### Chapter 2

# Linear β-linked Polyamides Target GAA Repeats and Alleviate Transcription Repression in Friedreich's Ataxia Cell Culture

The text of this chapter was taken in part from a manuscript co-authored with Ryan Burnett, Christian Melander, Leslie S. Son, Professor Robert D. Wells (Texas A & M Health Science Center), Professor Joel M. Gottesfeld (The Scripps Research Institute), and Professor Peter B. Dervan (California Institute of Technology).

(Burnett, R.; Melander, C.; Puckett, J. W.; Son, L. S.; Wells, R. D.; Dervan, P. B.; Gottesfeld, J. M. "DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA·TTC repeats in Friedreich's ataxia." *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 11497–11502.)

### **Chapter 2A**

# DNA sequence-specific polyamides alleviate transcription inhibition associated with long GAA·TTC repeats in Friedreich's Ataxia

### Abstract

The DNA abnormality found in 98% of Friedreich's ataxia (FRDA) patients is the unstable hyperexpansion of a GAA·TTC triplet repeat in the first intron of the frataxin gene. Expanded GAA·TTC repeats result in decreased transcription and reduced levels of frataxin protein in affected individuals.  $\beta$ -Alanine-linked pyrrole–imidazole polyamides bind GAA·TTC tracts with high affinity and disrupt the intramolecular DNA·DNA-associated region of the sticky-DNA conformation formed by long GAA·TTC repeats. Fluorescent polyamide-BODIPY conjugates localize in the nucleus of a lymphoid cell line derived from a FRDA patient. The synthetic ligands increase transcription of the frataxin gene in cell culture, resulting in increased levels of frataxin protein. DNA microarray analyses indicate that a limited number of genes are significantly affected in FRDA cells. Polyamides may increase transcription by altering the DNA conformation of genes harboring long GAA·TTC repeats or by chromatin opening.

### Introduction

The neurodegenerative disease Friedreich's ataxia (FRDA) is caused by hyperexpansion of GAA·TTC repeats in the first intron of a nuclear gene that encodes the essential mitochondrial protein frataxin.<sup>1-4</sup> Normal frataxin alleles have 6–34 repeats, whereas FRDA patient alleles have 66–1,700 repeats. Intronic GAA·TTC repeats interfere with gene transcription.<sup>5-7</sup> Longer repeats cause a more profound frataxin deficiency and are associated with earlier onset and increased severity of the disease.<sup>4</sup> Biochemical studies have documented that expanded GAA·TTC repeats adopt unusual non-B DNA structures, such as triplexes, containing two purine GAA strands along with one pyrimidine TTC strand, flanking a single-stranded pyrimidine region<sup>5,8</sup> as well as intramolecular "sticky" DNA.<sup>6,9–12</sup> Long (GAA·TTC), repeat sequences form sticky DNA with two separate long (GAA·TTC), repeating tracts associated within a single closed plasmid DNA. The interaction of the two tracts requires the repeats oriented in the direct repeat orientation, negative supercoiling and the presence of divalent metal ions to stabilize the DNA·DNA-associated region.<sup>8-10</sup> We have demonstrated the capacity of sticky DNA to form both in vitro<sup>10,13</sup> and in vivo.<sup>11-13</sup> Triplexes and/or sticky DNA block elongation by RNA polymerase II.<sup>5</sup> Another study using artificial transgenes has shown that expanded GAA·TTC repeats induce repressive heterochromatin *in vivo* in a manner reminiscent of position-effect variegation.<sup>14</sup> This effect was augmented by coexpression of heterochromatin protein 1. Here, we address whether minor-groove DNA-binding small molecules can alleviate transcription repression of the frataxin gene.

Molecules that reverse triplex/sticky DNA and/or heterochromatin formation in the frataxin gene may increase successful elongation through expanded GAA·TTC repeats, thereby relieving the deficiency in frataxin mRNA and protein in FRDA cells.<sup>11,14,15</sup> Pyrrole–imidazole polyamides are cell-permeable small molecules that can be programmed to recognize a broad repertoire of DNA sequences.<sup>16</sup> Two classes of polyamides are well established: hairpin polyamides bind mixed-sequence DNA with high affinity and

specificity,<sup>16,17</sup> and linear  $\beta$ -alanine-linked polyamides are available for targeting purine tracts of DNA, such as GAGAA·TTCTC repeats.<sup>18,19</sup>  $\beta$ -Alanine-linked polyamides have been shown to bind GAGAA repeats in *Drosophila* satellite DNA both *in vitro* and in cytological chromosome spreads.<sup>19</sup> These molecules induce chromatin opening and reverse heterochromatin-mediated position-effect gene silencing when administered to *Drosophila* embryos.<sup>19,20</sup>

Structural studies indicate that  $\beta$ -alanine-linked polyamides bind the minor groove of canonical B DNA.<sup>21</sup> Given the high affinity of  $\beta$ -alanine-linked polyamides for purine tracts,<sup>18</sup> these molecules might act as a thermodynamic "sink" and lock GAA·TTC repeats into double-stranded B DNA. Such an event would disfavor duplex unpairing, which is necessary for formation of FRDA triplexes and sticky DNA. Although single polyamides bound within coding regions of genes do not appear to block transcription elongation,<sup>22–24</sup> we cannot be certain that multiple polyamides have the potential to relieve transcription repression at expanded GAA·TTC repeats. Alternatively, polyamides may relieve heterochromatin-mediated repression by opening the chromatin domain containing the frataxin gene.<sup>20</sup>

### **Results and Discussion**

### **Targeting GAA·TTC Repeat DNA with High-Affinity Ligands.**

We synthesized two  $\beta$ -alanine-linked polyamides of different length, **1** (Im-Py- $\beta$ -Im-Py- $\beta$ -Im- $\beta$ -Dp, where Py is pyrrole, Im is imidazole,  $\beta$  is  $\beta$ -alanine, and Dp is dimethylaminopropylamine) to target the 9-bp site 5'-AAGAAGAAG-3' and **3** (Im-Py- $\beta$ -Im-Py- $\beta$ -Im-Py- $\beta$ -Im- $\beta$ -Dp) to target the 12-bp site 5'-AAGAAGAAGAAG-3' (Figure 2.1A). Quantitative DNase I footprinting<sup>25</sup> demonstrates that **1** binds to a radiolabeled PCR product containing a (GAA·TTC)<sub>6</sub> sequence with an apparent dissociation constant ( $K_d$ ) of 0.1 nM (Figure 2.2A and Table 2.1). Polyamide **3** exhibits a  $K_d$  of  $\approx$  3 pM in footprinting experiments performed at low DNA concentrations (Figure 2.2C, and Table



*Figure 2.1.* Polyamide structures, binding models and nuclear localization in cell culture. a) Structures of polyamides 1, 2, 3, and 4 (R = methyl) and their BODIPY conjugates 1-B, 2-B, and 3-B (R = *N*-propylbutanamide linked BODIPY FL C<sub>5</sub>). Polyamide structures are represented schematically as binding models. Filled and open circles are Im and Py rings, respectively; diamonds are  $\beta$ -alanine; and, the semicircle with a plus sign is dimethylaminopropylamine. Linear polyamides bind in a carboxyl- to amino-terminal orientation with respect to the 5' to 3' sequence of their DNA target site.<sup>18</sup> Mismatches formed with polyamides 2 and 4 are indicated with shaded boxes. b)-c) Deconvolution microscopy of unfixed lymphoid cells (GM15850, derived from an FRDA patient, shown in b), and GM15851 derived from a healthy sibling, shown in c)), incubated with BODIPY conjugates of each of the indicated polyamides, at 2  $\mu$ M concentration in cell culture medium for 16 h prior to visualization, as described.<sup>35</sup> The bars represent 10 microns.

2.1); however, this value may be an underestimation of the affinity of this molecule for GAA·TTC repeat DNA, because our  $K_d$  measurements are limited by a minimum DNA concentration of  $\approx 2$  pM in the binding reaction. Two mismatch controls, **2** (Im-Py- $\beta$ -Im-Im- $\beta$ -Py- $\beta$ -Dp) (Figure 2.1A) and **4** (Im-Py- $\beta$ -Im-Py- $\beta$ -Im-Im- $\beta$ -Py- $\beta$ -Dp), have binding affinities for the (GAA·TTC)<sub>6</sub> target sequence that are three orders of magnitude lower than those of the match polyamides (Figures 2B and 2D and Table 2.1). Polyamide **1** is



*Figure 2.2.* DNA binding properties of the polyamides. Quantitative DNase I footprint analysis for polyamide binding to a radiolabeled PCR product containing a  $(GAA \cdot TTC)_6$  repeat sequence, labeled on the purine strand. DNA (at  $\approx 20 \text{ pM}$  for b), c) and e) and  $\approx 2 \text{ pM}$  for d)) and polyamide were allowed to equilibrate for 16 h, with the indicated ranges of polyamide concentrations, prior to DNase digestion and gel analysis.<sup>25</sup> The phosphorimage of each gel is shown, with undigested DNA in the lane marked "-"; a G+A sequencing reaction of the same DNA is shown along with DNase-treated DNA in the absence of polyamide (in the lane marked "0"). a) An excerpt of the DNA sequence cloned in pCR2.1 DNA, used to generate the PCR product for footprinting reactions is referenced for b) – e). b) Polyamide 1 targeting 5'-AAGAAGAAG-3'; c) Polyamide 2, mismatch control for 1; d) Polyamide 3 targeting 5'-AAGAAGAAGAAG-3'; e) Polyamide 4, mismatch control for 3

