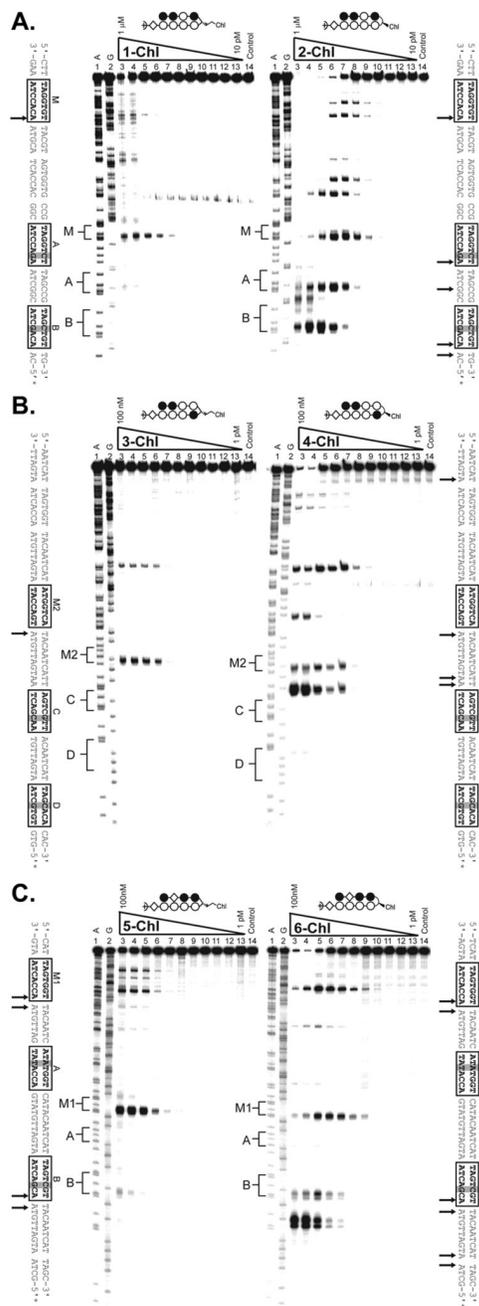
**Figure 2.**

Illustration of the *EcoRI/PvuII* restriction fragments derived from plasmids (A) pMFST2 and (B) pMFST. Polyamides **1** and **2** were characterized on pMFST2. Polyamides **3-6** were characterized on pMFST. The designed polyamide binding sites are indicated by boxes. Single base pair mismatches are indicated by shaded regions. For pMFST, the designed insert is shown twice with binding sites for (B, bottom) polyamides **3** and **4** and (B, top) polyamides **5** and **6**.



**Figure 3.**

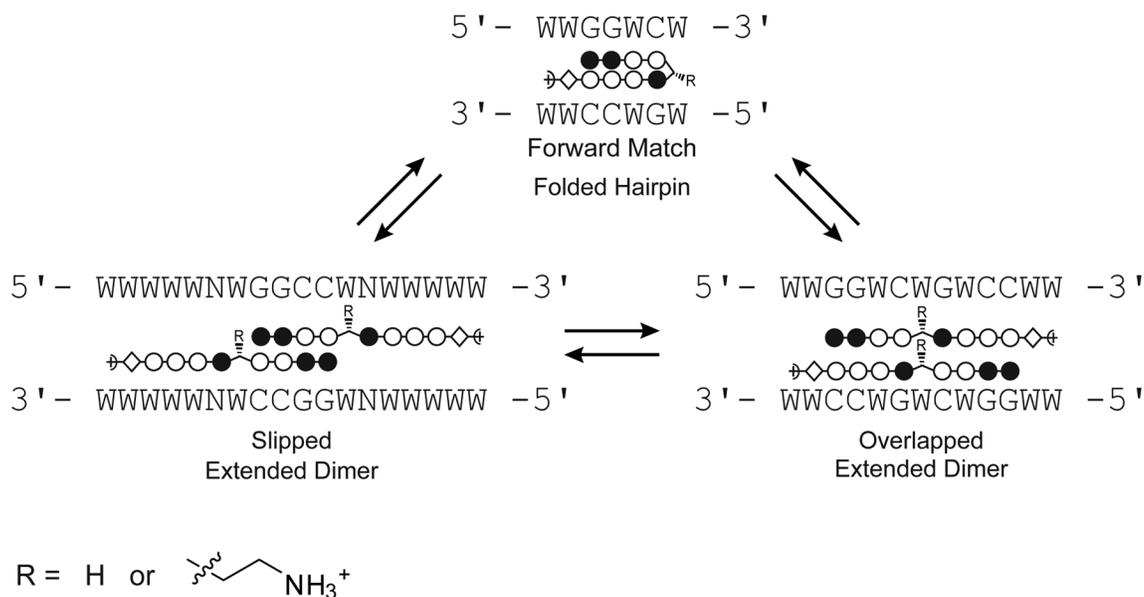
Quantitative DNase I footprinting titration experiments for polyamides (A) **1** and (B) **2** on the 280 base pair, 5' end-labeled PCR product of plasmid pMFST2; and for polyamides (C) **3**, (D) **4**, (E) **5**, and (F) **6** on the 284 base pair, 5' end-labeled PCR product of plasmid pMFST: lanes 1–11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. (Below) Each footprinting gel is accompanied by its respective binding isotherms.