Polvamide sequence	(GAA·TTC) <sub>n</sub> repeat no. in target site. n	Binding affinity (Kd. nM)*
<b>1</b> : Im-Pv-β-Im-Pv-β-Im-β-Dp	3	0.11 ± 0.02
<b>2</b> : Im-Py- <i>β</i> -Im- <u>Im</u> - <i>β</i> - <u>Py</u> - <i>β</i> -Dp	3	>100
<b>3</b> : Im-Py- <i>β</i> -Im-Py- <i>β</i> -Im-Py- <i>β</i> -Im- <i>β</i> -Dp	4	$0.003 \pm 0.001$
<b>4</b> : Im-Py- <i>β</i> -Im-Py- <i>β</i> -Im- <u>Im</u> - <i>β</i> - <u>Py</u> -β-Dp	4	$2.0 \pm 0.4$

Table 2.1. Polyamides designed to target GAA·TTC repeats in the frataxin gene

Mismatch amino acids are underlined. Im, imidazole; Py, pyrrole,  $\beta$ ,  $\beta$ -alanine; Dp, dimethylaminopropylamine

\*Binding affinities (mean values of the  $K_d$  from a minimum of two determinations, and standard deviations) determined by quantitative DNase I footprinting, as in Figure 2.2



also able to bind extended regions of GAA·TTC repeats  $[(GAA·TTC)_{33}]$  with no loss in affinity, with several molecules of 1 bound per DNA molecule (Figure 2.3). As another test for sequence specificity, footprinting experiments with 1 and 3 and a radiolabeled DNA fragment containing a mismatch DNA sequence (5'-...AGGAGGAGGAGGAGGAGGAGGA. ...-3') were performed. Neither polyamide 1 nor polyamide 3 bound this DNA sequence at polyamide concentrations up to 100 nM (Figure 2.4). Polyamide 1 binds other polypurine DNA tracts with single G-to-A transitions, maintaining similar or slightly reduced binding affinities compared with the 5'-AAGAAGAAG-3' site. A sequence harboring two transitions (A-to-G and G-to-A), giving rise to 5'-AAAAGGAAG-3', increases the  $K_d$  of 1 by more than three orders of magnitude. Truncation of 1 to Im-Py- $\beta$ -Im- $\beta$ -Dp increases its  $K_{d}$  on 5'-AAGAAGAAG-3' by more than four orders of magnitude (see Chapter 2B for data and discussion).

# Nuclear Localization of Fluorescent Polyamides.

Fluorescent versions of the match polyamides 1 and 3 (1-B and 3-B, respectively) and mismatch polyamide 2 (2-B) were synthesized, where the dye BODIPY FL (C5) was attached at the carboxyl terminus of the polyamide (Figure 2.1A). Quantitative DNase I footprinting demonstrated that polyamides 1- and 3-B exhibit 13- to 20-fold losses in binding



*Figure 2.4.* Quantitative DNase I footprint analysis for polyamide **1-B**, the BODIPY conjugate of **1**, binding to a radiolabeled (GAA·TTC)<sub>6</sub> PCR product. Lanes are labeled as per Figures 2.2 and 2.3. Quantitation of this experiment by PhosphorImage analysis reveals a  $K_d$  of 1.3 nM. The amplicon sequence is depicted in Figure 2.2A.

affinity for  $(GAA \cdot TTC)_6$  DNA, compared with the parent polyamides (for **1-B**,  $K_d = 1.3$  nM) (Figure 2.5; for **3-B**,  $K_d = 0.04$  nM). Epstein–Barr virus-transformed lymphoblast cell lines from an FRDA patient (line GM15850) and from an unaffected sibling (line GM15851) were obtained from the NIGMS Human Genetic Cell Repository (Coriell Institute, Camden, NJ). Both the match **1-B** and mismatch **2-B** conjugates localize in the nucleus of live, unfixed normal and FRDA lymphoid cells after 16 h incubation in culture medium, as determined by deconvolution microscopy (Figure 2.1B). The BODIPY-conjugate of the longer polyamide **3**, **3-B**, also localizes in the nucleus of the FRDA cells.

### GAA·TTC-Specific Polyamides Increase Frataxin mRNA and Protein.

To assess whether polyamides alleviate transcription inhibition caused by expanded GAA·TTC repeats in the frataxin gene, we used real-time quantitative qRT-PCR to monitor



Figure 2.5. Polyamide 1 increases the levels of frataxin mRNA and protein in an FRDA lymphoid cell line. a) FRDA and control cells with Measurement of frataxin mRNA levels, relative to that of GAPDH, in cell lines derived from an unaffected individual (GM15851), and an various concentrations of each FRDA patient (GM15850) by qRT-PCR.<sup>35</sup> Polyamides 1 and 2 were included in the cell culture medium at the indicated concentrations, of the polyamides for various and frataxin and GAPDH mRNA were determined after 7 days, with media and polyamide replenished on days 3 and 5. Error bars are de- lengths of time and found that rived from the percent error of the average and standard deviation of the change in cycle threshold between frataxin and GAPDH for trip- only polyamide 1 increased licate experiments, with triplicate qRT-PCR determinations for each experiment. b) Effects of polyamides on frataxin protein in cultured frataxin mRNA levels after lymphoid cells. FRDA or control cells were incubated as in panel a) prior to western blot analysis with antibody to human frataxin or actin. 7 days incubation in culture Equivalent amounts of total cell extract protein were loaded in each lane.

frataxin mRNA levels in the lymphoid cell lines described above, with the levels of GAPDH mRNA as an internal control for each RNA sample. No differences in GAPDH mRNA levels were found between the two cell lines. As expected, the FRDA cell line had a markedly lower level of frataxin mRNA compared with the cell line from the normal individual (Figure 2.6A; 13  $\pm$  6%, for 20 determinations, Table 2.2). We incubated the medium. No changes in frataxin

mRNA levels were observed on shorter incubation times. Over the concentrations of 1–8  $\mu$ M, we find that polyamides are not cytotoxic to the lymphoid cell lines (as determined by trypan blue exclusion and measurements of ATP levels) and do not affect cell growth rates. Importantly, the level of frataxin mRNA in the FRDA GM15850 cell line was increased

2.5-fold by incubation with polyamide 1 (at 2  $\mu$ M, Figure 2.6A). The average fold increase observed with 2  $\mu$ M 1 in the FRDA cell line is 2.2 ± 0.6 (in 20 experiments), resulting in an average of ≈27% of the level of frataxin mRNA found in the normal cell line (Table 2.2). We note that polyamide concentrations greater than the  $K_d$  for *in vitro* binding are required for increasing frataxin transcription, presumably because of the large number of potential polyamide-binding sites in the human genome and availability of these sites in the cell nucleus. Despite the fact that 1 binds the 9-bp 5'-AAGAAGAAG-3' repeat with high affinity, the sequence landscape of all possible high-affinity sites for this class of  $\beta$ -linked oligomers has not yet been fully characterized.<sup>26</sup> Neither higher concentrations of 1 nor longer incubation times increased frataxin transcription above the levels observed at 2  $\mu$ M on 7 day incubations. Polyamide 1 did not increase frataxin mRNA levels in the cell line derived from the normal individual. Similar incubations with the mismatch polyamide **2** were without significant effect in either cell line. The levels of GAPDH mRNA were not changed by polyamide treatment in either cell line.

	GM15851	GM15850	GM15850+1	Ave	Rel. exp.	Rel. exp.	Fold change
Exp.	Ave ∆Ct	Ave ∆Ct	Ave ∆Ct	∆∆Ct	850:851	850+1:851	850+1:850
1	8.088	11.234	10.628	3.146	11.30	17.19	1.52
2	7.584	11.088	9.923	3.505	8.81	19.76	2.24
3	7.484	10.850	10.057	3.366	9.70	16.81	1.73
4	7.363	11.230	10.785	3.868	6.85	9.33	1.36
5	8.566	10.765	10.083	2.200	21.77	34.94	1.61
6	7.683	10.649	9.206	2.966	12.80	34.80	2.72
7	8.523	12.130	10.748	3.607	8.21	21.39	2.61
8	7.973	11.468	10.959	3.496	8.87	12.62	1.42
9	9.176	12.885	11.241	3.710	7.64	23.89	3.13
10	7.823	11.244	9.837	3.421	9.34	24.75	2.65
11	8.387	11.976	10.607	3.589	8.31	21.46	2.58
12	8.636	11.636	10.551	3.000	12.50	26.52	2.12
13	7.855	11.638	9.812	3.783	7.26	25.74	3.54
14	8.608	12.082	10.804	3.474	9.00	21.82	2.42
15	8.150	10.601	9.653	2.451	18.29	35.29	1.93
16	8.607	11.589	10.402	2.982	12.66	28.83	2.28
17	8.540	11.165	10.316	2.625	16.21	29.19	1.80
18	8.271	9.959	9.425	1.688	31.04	44.93	1.45
19	7.739	10.136	8.515	2.397	18.99	58.43	3.08
20	8.586	10.763	10.024	2.177	22.12	36.92	1.67
Averages	8.182	11.254	10.179	3.072	13.08	27.23	2.19
SD (±)	0.482	0.707	0.662	0.623	6.42	11.44	0.64

*Table 2.2.* Collected data showing the average  $\Delta$ Ct for normal cells (GM15851) and FRDA cells (GM15850) before and after treatment with 1 at 2  $\mu$ M

Exp., experiment; Ave, average; Rel. exp., relative expression; 850, GM15850 cell line; 851, GM15851 cell line

We next examined the effect of removal of polyamide 1 from the culture medium on frataxin transcription. After induction of frataxin mRNA synthesis by 1 (7 days at 2  $\mu$ M), transfer of the cells to fresh medium lacking polyamide caused frataxin mRNA levels to decrease to their original levels after 96 h (data not shown). Thus, polyamides must be continuously present to maintain active transcription of FRDA frataxin alleles. The finding that incubation periods of 7 days or more are necessary to observe increases in frataxin mRNA suggests that multiple rounds of DNA replication are necessary for the compound to alter either the DNA or chromatin structure of expanded frataxin alleles, leading to active transcription, and removal of the polyamide causes the frataxin gene to readopt its inactive DNA or chromatin conformation. To test whether cellular proliferation is required for frataxin gene activation, we serum-starved FRDA and control cells, leading to cell cycle arrest (confirmed by fluorescence activated cell sorting, data not shown), and then incubated the arrested cells with polyamide 1 for 7 days before qRT-PCR analysis. We find no increase in frataxin mRNA with polyamide 1 under these conditions, suggesting that cell division is a requirement for upregulation of transcription by the polyamide.

It was curious that the highest-affinity compound, **3**, did not increase frataxin mRNA levels, because the fluorescent version of this molecule, **3-B**, localized in the nucleus of FRDA lymphoid cells. Previous studies have established that nuclear localization is sensitive to polyamide composition and structure and, especially, the nature of the carboxyl terminus;<sup>27,28</sup> therefore, the nonfluorescent version of **3** may not enter the nucleus. To test this hypothesis, we monitored the levels of frataxin mRNA after incubation with **3-B**, and found an approximately two to threefold increase in relative levels of frataxin mRNA (compared with GAPDH) after 2- to 4-day incubations (data not shown). Thus, polyamide **3** may not have the optimum chemical properties for nuclear localization or DNA binding in the context of cellular chromatin. In contrast, polyamide **1-B** did not increase frataxin mRNA levels in experiments where positive effects were found with polyamide **1**.

Because the primary transcripts from FRDA frataxin genes contain long stretches of

GAA repeat RNA sequence, it is conceivable that this RNA will not be correctly processed into mature frataxin mRNA, and frataxin protein will not be produced. To test whether polyamide **1** leads to increased levels of frataxin protein in treated lymphoid cells, totalcell extracts from polyamide-treated (1–2  $\mu$ M for 7 days), and untreated GM15851 control and GM15850 FRDA cells were subjected to SDS/PAGE and the corresponding blots probed with anti-frataxin or anti-actin antibodies (Figure 2.6B). An approximately two to threefold increase in frataxin protein is observed with **1** in the FRDA cells, which correlates well with the observed increase in frataxin mRNA (Figure 2.6A).

### Effects of Polyamides on Global Gene Expression.

DNA microarray analyses were performed with RNA isolated from GM15850 FRDA and GM15851 normal lymphoid cells that were either untreated or treated with polyamides 1 (at 1 and 2  $\mu$ M) or 2 (at 2  $\mu$ M) for 7 days on Affymetrix Human Genome U133 Plus 2.0 GeneChips. Polyamide 1 was found to affect the mRNA levels for a limited number of genes in the FRDA cell line (at  $P \le 0.005$ , 51 genes affected by 1  $\mu$ M 1, 16 genes affected by  $2 \mu M 1$ ) and only 2 genes in the normal cell line (Table 2.3). Although more genes were called affected by 1 at 1  $\mu$ M than at 2  $\mu$ M, this difference is largely due to genes whose mRNA levels change by  $\leq 25\%$  in either direction. At 2  $\mu$ M 1, 15 genes were increased in expression by >50%, and 1 gene was decreased by 45%. At 1  $\mu$ M 1, only 3 genes had comparable changes in their mRNA levels. For GM15851 cells, 2 genes were up-regulated by 1, and no genes were down-regulated. For the frataxin gene, untreated GM15850 cells showed 17% of the frataxin mRNA found in untreated GM15851 cells, and incubation with 1 at 2 µM increases frataxin mRNA by 2.5-fold, bringing the frataxin mRNA level in GM15850 cells to 42% of that found in GM15851 cells. These values are comparable with those obtained by qRT-PCR (Figure 2.6A). Transcript levels for frataxin were not changed by 2 in either cell line, and 2 affected only a small number of genes in either cell line (at  $P \le 0.005$ , 3 genes affected by 2  $\mu$ M 2 in GM15850 cells, and 1 gene

*Table 2.3.* Potential match and degenerate binding sites for polyamide 1 in significant genes listed for class comparisons where  $P \le 0.005$ . GeneChip data were normalized using RMAExpress (version 0.4 alpha 7) and analyzed by class comparisons among groups of arrays by BRB ArrayTools. Genes are listed in order from increasing to decreasing expression levels according to each comparison. Genomic sequences were identified by using the Affymetrix probe set target sequence to perform a BLAT search (http://genome.ucsc.edu/). When a representative gene was present, both strands were analyzed to determine the total number of occurrences of each sequence.

					Numbe	r of times seau	ence appears on	both strands
Class	Parametric	Fold	Probe set			in g	enomic DNA	
comparison	P value	change	ID	Description	AAGAAG	AAGAAGA	AAGAAGAA	AAGAAGAAG
GM15851 control vs	GM15851 1µM 1							
1	4.65E-03	1.56	219298_at	Enoyl Coenzyme A hydratase domain containing 3	27	9	3	1
GM15851 control vs	GM15851 2µM 1							
1	4.83E-03	1.74	1569106_s_at	Hypothetical protein FLJ10707	54	16	7	0
GM15851 control vs	GM15851 2µM 2							
1	1.89E-03	1.70	219298_at	Enoyl Coenzyme A hydratase domain containing 3	27	9	3	1
GM15850 control vs	GM15850 1µM 1							
1	4.51E-03	2.24	213060_s_at	Chitinase 3-like 2	14	3	1	1
2	3.41E-03	1.66	239862_at	Tumor protein D52	150	43	17	3
3	3.37E-03	1.52	237626_at	RB1-inducible colled-coll 1	94	49	22	4
4	2.64E-03	1 39	243149_at	Chromosome 14 open reading frame 24	201	14	10	15
6	1.97E-03	1.38	227232 at	Enah/Vasp-like	155	55	23	6
7	2.77E-03	1.33	242456_at	MRE11 meiotic recombination 11 homolog A (S. cerevisiae)	81	27	11	3
8	3.23E-03	1.32	240174_at	Hypothetical LOC388903	245	93	43	18
9	9.00E-07	1.28	232773_at	Hypothetical protein MGC13057	17	7	3	1
10	3.38E-03	1.20	1564639_at	Hypothetical LOC389908		No repr	esentative gene	
11	4.37E-03	1.20	1569714_at	Zinc finger, FYVE domain containing 20	19	9	4	1
12	4.09E-04 1.84E-04	1.17	239648_at 231958_at	Hypothetical protein MGC48972	43	15	4	0
13	2.20E-03	1.11	233681 at	Keratin associated protein 3-3	0	0	0	0
15	3.83E-03	1.10	1562516 at	Hypothetical protein FLJ10300	64	21	12	3
16	4.97E-03	1.09	227532_at	Hypothetical protein MGC14816	31	15	9	0
17	2.67E-03	1.08	1560021_at	Ribosomal protein S20		No repr	esentative gene	
18	5.76E-04	1.07	239773_at	Hypothetical protein BC004923	43	18	7	1
19	4.45E-03	1.06	1553618_at	Tripartite motif-containing 43	9	4	1	0
20	1.69E-03	1.04	1563540_at	MRNA; cDNA DKF2p68600511 (from clone DKF2p68600511)	240	No repi	esentative gene	22
21	4.15E-03 2.51E-03	1.04	233616_at 213453 x at	Candidate tumor suppressor protein	318	126	58	22
23	1.71E-03	0.93	242404 at	Transcribed locus	-	No rep	esentative dene	0
24	1.54E-03	0.90	222130 s at	FtsJ homolog 2 (E. coli)	4	3	2	0
25	3.83E-03	0.89	1562524_at			No repr	esentative gene	
26	4.76E-03	0.89	1554820_at	Hypothetical protein LOC340351	155	70	24	4
27	2.83E-03	0.87	1559403_at	High mobility group nucleosomal binding domain 3	28	8	1	0
28	6.44E-04	0.87	237619_at	Chromosome 6 open reading frame 146	5	1	0	0
29	2.22E-03	0.87	237124_at	Transcribed locus	407	No repi	esentative gene	
30	3.48E-03	0.87	1553672_at	Enabled homolog (Drosophila)	167	64 No ropi	2b	4
32	3.15E-03	0.85	244770_at 228787 s at	Reast carcinoma amplified sequence 4	41	14	q q	2
33	4.43E-03	0.84	218494 s at	SI C2A4 regulator	1	0	0	0
34	4.13E-03	0.83	1561391 at	Staufen, RNA binding protein, homolog 2 (Drosophila)	354	128	56	8
35	3.97E-04	0.83	207253_s_at	Ubinuclein 1	32	15	8	3
36	4.81E-03	0.82	210776_x_at	Transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47)	17	4	0	0
37	1.20E-03	0.81	201793_x_at	Chromosome 1 open reading frame 16	94	36	15	5
38	3.23E-03	0.81	222138_s_at	WD repeat domain 13	2	0	0	0
39	4.21E-04	0.80	229269_x_at	Single-stranded DNA-binding protein 4	170	1	0	0
40	4.00E-03	0.80	215176_X_at	HRV Fab 027-VL	479	No rep	I O	10
41	2.62E-03	0.80	233982 x at	Dual-specificity phosphatase 24 (putative)		No rep	esentative gene	
43	4.91E-04	0.78	201234 at	Integrin-linked kinase	7	3	0	0
44	3.28E-03	0.78	1554907_a_at	Hydrocephalus inducing	242	85	35	5
45	6.76E-04	0.77	205112_at	Phospholipase C, epsilon 1	334	118	44	13
46	4.31E-03	0.75	218954_s_at	BRF2, subunit of RNA polymerase III transcription initiation factor, BRF1-like	3	0	0	0
47	2.48E-03	0.75	202383_at	Smcy homolog, X-linked (mouse)	27	9	1	0
48	4.57E-03	0.75	204908_s_at	B cell CLL/lymphoma 3	4	2	0	0
49	2.21E-U3	0.75	1004129_a_at	cumonnate phosphoribosylitaristerase (nicotinate-nucleotide pyrophosphorylase (carboxylating))	1	2	1	U
50	2 36E-03	0.72	220306 at	Family with sequence similarity 46 member C	23	5	2	1
51	2.97E-03	0.63	219348 at	Uncharacterized hematopoietic stem/progenitor cells protein MDS032	3	1	1	0
CM45050 control up	CM45950 2M 4							
GIVIT5650 CONTROLVS	3 80 E-05	19.61	203200 at	Major histocompatibility complex, class II, DO alpha 1	4	1	0	0
2	2.15E-05	6.02	213060 s at	Chitinase 3-like 2	14	3	1	1
3	2.20E-03	3.12	211633 x at	Immunoglobulin heavy constant gamma 1 (G1m marker)	371	175	56	10
4	4.87E-03	2.95	203324_s_at	Caveolin 2	12	5	1	0
5	3.60E-03	2.75	211634_x_at	Immunoglobulin heavy constant mu	837	378	121	24
6	2.62E-03	2.48	205565_s_at	Frataxin	36	13	10	8
7	1.92E-03	2.32	230735_at	Interferon (alpha, beta and omega) receptor 2	36	13	8	2
8	4.62E-03	2.25	217022_s_at	Immunoglobulin heavy constant alpha 1 /// immunoglobulin heavy constant alpha 2 (A2m marker) /// hypothetical protein MGC27165	320	141	48	10
٩	1.52E-03	2.18	208029 s at	Lysosomal associated protein transmembrane 4 beta	57	15	7	1
10	1.05E-03	2.15	208767_s_at	Lysosomal associated protein transmembrane 4 beta	57	15	7	1
11	4.49E-03	2.02	225133_at	Kruppel-like factor 3 (basic)		No repr	esentative gene	
12	3.67E-03	2.00	227198_at	Lymphoid nuclear protein related to AF4		No repr	esentative gene	
13	2.33E-03	1.71	242399_at	Hypothetical protein FLJ33708	319	122	52	12
14	2.58E-03	1.65	1558186_s_at	Noncoding transcript (CLLU1 gene), splice variant 3		No repr	esentative gene	
15	4.45E-03	1.50	230961_at	WAP four-disulfide core domain 2	0	No repi	esentative gene	c
16	3.01E-U3	0.55	203092_at	WAL TOUR-GROUNDE COTE CONTIAIN 2	ő	٢	1	U
GM15850 control vs	GM15850 2µM 2		4550005		<i>c</i>	6		<i>c</i>
1	3.78E-03	1.83	1552287_s_at	AFG3 A Fase family gene 3-like 1 (yeast)	8	3	1	0
2 3	4.43E-03	0.60	201169 s at	Basic helix-loon-helix domain containing class R 2	4	1 NO TEPI	0	0
5		5.00	a				-	2