**Figure 4.**

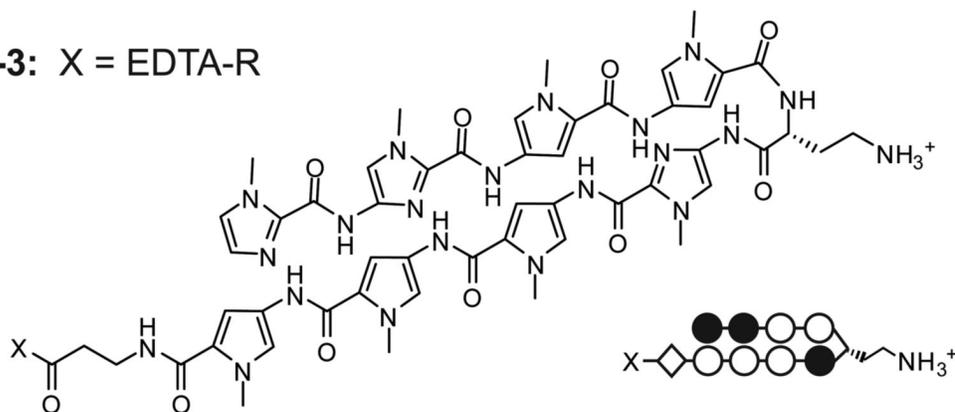
(A) Thermal cleavage assay experiments with polyamide-chlorambucil conjugates (left) **1-Chl** and (right) **2-Chl** on the on the 280 base pair, 5' end-labeled PCR product of plasmid pMFST2: lane 1, A reaction; lane 2, G reaction; lanes 3–13, 1 μM, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, and 10 pM polyamide, respectively; lane 14, intact DNA. (C-D) Thermal cleavage assay experiments with polyamide-chlorambucil conjugates (C, left) **3-Chl** and (C, right) **4-Chl** and (D, left) **5-Chl** and (D, right) **6-Chl** on the on the 284 base pair, 5' end-labeled PCR product of plasmid pMFST: lane 1, A reaction; lane 2, G reaction; lanes 3–13, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 14, intact DNA. Putative major sites of

alkylation on the DNA fragments are indicated by arrows on the sequences adjacent to each gel.



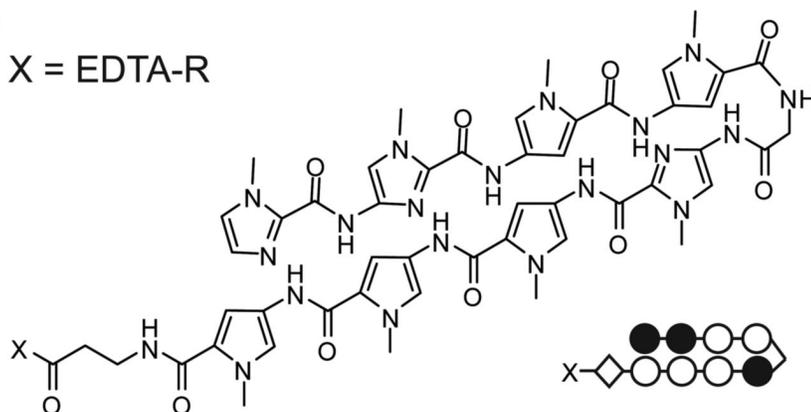
**Figure 5.** Illustration of the binding modes for polyamides **3** and **7**. W signifies A or T; N represents any nucleotide. Sequence specificity for the slipped binding mode has been presumed based on the polyamide literature.