affected in GM15851 cells) (Table 2.3).

To examine the overall changes in gene-expression profiles in treated and untreated populations of FRDA and control cells, we first determined the genes whose expression was called significantly different among all conditions of untreated and treated GM15851 and GM15850 cells. At a *P* value of  $\leq 0.005$ , this class comparison generated a total of 632 genes. We then generated a correlation graph of the difference in geometric means of intensities for these genes between untreated and treated cells for each experimental condition (Figure 2.7). Neither the match polyamide **1** nor mismatch **2** affected the profile of GM15851 cells (slope of the correlation between conditions = -0.04 to +0.04, Figure 2.7A–D), whereas polyamide 1-treated GM15850 FRDA cells (at 2  $\mu$ M) have a gene-expression profile that approaches that of untreated GM15851 cells compared with untreated FRDA cells (Figure 2.7, compare G with E; slope = 0.69 compared with 0.97,



*Figure 2.6.* Microarray analysis of polyamide effects on global gene expression. Correlation of significant genes called from a class comparison of all arrays (n = 632 genes at  $P \le 0.005$ ). Cells were cultured with and without the addition of 1 at 1 µM or 2 µM or 2 at 2 µM for 7 days prior to RNA extraction. Each graph represents the difference between the geometric mean of intensities (from an average of the logged RMA data for each condition<sup>38</sup>) of untreated GM15851 cells (denoted 851 in the graphs) and untreated GM15850 cells (denoted 850) plotted versus the difference between the average from all GM15851 geometric means of intensities and each of the individual cell types and conditions. Thus the untreated controls for GM15851 and GM15850 cells should give slopes for the least squares fit line approaching zero and one, respectively. Graphs for normal GM15851 cells are shown for a) the untreated control and b) cells treated with 1 at 1 µM, c) 2 µM, and d) 2 at 2 µM. Graphs for FRDA GM15850 cells are shown for e) the untreated control and f) cells treated with 1 at 1 µM, g) 2 µM, and h) 2 at 2 µM.

respectively). This effect is seen to a lesser degree at 1  $\mu$ M polyamide 1 (Figure 2.7F), but was not seen with the mismatch polyamide 2 (Figure 2.7H). These changes in gene expression in the affected cell line may be a consequence of changes in frataxin protein levels,<sup>29,30</sup> or some could be direct effects of the polyamide, as suggested by the occurrence of 1 binding sites in up-regulated genes (Table 2.2). Taken together, these data lead us to conclude that polyamide 1 increases frataxin gene expression and, perhaps, downstream targets of frataxin, but this molecule has a limited effect on global gene expression. A search of the GenBank database reveals that most regions of the GAA·TTC DNA sequence [(GAA·TTC)<sub>6</sub> or longer] are present in nontranscribed repetitive DNA elements (including Alu sequences) in the human genome.<sup>31</sup> Polyamide 1 has no effects on lymphoid cell morphology, metabolism, or growth in culture.

### Influence of Polyamides on Sticky-DNA Conformation.

The capacity of sequence-specific polyamides to disrupt the intramolecular sticky-DNA structure formed by GAA·TTC repeat tracts was investigated. Plasmids harboring the sticky-DNA structure are visualized by gel electrophoresis after restriction endonuclease cleavage (Figure 2.7A). Linear DNA is indicative of disruption of the sticky-DNA structure by a polyamide, whereas the cleaved sticky-DNA band that migrates with a much slower mobility reveals no influence of the ligand. Plasmid pRW4886, which contains two tracts of (GAA·TTC)<sub>176</sub> in a direct-repeat orientation,<sup>10</sup> was incubated with each of the polyamides **1–4** at concentrations ranging from 0 to 50  $\mu$ M. The polyamide-bound DNA was then digested with EcoNI and electrophoresed on 1% agarose gels to determine the amount of the EcoNI-cleaved sticky-DNA retarded band present (Figure 2.7B). Incubation of pRW4886 DNA with **3** shifted the equilibrium from a maximum amount of sticky DNA to a complete loss of the EcoNI-cleaved sticky-DNA retarded band at a concentration of 50 nM. Polyamide **4**, which is a mismatch of **3** and has a  $K_d$  value ≈1,000 times higher than **3** (Table 2.1), did not affect the stability of sticky DNA in pRW4886 until a concentration



*Figure 2.7.* Effect of polyamide binding to plasmid DNA on sticky DNA stability. a) Illustration showing the assay for influence of polyamides on sticky DNA conformation. b) The capacity of a polyamide to disrupt the DNA·DNA associated region in the sticky DNA structure was revealed by the formation of linearized pRW4886 rather than the EcoNI cleaved sticky DNA, which showed the absence of a perturbing influence of the ligands. The polyamides that had higher  $K_d$  values required higher concentrations to observe the disruption of the DNA·DNA associated region of sticky DNA. The polyamides used were: 1 (green), 2 (blue), 3 (black) and 4 (red).

of 5  $\mu$ M. For polyamide 1, a 1  $\mu$ M concentration was needed to dissociate the DNA·DNA structure-forming region. The mismatched polyamide 2, having the highest  $K_d$  value of all of the polyamides tested, showed no effect on sticky-DNA stability, even at a 100  $\mu$ M concentration. Thus, the binding affinities of the polyamides for the GAA·TTC sequence had an intimate relationship with the concentration needed to shift the equilibrium from the DNA·DNA-associated structure to the duplex conformation (Figure 2.7B). The absence of the EcoNI-cleaved sticky-DNA-retarded band demonstrated the capacity of the sequence specific polyamide binding to shift the non-B–B-DNA equilibrium toward a conventional DNA duplex conformation in supercoiled plasmids. Because sticky DNA inhibits transcription,<sup>5,32</sup> and because the polyamides destabilize this conformation by shifting the structural equilibrium to duplex B DNA, increases in frataxin mRNA observed with polyamide 1 may be due to this structural transition, although other mechanisms, such as heterochromatin desilencing, must be considered.<sup>19</sup>

### Conclusions

Increased frataxin transcription with GAA·TTC-specific polyamides may be due to a change in DNA structure, as suggested by reversal of sticky DNA, which then allows for transcription through expanded GAA·TTC-repeat DNA, or by reversing heterochromatin caused by expanded GAA·TTC repeats.<sup>14</sup> Polyamides designed to target the related sequence GAGAA·TTCTC repeat DNA, which is found in *Drosophila* satellite V, have been shown to alter gene expression in developing embryos by opening the chromatin domains containing these repeated sequences<sup>14,19</sup> and to displace heterochromatin protein 1 (HP1) and the other chromosomal proteins D1 and topoisomerase II.<sup>33</sup> Interestingly, in the transgene study reported by Festenstein and colleagues,<sup>14</sup> HP1 was found to promote gene silencing by expanded GAA·TTC-repeat DNA. Thus, polyamides targeting GAA·TTC-repeat DNA may increase transcription of genes that harbor these sequences by displacement of similar repressor proteins and reversal of inactive heterochromatin to a more active chromatin structure.

#### **Materials and Methods**

### Polyamide Synthesis and Characterization.

Polyamides were synthesized by solid-phase methods<sup>18,34</sup> and their identity and purity verified by MALDI-TOF MS and analytical HPLC. Fluorescent conjugates were prepared by coupling BODIPY FL C5 (Molecular Probes) to the carboxyl terminus.<sup>27</sup> Binding affinities for match and mismatch sites were determined by quantitative DNase I footprinting. A plasmid harboring six GAA·TTC repeats was constructed by cloning the oligonucleotide 5'-GCCTTACGGTTACACTTGATGAAGAAGAAGAAGAAGAAGA ATTCGCAATGCCATTGCGCTATGA-3'·3'-ACGGAATGCCAATGTGAACTACTT CTTCTTCTTCTTCTTAAGCGTTACGGTAACGCGATAC-5' in the pCR2.1 TOPO vector (Invitrogen), and 251-bp singly end-labeled PCR product was generated from this plasmid with the following oligonucleotides: 5'-GAAAGACCCGTGTGTAAAGCC-3' and 5'-CTCGATATCTGCAGAATTGCC-3', where the second oligonucleotide was labeled with  $\gamma$ -[<sup>32</sup>P]-ATP and polynucleotide kinase, by using standard procedures, to generate a PCR product labeled on the GAA strand. A 204-bp singly end-labeled PCR product was derived from plasmid pMP142 DNA, containing 33 GAA·TTC repeats,<sup>5</sup> with the following oligonucleotides: 5'-GGCCAACATGGTGAAACC-3' and 5'-GTAGCTGGGATTACAGGCGC-3'. The first oligonucleotide shown was radiolabeled as above to generate a PCR product labeled on the GAA strand. A 150-bp PCR product containing a (GGA·TCC)<sub>3</sub> mismatch sequence was derived from the erbB2 (Her2-neu) promoter in human genomic DNA with the following oligonucleotides: 5'-CTTGTTGGAATGCAGTTGGA-3' and 5'-GGTTTCTCCGGTCCCAAT-3', with the first oligonucleotide radiolabeled.

### Cell Culture.

Epstein–Barr virus-transformed lymphoblast cell lines GM15850 from a FRDA patient (alleles with 650 and 1,030 GAA·TTC repeats in the frataxin gene, from the Coriell Cell Repository) and GM15851 from an unaffected sibling (normal range of repeats) were propagated in RPMI medium 1640 with 2 mM L-glutamine and 15% FBS at 37°C in 5% CO<sub>2</sub>. Cell growth and morphology were monitored by phase-contrast microscopy and viability by trypan blue exclusion and an ATP assay (ApoSENSOR; BioVision). Polyamides were added directly to the culture medium in PBS, and incubations were for the times indicated in the text and figure legends. Nuclear localization of the polyamides was verified by deconvolution microscopy, as described.<sup>35</sup>

### **Real-Time qRT-PCR.**

Real-time qRT-PCR analysis was performed essentially as described,<sup>36</sup> by using the following primers for the frataxin gene: 5'-CAGAGGAAACGCTGGACTCT-3' and 5'-AGCCAGATTTGCTTGTTTGG-3'. RNA was standardized by quantification of GAPDH mRNA,<sup>37</sup> and all values are expressed relative to GAPDH. qRT-PCR was performed by using iScript One-Step RT-PCR kit with SYBR green (Bio-Rad). Statistical analysis was performed on three independent qRT-PCR experiments for each RNA sample, and triplicate cell incubations were performed.

### Western Blot Analysis.

Total-cell extracts were used for SDS/PAGE and Western blotting with antibodies to human frataxin (Chemicon) or actin (Santa Cruz Biotechnology) as a control for cell number and protein loading. Signals were detected by chemiluminescence after probing the blot with HRP-conjugated secondary antibody (Supersignal West; Pierce). To quantify the relative levels of proteins, autoradiograms (within the linear response range of x-ray film) were converted into digital images and the signals quantified by using Molecular Dynamics ImageQuant software.

### **DNA Microarrays.**

FRDA and control lymphoid cells were incubated with polyamide 1 (at 1 or 2  $\mu$ M) or 2 (at 2  $\mu$ M), or in the absence of polyamide, in triplicate for 7 days before RNA purification and microarray analysis at the California Institute of Technology microarray facility. Affymetrix U133A Plus 2.0 GeneChips were hybridized in groups of eight for each of the three replicates. Raw GeneChip data were normalized with RMAExpress,<sup>38</sup> and the normalized data were filtered to remove probe sets called absent on 24 of 24 chips from class comparisons. The Affymetrix probe set-level data were imported to BRB Arraytools (Version 3.3.0 Beta 3a), selecting the U133 chips used in the experiment and leaving all filters off. For class comparisons between groups of arrays, unpaired samples were used, and the random variance model was selected, with the univariate significance threshold set to 0.005. The restrictions for the univariate test were maintained as the default values of 10 for the maximum number of false discovered genes, 0.1 for the maximum proportion

of false discoveries, and a 90% confidence level. Because of poor data correlation in one set of replicates, class comparisons were performed by using all chips for the control group versus two of the three replicates for the treatment group (five groups are the minimum number required for class comparisons). Microarray data (accession no. GSE5040) have been deposited at Gene Expression Omnibus.

### Effects of Polyamides on Sticky DNA.

Plasmid pRW4886, which contains two tracts of  $(GAA \cdot TTC)_{176}$  in a direct repeat orientation<sup>10</sup> and forms sticky DNA,<sup>13</sup> was treated with polyamides at concentrations of 0–50 µM at 37°C for 1 h. Restriction digestion with EcoNI after the polyamide incubation enabled visualization of the presence or absence of an EcoNI-cleaved sticky-DNA band that runs with decreased mobility compared with the linearized plasmid on 1% agarose gels.<sup>13</sup> Quantitation was by densitometric analysis using FluorChem (Version 3.04; Alpha Innotech).

We thank Malcolm Wood for assistance with deconvolution microscopy, Steve Head for statistical analysis of microarray data, and members of the Millard and Muriel Jacobs Genetics and Genomics Laboratory at California Institute of Technology for technical assistance and data analysis. This work was supported by National Institutes of Health Grants NS048989 (to J.M.G.), GM27681 (to P.B.D.), and NS37554 and ES11347 (to R.D.W.); the Friedreich's Ataxia Research Alliance; and the Seek-a-Miracle Foundation (Muscular Dystrophy Association).

### References

- Campuzano, V.; Montermini, L.; Molto, M. D.; Pianese, L.; Cossee, M.; Cavalcanti,
   F.; Monros, E.; Rodius, F.; Duclos, F.; Monticelli, A.; Zara, F.; Canizares, J.;
   Koutnikova, H.; Bidichandani, S. I.; Gellera, C.; Brice, A.; Trouillas, P.; De Michele,
   G.; Filla, A.; De Frutos, R.; Palau, F.; Patel, P. I.; Di Donato, S.; Mandel, J. L.;
   Cocozza, S.; Koenig, M.; Pandolfo, M. *Science* 1996, *271*, 1423–7.
- Montermini, L.; Richter, A.; Morgan, K.; Justice, C. M.; Julien, D.; Castellotti, B.; Mercier, J.; Poirier, J.; Capozzoli, F.; Bouchard, J. P.; Lemieux, B.; Mathieu, J.; Vanasse, M.; Seni, M. H.; Graham, G.; Andermann, F.; Andermann, E.; Melancon, S. B.; Keats, B. J.; Di Donato, S.; Pandolfo, M. *Ann. Neurol.* **1997**, *41*, 675–82.
- Wells, R. D.; Warren, S. T., eds. *Genetic Instabilities and Hereditary Neurological Diseases*; Academic Press: San Diego, 1998.
- (4) Pandolfo, M. Semin. Pediatr. Neurol. 2003, 10, 163–72.
- (5) Ohshima, K.; Montermini, L.; Wells, R. D.; Pandolfo, M. J. Biol. Chem. 1998, 273, 14588–95.
- (6) Bidichandani, S. I.; Ashizawa, T.; Patel, P. I. Am. J. Hum. Genet. **1998**, 62, 111–21.
- (7) Grabczyk, E.; Usdin, K. *Nucleic Acids Res.* **2000**, *28*, 2815–22.
- Sakamoto, N.; Chastain, P. D.; Parniewski, P.; Ohshima, K.; Pandolfo, M.; Griffith,
  J. D.; Wells, R. D. *Mol. Cell* 1999, *3*, 465–75.
- (9) Vetcher, A. A.; Napierala, M.; Wells, R. D. J. Biol. Chem. 2002, 277, 39228–34.
- (10) Vetcher, A. A.; Napierala, M.; Iyer, R. R.; Chastain, P. D.; Griffith, J. D.; Wells, R. D. J. Biol. Chem. 2002, 277, 39217–27.
- (11) Napierala, M.; Dere, R.; Vetcher, A.; Wells, R. D. J. Biol. Chem. 2004, 279, 6444–54.
- (12) Vetcher, A. A.; Wells, R. D. J. Biol. Chem. 2004, 279, 6434–43.
- (13) Son, L. S.; Bacolla, A.; Wells, R. D. J. Mol. Biol. 2006, 360, 267–84.