**EDTA-3:** X = EDTA-R

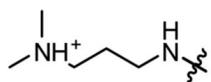


**7:** X = Dp

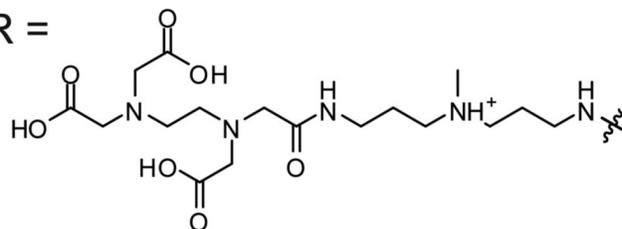
**EDTA-7:** X = EDTA-R



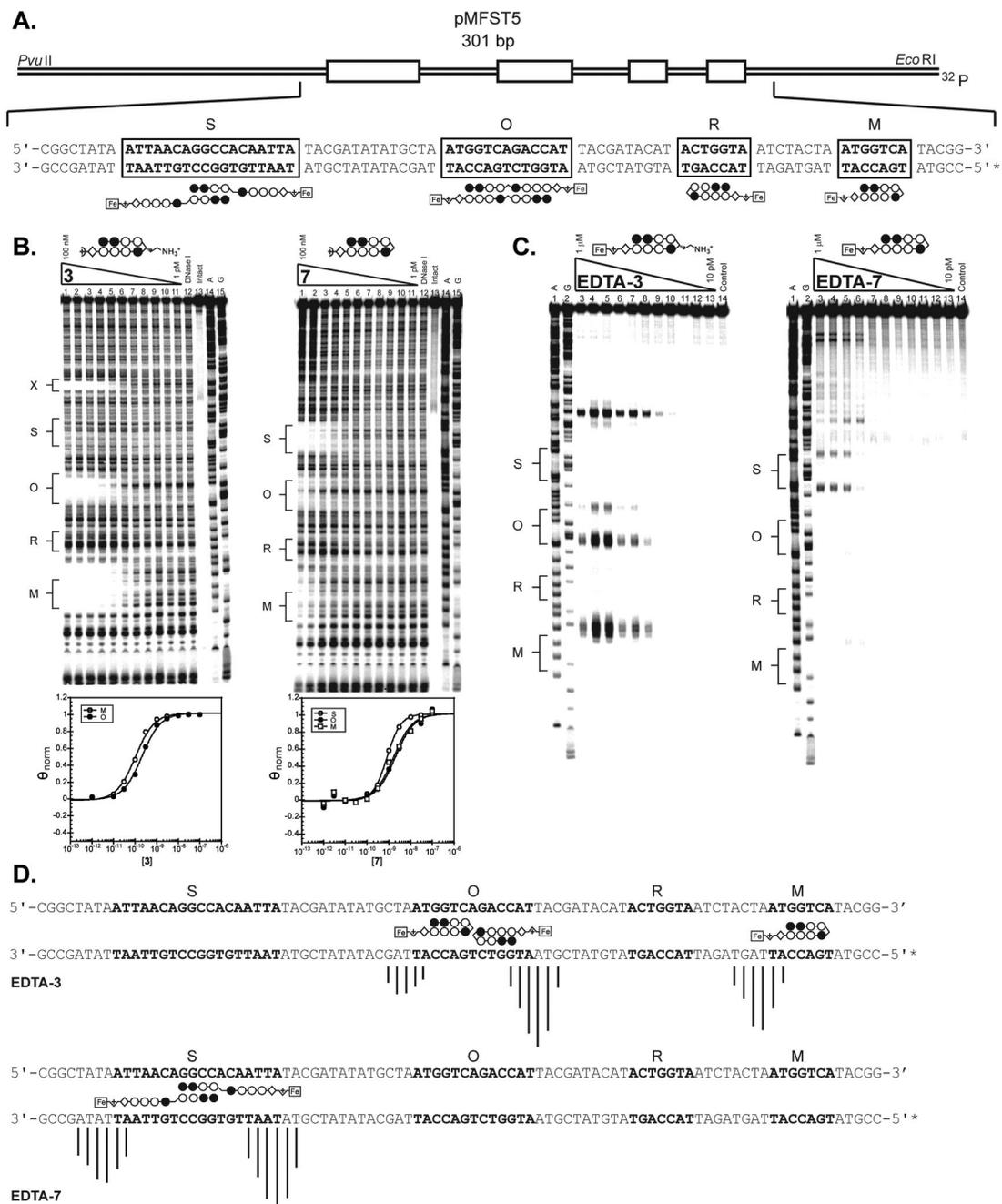
Dp =



EDTA-R =



**Figure 6.** Chemical and ball-and-stick structures of **EDTA-3**, and of the glycine-linked polyamide **7** and its EDTA conjugate **EDTA-7**.