- (14) Saveliev, A.; Everett, C.; Sharpe, T.; Webster, Z.; Festenstein, R. *Nature* 2003, 422, 909–13.
- (15) Grabczyk, E.; Usdin, K. Nucleic Acids Res. 2000, 28, 4930–7.
- (16) Dervan, P. B.; Edelson, B. S. Curr. Opin. Struct. Biol. 2003, 13, 284–99.
- (17) Trauger, J. W.; Baird, E. E.; Dervan, P. B. Nature 1996, 382, 559–61.
- (18) Urbach, A. R.; Dervan, P. B. Proc. Natl. Acad. Sci. U. S. A. 2001, 98, 4343-8.
- (19) Janssen, S.; Durussel, T.; Laemmli, U. K. Mol. Cell 2000, 6, 999–1011.
- (20) Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U. K. Mol. Cell 2000, 6, 1013–24.
- (21) Urbach, A. R.; Love, J. J.; Ross, S. A.; Dervan, P. B. J. Mol. Biol. 2002, 320, 55–71.
- (22) Gottesfeld, J. M.; Belitsky, J. M.; Melander, C.; Dervan, P. B.; Luger, K. J. Mol.
   *Biol.* 2002, 321, 249–63.
- (23) Shinohara, K.; Narita, A.; Oyoshi, T.; Bando, T.; Teraoka, H.; Sugiyama, H. J. Am.
   *Chem. Soc.* 2004, *126*, 5113–8.
- (24) Dickinson, L. A.; Burnett, R.; Melander, C.; Edelson, B. S.; Arora, P. S.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* 2004, *11*, 1583–94.
- (25) Trauger, J. W.; Dervan, P. B. Methods Enzymol. 2001, 340, 450–66.
- Warren, C. L.; Kratochvil, N. C.; Hauschild, K. E.; Foister, S.; Brezinski, M. L.;
  Dervan, P. B.; Phillips, G. N., Jr.; Ansari, A. Z. Proc. Natl. Acad. Sci. U. S. A. 2006, 103, 867–72.
- (27) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S. A.* 2003, *100*, 12063–8.
- (28) 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*, 2802–18.
- (29) Coppola, G.; Choi, S. H.; Santos, M. M.; Miranda, C. J.; Tentler, D.; Wexler, E. M.;
   Pandolfo, M.; Geschwind, D. H. *Neurobiol. Dis.* 2006, *22*, 302–11.
- (30) Schoenfeld, R. A.; Napoli, E.; Wong, A.; Zhan, S.; Reutenauer, L.; Morin, D.;

Buckpitt, A. R.; Taroni, F.; Lonnerdal, B.; Ristow, M.; Puccio, H.; Cortopassi, G. A. *Hum. Mol. Genet.* **2005**, *14*, 3787–99.

- (31) Clark, R. M.; Dalgliesh, G. L.; Endres, D.; Gomez, M.; Taylor, J.; Bidichandani, S.
  I. *Genomics* 2004, *83*, 373–83.
- (32) Sakamoto, N.; Ohshima, K.; Montermini, L.; Pandolfo, M.; Wells, R. D. J. Biol.
   *Chem.* 2001, 276, 27171–7.
- (33) Blattes, R.; Monod, C.; Susbielle, G.; Cuvier, O.; Wu, J. H.; Hsieh, T. S.; Laemmli, U. K.; Kas, E. *EMBO J.* 2006, *25*, 2397–408.
- (34) Baird, E. E.; Dervan, P. B. J. Am. Chem. Soc. 1996, 118, 6141–6146.
- (35) Dudouet, B.; Burnett, R.; Dickinson, L. A.; Wood, M. R.; Melander, C.; Belitsky, J. M.; Edelson, B.; Wurtz, N.; Briehn, C.; Dervan, P. B.; Gottesfeld, J. M. *Chem. Biol.* 2003, *10*, 859–67.
- (36) Chuma, M.; Sakamoto, M.; Yamazaki, K.; Ohta, T.; Ohki, M.; Asaka, M.; Hirohashi,S. *Hepatology* 2003, *37*, 198–207.
- (37) Pattyn, F.; Speleman, F.; De Paepe, A.; Vandesompele, J. Nucleic Acids Res. 2003, 31, 122–3.
- (38) Bolstad, B. M.; Irizarry, R. A.; Astrand, M.; Speed, T. P. *Bioinformatics* 2003, *19*, 185–93.

### **Chapter 2B**

# Footprinting Studies of Polyamides Targeting the [AAG]<sub>3</sub> Triplet Repeat and Related Sequences

### Abstract

The recent discovery of a linear, beta-alanine linked polyamide that reverses the transcription defect of Friedreich Ataxia (FRDA) has generated the impetus to more fully characterize the binding properties of this polyamide and other related ligands. Polyamides are a class of programmable, sequence-specific minor groove DNA-binding small molecules, derived from the natural product distamycin, that hold promise for sequencespecific transcriptional modulation. The newly discovered ligand belongs to a subset of polyamides that bind narrower tracts of B DNA in a 1:1 ligand:DNA stoichiometry. This class of polyamides has received less attention due to its presumed sequence promiscuity. Interestingly, few genes beyond frataxin were found to be affected by the polyamide in genomic microarray transcript analysis. To more fully understand the binding properties of this and other related linear, beta-linked polyamides, we have synthesized eleven polyamides to study representative effects of polyamide length, polyamide amino acid mismatches, polyamide dye conjugates, and an altered polyamide tail on binding affinity and specificity for polypurine tracts of DNA. Polyamides with fewer than seven monomers appear to be the poorest DNA binders. The other synthesized polyamides maintained binding affinities of at least 10<sup>8</sup> M<sup>-1</sup>.

### Introduction

Friedreich Ataxia (FRDA), a neurodegenerative disorder characterized by the intronic 5'-GAA-3' trinucleotide repeat hyperexpansion (Figure 2.8A) that affects roughly 1 of 30,000 Caucasians,<sup>1</sup> may be caused by an equilibrium shift from B DNA toward a higher-order DNA structure<sup>2–6</sup> or by heterochromatin stabilization<sup>7</sup> that impedes transcription of the frataxin gene.<sup>8–10</sup> A ligand that stabilizes B DNA may shift the equilibrium towards B DNA or may aide heterochromatin opening,<sup>7,11</sup> enabling transcription. Polyamides, minor-groove binding, B DNA stabilizing small molecules, effect increased frataxin transcription in immortalized FRDA cell culture at micromolar concentrations.<sup>12</sup> Traditionally, the polyamide core contains *N*-methylimidazoles (Im), *N*-methylpyrroles (Py), and beta-alanines ( $\beta$ ) that specify bases read from minor groove hydrogen bond contacts. A binding

code relating polyamide structure to DNA sequence exists for hairpin polyamides and their unlinked and less entropically favored progenitors, the 2:1 binding polyamides.<sup>13–17</sup>

Despite the breadth of data studied for the hairpin and 2:1 binding polyamides, a 1:1 binding linear  $\beta$ -linked polyamide aptly targeted 5'-AAGAAGAAG-3' (Figure 2.8B) and increased frataxin mRNA transcript levels.<sup>12</sup> It is believed that this class of polyamides binds narrower DNA tracts, such as some polypurine sequences. Previous studies on a linear  $\beta$ -linked



*Figure 2.8.* Frataxin gene structure and model of polyamide-DNA binding. a) A schematic showing the location of the 5'-GAA-3' repeat expansion in the frataxin gene. For afflicted individuals, x < n < 1,800. b) A model of polyamide 1 binding its target sequence. Red residues represent imidazoles, yellow residues pyrroles, white residues beta-alanines, and gray, the Dp tail. A ball-and-stick model oriented analogously to the 3D-model is shown below with filled circles representing imidazoles, hollow circles pyrroles, greyed diamonds beta-alanines, and the semi-circle with a plus, the Dp tail.

polyamide known to bind 5'-AAGAG-3' telomeric repeats<sup>11,18</sup> provided the foundations for further examination of binding preferences for this class of polyamides.<sup>19–21</sup> A preliminary binding code derived from quantitative DNase I footprinting and from Fe(II)-EDTA affinity cleavage experiments of the polyamide  $\text{Im}-\beta$ -Im-Py- $\beta$ -Im-Py- $\beta$ -Dp (-EDTA) aligns the polyamide from N-terminus to C-terminus with respect to the polypurine DNA strand read from 3' to 5', pairing *N*-methylimidazole (Im) with any base pair and *N*-methylpyrrole (Py) or  $\beta$ -alanine ( $\beta$ ) with A·T or T·A base pairs.<sup>19-21</sup> These rules enabled the design of Im-Py- $\beta$ -Im-Py- $\beta$ -Im- $\beta$ -Dp (Figure 2.9, polyamide 1), targeting 5'-AAGAAGAAG-3'. A constitutional isomer of 1, Im-Py- $\beta$ -Im-Im- $\beta$ -Py- $\beta$ -Dp (Figure 2.9, polyamide 2) fails to bind 5'-AAGAAGAAG-3' and consequently does not increase frataxin transcription.<sup>12</sup> Strikingly, while polyamide 1 may bind any sequence of the composition 5'-WWNWWNWN-3', only a handful of genes beyond frataxin are modulated in its presence. The binding preferences of polyamides 1 and 2, have not been elucidated. One goal of this chapter is to establish the preferred binding orientation, to study the preferred binding stoichiometry, and to examine other polypurine tracts polyamide 1 may bind. Likewise, this chapter seeks to establish that polyamide 2 binds DNA.

Because of the biological ramifications of polyamides **1** and **2**, it is necessary to explore other small molecules that may alleviate the transcription defect of the frataxin gene. While polyamide **1** upregulates frataxin expression in Epstein-Barr virus immortalized cell culture samples,<sup>12</sup> a variety of polyamides harboring similar binding affinities and specificities may increase chances of successful frataxin upregulation in an FRDA mouse model.<sup>22</sup> A polyamide consists of aromatic heterocycle and aliphatic amino acid monomers as well as a tail, which is typically 3-(dimethylamino)-1-propylamine (Dp). Polyamide **1** offers two obvious modification points, the truncation of eight monomers to fewer and the alteration of the Dp tail.