**Figure 7.**

(A) Illustration of the *EcoRI/PvuII* restriction fragment derived from plasmid pMFST5. The designed polyamide binding sites are indicated by boxes. Binding models are illustrated below the DNA sequence by ball-and-stick structures of **EDTA-7**. (B) Quantitative DNase I footprinting titration experiments for polyamides (left) **3** and (right) **7** on the 301 base pair, 5' end-labeled PCR product of plasmid pMFST5: lanes 1–11, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide, respectively; lane 12, DNase I standard; lane 13, intact DNA; lane 14, A reaction; lane 15, G reaction. Site X indicates an undesigned match site inherent in the plasmid sequence. (Below) Each footprinting gel is accompanied by its respective binding isotherms. (C) Affinity cleavage experiments with

polyamide-EDTA conjugates (left) **EDTA-3** and (right) **EDTA-7** on the 301 base pair, 5' end-labeled PCR product of plasmid pMFST5: lane 1, A reaction; lane 2, G reaction; lanes 3–13, 1  $\mu$ M, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, and 10 pM polyamide, respectively; lane 14, intact DNA. (D) Summary of affinity cleavage patterns for (top) **EDTA-3** and (bottom) **EDTA-7**. Bar heights are proportional to the relative cleavage intensities at each base pair, normalized to each gel. Ball-and-stick structures illustrating their respective proposed modes of binding are shown using **EDTA-7** for clarity. Designed binding sites are indicated in bold in the DNA sequence.

Binding affinities ( $M^{-1}$ ) for parent polyamides on plasmid pMFST2 (1–2) and plasmid pMFST (3–6). Equilibrium association constants reported are mean values from three DNase I footprinting titration experiments. Standard deviations are shown in parentheses. Bolded numbers in brackets indicate fold decrease in match site binding affinity for  $\alpha$ -DABA-linked versus  $\gamma$ -DABA-linked polyamides. Mismatch base pairs are indicated in bold.

Table 1

Polyamide	Match Site	$K_a$ ( $M^{-1}$ )	Mismatch Site	$K_a$ ( $M^{-1}$ )	Mismatch Site	$K_a$ ( $M^{-1}$ )
1	5'-TAGGTGT-3'	3.0( $\pm$ 1.3) $\times 10^8$	5'-TAGGTCT-3'	8.5( $\pm$ 3.1) $\times 10^7$	5'-TAGCTGT-3'	—
2	5'-TAGGTGT-3'	4.4( $\pm$ 0.4) $\times 10^{10}$	5'-TAGGTCT-3'	1.1 ( $\pm$ 0.8) $\times 10^{10}$	5'-TAGCTGT-3'	1.1 ( $\pm$ 0.3) $\times 10^9$
3	5'-ATGGTCA-3'	4.4( $\pm$ 0.6) $\times 10^9$	5'-AACGACT-3'	1.5( $\pm$ 0.3) $\times 10^8$	5'-TAGCACA-3'	—
4	5'-ATGGTCA-3'	4.9( $\pm$ 0.7) $\times 10^{10}$	5'-AACGACT-3'	1.8( $\pm$ 0.3) $\times 10^{10}$	5'-TAGCACA-3'	3.9( $\pm$ 1.0) $\times 10^8$
5	5'-TAGTGGT-3'	8.0( $\pm$ 0.6) $\times 10^8$	5'-ATATGGT-3'	—	5'-TAGTCGT-3'	6.2( $\pm$ 3.0) $\times 10^7$
6	5'-TAGTGGT-3'	4.3( $\pm$ 0.3) $\times 10^{10}$	5'-ATATGGT-3'	9.9( $\pm$ 0.3) $\times 10^8$	5'-TAGTCGT-3'	6.2( $\pm$ 0.2) $\times 10^9$

**Table 2**

Binding affinities ( $M^{-1}$ ) for parent polyamides **3** and **7** on plasmid pMFST5. Equilibrium association constants reported are mean values from three DNase I footprinting titration experiments. Standard deviations are shown in parentheses.

Polyamide	M	O	S
3	$9.8(\pm 1.3) \times 10^9$	$4.7(\pm 1.3) \times 10^9$	--
7	$7.0(\pm 1.0) \times 10^8$	$5.9(\pm 2.1) \times 10^8$	$1.3(\pm 0.2) \times 10^9$