While polyamide **2** effected no change in frataxin mRNA transcript levels,<sup>12</sup> this polyamide contains two mutations from polyamide **1**. It is unclear if this double mutation

is necessary to prevent binding to 5'-AAGAAGAAG-3'. Single amino acid mutation polyamides should bind 5'-AAGAAGAAG-3' more poorly than polyamide **1**. Their relative affinity for this site as compared to **2** will be studied.

Polyamide-dye conjugate compounds enable the visualization of polyamide nuclear localization<sup>12,23–25</sup> and may aide the likelihood of nuclear localization and subsequent cellular activity.<sup>25,26</sup> A final consideration studies the energetic penalty of attaching a fluorescent dye to polyamide **1**.

The characterization of this series of ten polyamides by quantitative DNase I footprinting and in one instance methidium-propyl EDTA (MPE) footprinting, and a polyamide-EDTA conjugate by affinity cleavage footprinting marks a first step toward expanding and understanding the known repertoire of 1:1 binding linear  $\beta$ -linked polyamides.

### Results

Eleven polyamides have been synthesized (Figure 2.9) related to the match polyamide **1** and the mismatch polyamide **2** that were originally used in immortalized cell culture frataxin expression studies.<sup>12</sup> Polyamides were incubated with pJWP2, pJWP5, or pJWP10, as noted, for 14 h prior to DNase I cleavage, methidium propyl EDTA (MPE) digestion, or affinity cleavage. Initial specificity studies of polyamides **1** and **2** are reported. Furthermore, the effects of four variables relating to polyamide **1** or **2** on binding affinity were studied—the length of the DNA binding ligand, the composition of the charged carboxy-terminal tail appended to the polyamide, the mutation of a single amino acid on the ligand, and the presence of the BODIPY FL dye (Figure 2.9).

### **Plasmid Design**

Three plasmids, pJWP2, pJWP5, and pJWP10 were designed to harbor binding sites for polyamides **1–8** and polyamides **9–10** (Figure 2.10). Plasmid pJWP2 contains



*Figure 2.9.* A list of polyamide structures utilized in footprinting studies on  $[AAG]_3$  repeats. The ball-and-stick structures contain filled circles (*N*-methylimidazoles), empty circles (*N*-methylpyrroles), and greyed diamonds ( $\beta$ -alanines), connected by straight lines (amide bonds). The semi-circle with a plus represents 3-(dimethylamino)-1-propylamine, the filled hexagon represents 4-(2-aminoethyl) morpholine, and the double-plus or a plus in parentheses represents 3,3'-diamino-*N*-methyl-dipropylamine. Dyes and EDTA are represented by their written name appended to the C-terminal tail.

four binding sites – 5'-AAGAAGAAG-3' (Figure 2.10, Site III) for polyamide 1, 5'-AAAAGGAAG-3' (Figure 2.10, Site IV) for polyamide 2, 5'-AAAAAGAAG-3' (Figure 2.10, Site V) for polyamide 7, and 5'-AAGAAAAG-3' (Figure 2.10, Site VI) for polyamide 8. Binding sites V and VI contain single mismatch transitions (G to A) relative to binding site III; binding site IV contains a double mismatch transition (G to A and A to G), enabling specificity of *N*-methylpyrrole and *N*-methylimidazole to be tested for several of the eight monomer polyamides. The spacer sequence between binding sites has been held constant – 5'-CCGGGGCC-3' – analogous to the spacers originally used to



*Figure 2.10.* Plasmid insert sequences utilized for footprinting [AAG]<sub>3</sub>-targeting polyamides. a) Schematic showing the sequence of the hybridized insert, 5'-<sup>32</sup>P label position, and PCR amplicon for pJWP10 plasmid containing the boxed 11-base pair and 10-base pair 1:1 and 2:1 stoichiometry binding sites, I and II, respectively. b) Schematic showing the sequence of the hybridized insert, 5'-<sup>32</sup>P label position, and PCR amplicon for pJWP2 plasmid containing the boxed 9-base pair target sites III, IV, V, and VI. c) Schematic showing the sequence of the hybridized insert, 5'-<sup>32</sup>P label position, and PCR amplicon for pJWP5 plasmid containing the boxed 12-base pair target sites VII, VIII, IX, and X.

study binding specificity and affinity of 1:1 binding  $\beta$ -linked linear polyamides.<sup>20</sup> Plasmid pJWP5 contains the same sequence composition as pJWP2 but with 5'-AAG-3' appended immediately flanking the 5' portion of each binding site (Figure 2.10)—dye conjugate **10** is substantially longer than polyamide **1**. Plasmid pJWP10 contains binding sites for the 1:1 ligand:DNA stoichiometry, 5'-AAGAAGAAGAA-3' (Figure 2.10, Site I) and 2:1 ligand:DNA stoichiometry 5'-AGCAGCAGCA-3' (Figure 2.10, Site II). Site I enables the elucidation of the binding orientation for polyamide **1** against the polypurine strand of DNA.

### **Characterization of Polyamide 1 Binding Affinity, Stoichiometry, and Orientation**

Polyamide **1** has been footprinted on the amplicon from pJWP10 to examine preference for 1:1 polyamide:DNA binding stoichiometry versus 2:1 polyamide:DNA binding stoichiometry. Figure 2.11A shows the DNase I footprints of polyamide **1** for



*Figure 2.11.* Quantitative DNase I footprint titrations for an [AAG]<sub>3</sub>-targeting polyamide and a 4-(2-aminoethyl) morpholine tail variant. Binding stoichiometry preferences (A) and a limited set of binding preferences (B) for polyamide 1 and (C) for polyamide 2. Lane 1 contains intact DNA from pJWP10 (A) or pJWP2 (B, C); lanes 2 and 3 contain G and A chemical sequencing reactions; and lane 4 contains DNase I digested DNA. Lanes 5–15 contain 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide concentrations, respectively along with DNA from pJWP10 (A) or pJWP2 (B, C). DNase I digestion of lanes 5–15 proceeds under identical conditions to lane 4. Binding sites from pJWP10 and pJWP2 are identified as roman numerals to the right of each gel. Binding isotherms are shown directly beneath each gel. Hollow circles represent binding site I and filled circles binding site II for A. Hollow triangles represent binding site III, hollow circles binding site V, and filled circles binding site VI for B.

these two binding sites and the corresponding binding isotherms derived from these footprints. Table 2.4 shows the corresponding binding affinity data for polyamide 1 on pJWP10. Like the previously reported polyamide, Im- $\beta$ -Im-Py- $\beta$ -Im- $\beta$ -Im-Py- $\beta$ -Dp that targets 5'-AAAGAGAAGAG-3' with 31.5-fold preference over 5'-AGCGCAGCGCT-3',<sup>21</sup>

**Table 2.4.** Binding affinities  $(K_a, M^{-1})$  of polyamide **1** to sites **I** and **II**. Columns are divided by binding site, with sequences listed 5' to 3' below the binding site number. Values reported are the mean values from at least three DNase I footprinting titration experiments, with the standard deviation given in parentheses. Bracketed numbers below the binding affinities represent the fold decrease in binding affinity relative to the highest affinity site in the row.

		I	
Numbe	r Polyamide A	AGAAGAAGAA	AGCAGCAGCA
1	+)~~~~~~	4.6 (± 0.5) × 10 <sup>9</sup> [1]	$8.2 (\pm 0.3) \times 10^7$ [57]

polyamide 1 prefers to bind DNA in a 1:1 stoichiometry (site I, non-cooperative binding isotherms n = 1) with 57-fold preference over the 2:1 polyamide:DNA binding mode (site II, cooperative binding isotherms n = 2). Affinity cleavage experiments of 1E on pJWP10 (Figure 2.12) present cleavage patterns in concordance with the observed binding modes from the DNase I footprint derived isotherms. Additionally, the cleavage patterns of 1E at site I confirm polyamide orientation of N-terminus to C-terminus read from 3' to 5' against the polypurine strand of DNA. This binding orientation agrees with that previously



Figure 2.12. Methidium-propyl EDTA (MPE) footprinting of polyamide 1 and affinity cleavage of polyamide 1E on pJWP10. The affinity cleavage sites for the 1:1 binding stoichiometry site (I) and 2:1 binding stoichiometry site (II) are shown schematically to the right of the gel. Arrows representing the affinity cleavage sites are drawn proportional to the radiation intensity, as calculated from lane 11. Lane 1 contains intact DNA from pJWP10; lanes 2 and 3 contain G and A chemical sequencing reactions; and lane 4 contains MPE digested DNA. Lanes 5-9 and lanes 10-14 contain 100 nM, 30 nM, 10 nM, 3 nM, and 1 nM polyamide 1 or polyamide 1E concentrations, respectively as indicated, along with DNA from pJWP10. Lanes 5-9 were treated with MPE in an identical fashion to lane 4.

binding orientation agrees with that previously observed for Im- $\beta$ -Im-Py- $\beta$ -Im-Py- $\beta$ -Im-Py- $\beta$ -Dp.<sup>21</sup>

Intriguingly, polyamide 1 exhibits a 2.4-fold higher affinity to binding site III (Figure 2.11B and Table 2.5) than to binding site I (Figure 2.11A and Table 2.4) even though the core 5'-AAGAAGAAG-3' sequence remains the same. This disparity in affinity is likely due to unique DNA microstructures resulting from a difference in 3'-flanking sequence between the two binding sites. Site I has a 3'-flanking sequence of 5'-AACC-3', whereas

#### **Polyamide Length/Affinity Titration**

Polyamides 1 and 3–5 were synthesized to study the effect of number of monomers per polyamide (length) on binding affinity to 5'-AAGAAGAAG-3' (Figure 2.9). Figures 2.11 and 2.13 show the DNase I footprints of these compounds and the corresponding binding isotherms derived from these footprints. Table 2.5 summarizes the binding affinity data for polyamides 1 and 3–5. Polyamide 1, composed of 8 monomers, binds site III with highest affinity in this series ( $K_a = 1.1 \times 10^{10} \text{ M}^{-1}$ ). The seven-monomer polyamide 3 removed the C-terminal  $\beta$ -alanine, creating a fourfold energetic penalty for binding site III ( $K_a = 2.5 \times 10^9 \text{ M}^{-1}$ ). Further truncation proves to be energetically unfavorable. Two 5-monomer, truncated polyamides were synthesized, polyamide 4 and 5. Polyamide 4 removes three monomers from the C-terminus of polyamide 1; polyamide 5 removes three monomers from the N-terminus of polyamide 1 (Figure 2.9). While polyamide 4 harbors a threefold higher binding affinity than polyamide 5 ( $K_a = 1.8 \times 10^6 \text{ M}^{-1}$  for 4 versus  $K_a$ = 6.3 × 10<sup>5</sup> M<sup>-1</sup> for 5), polyamide 4 almost fully coats DNA at 100 mM concentration (Figure 2.13). No signs of coating were observed for polyamide 5. In comparison to polyamide 1, polyamides 4 and 5 have 6,100-fold and 17,000-fold lower binding affinities,

**Table 2.5.** Polyamide length titration series. Polyamide ball-and-stick structures and their corresponding numbers are identified in columns to the left of the reported  $K_a$  values (M<sup>-1</sup>). Values reported are the mean values from at least three DNase I footprinting titration experiments, with the standard deviation given in parentheses. Bracketed values underneath each  $K_a$  represent the fold decrease in binding affinity relative to this highest affinity binder for the row. Bracketed values to the right of each  $K_a$  represent the fold decrease in binding affinity relative to the highest affinity binder in the column. Assays were performed at 25 °C.

		ш	IV	V	VI
Number	r Polyamide	AAGAAGAAG	AAAAGGAAG	AAGAAAAAG	AAAAAGAAG
1	+)~~~~~~~	$\begin{array}{cc} 1.1 \ (\pm \ 0.05) \times 10^{10} & [1] \\ [1] \end{array}$	$\leq 1 \times 10^7$ [ $\geq 1,100$ ]	9.4 $(\pm 0.5) \times 10^9$ [1] [1.2]	$\begin{array}{cc} 2.7 \ (\pm \ 0.3) \times 10^9 & \textbf{[1]} \\ \textbf{[4.1]} \end{array}$
3	+)●◇○●◇○●	$\begin{array}{c} 2.5 \ (\pm \ 0.2) \times 10^9 \qquad \textbf{[4.4]} \\ \textbf{[1]} \end{array}$	$\leq 1 \times 10^7$ [≥250]	$\begin{array}{cc} 2.2 \ (\pm \ 0.2) \times 10^9 & \textbf{[4.3]} \\ \textbf{[1.1]} \end{array}$	$\begin{array}{ccc} 1.0 \ (\pm \ 0.2) \times 10^9 & \textbf{[2.7]} \\ \textbf{[2.5]} \end{array}$
4	+)))	$\begin{array}{ccc} 1.8 \ (\pm \ 0.6) \times 10^6 & \textbf{[6,100]} \\ \textbf{[1.1]} \end{array}$	$\leq 3.3 \times 10^5$ [26.1]	$1.2 (\pm 0.3) \times 10^6$ [7,800] [1.7]	$\begin{array}{c} 2.0 \; (\pm \; 0.4) \times 10^6  \textbf{[1,300]} \\ \textbf{[1]} \end{array}$
5	+)���	$\begin{array}{c} 6.3 \ (\pm \ 1.6) \times 10^5  \textbf{[17,000} \\ \textbf{[1]} \end{array}$	$ \leq 1 \times 10^4 $	$\begin{array}{c} 2.4 \ (\pm \ 0.6) \times 10^5 \ \textbf{[39,000]} \\ \textbf{[2.6]} \end{array}$	$5.0 \ (\pm \ 0.7) \times 10^5 \ \ \textbf{[5,400]} \\ \textbf{[1.3]}$
6		$\begin{array}{cc} 3.8 \ (\pm \ 0.7) \times 10^9 & \textbf{[2.9]} \\ \textbf{[1]} \end{array}$	$\leq 1 \times 10^7$ [ $\geq 380$ ]	$3.2 (\pm 0.6) \times 10^9$ [2.9] [1.2]	8.0 $(\pm 1.1) \times 10^8$ [3.4] [4.8]



*Figure 2.13.* Quantitative DNase I footprint titrations showing effects of polyamide length on binding preferences and affinities. Lane 1 contains intact DNA from pJWP2; lanes 2 and 3 contain G and A chemical sequencing reactions; and lane 4 contains DNase I digested DNA. Lanes 5–15 contain 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM of polyamide **3**, respectively for **A**. Lanes 5–15 contain 100 mM, 30 mM, 10 mM, 3 mM, 1 mM, 300 nM, 100 nM, 30 nM, 100 nM, 30 nM, 100 nM, 30 mM, 100 mM, 3 mM, 1 mM, 300 nM, 100 nM, 30 nM, 100 nM, 100 nM, 30 nM, 100 nM, 100 nM, 30 nM, 100 nM, 100 nM, 100 nM, 30 nM, 100 nM, 30 nM, 100

respectively.

### Binding Affinity and Specificity of Polyamide 1 and Its Ethyl-Morpholino Derivative

Polyamides 1 and 6 were synthesized to study their relative binding affinities for designed match and mismatch sites. It is believed that the 4-(2-aminoethyl)-morpholine

tail may enhance the *in vitro* and *in vivo* potency of the polyamide bearing the core sequence Im-Py- $\beta$ -Im-Py- $\beta$ -Im- $\beta$ .<sup>27</sup> Figures 2.11 and 2.13 show the DNase I footprints of these compounds and the corresponding binding isotherms derived from these footprints. Table 2.5 summarizes the binding affinity data comparing polyamides 1 and 6. Polyamide 1 binds site I with threefold higher affinity than polyamide 6 ( $K_a = 3.8 \times 10^9 \text{ M}^{-1}$ ). The preference for site III over sites IV–VI is roughly comparable between polyamides 1 and 6. Neither compound binds the double transition mismatch site IV at 30 nM concentrations. While some binding begins to occur at a concentration of 100 nM, there is an insufficient concentration range to accurately calculate the binding affinity at this site. Both compounds display little preference for the single transition mismatch site V over binding site III, and each has four or fivefold preference for site III over site VI, another single transition mismatch site.

#### **Polyamide Single and Double Amino Acid Mutations**

Three mismatch polyamides were designed and synthesized that should fail to bind site III of pJWP2 (Figure 2.9, polyamides 2, 7, and 8). Figure 2.14 shows the DNase I footprints of these compounds and the corresponding binding isotherms derived from these footprints. Table 2.6 summarizes the binding affinity data comparing polyamides 2, 7, and 8. Here, we report that polyamide 2 binds site IV with modest affinity ( $K_a = 2.4 \times 10^8 \text{ M}^{-1}$ ) and fails to bind sites III and V. A footprint begins to appear in binding site VI, a single G to A transition placed adjacent to an imidazole, although insufficient data precludes the calculation of an affinity for this site.

Polyamide **2** was formed by swapping two residues from polyamide **1**, an Im and a Py, to generate a constitutional isomer of polyamide **1**. Polyamides **7** and **8** represent single amino acid mutations from polyamide **1** (Figure 2.9). Polyamide **7** mutates an internal Im amino acid of polyamide **1** to a Py, forming a compound that should target binding site **V**. Polyamide **8** mutates the C-terminal Im amino acid of polyamide **1** to a Py, forming a



*Figure 2.14.* Quantitative DNase I footprint titrations examining binding affinities and preferences of double amino acid mismatch versus single amino acid mismatch polyamides on pJWP2. Lane 1 contains intact DNA from pJWP2; lanes 2 and 3 contain G and A chemical sequencing reactions; and lane 4 contains DNase I digested DNA. Lanes 5–15 contain 3  $\mu$ M, 1  $\mu$ M, 300 nM, 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, and 30 pM polyamide **2**, respectively for **A**. Lanes 5–15 contain 100 nM, 30 nM, 10 nM, 3 nM, 1 nM, 300 pM, 100 pM, 30 pM, 100 pM, 30 pM, 10 pM, 3 pM, and 1 pM polyamide 7 or polyamide **8**, respectively for **B** and **C**. DNase I digestion of lanes 5–15 proceeds under identical conditions to lane 4. Binding sites from pJWP2 are identified as roman numerals to the right of each gel. Binding isotherms are shown directly beneath each gel. Hollow triangles represent binding site **III**, filled triangles binding site **IV**, hollow circles binding site **V**, and filled circles binding site **V**.

compound that should target site VI. As predicted, polyamide 7 binds site V with  $K_a = 4.3 \times 10^{10} \text{ M}^{-1}$ , fourfold more strongly than polyamide 1 binds site III; polyamide 8 binds site VI with  $K_a = 4.2 \times 10^9 \text{ M}^{-1}$ , 2.6-fold more poorly than polyamide 1 binds site III. A single

**Table 2.6.** Polyamide alternative mismatch control series. Polyamide ball-and-stick structures and their corresponding numbers are identified in columns to the left of the reported  $K_a$ 's (M<sup>-1</sup>). Values reported are the mean values from at least three DNase I footprinting titration experiments, with the standard deviation given in parentheses. Bracketed values underneath each  $K_a$  represent the fold decrease in binding affinity relative to this highest affinity binder for the row. Bracketed values to the right of each  $K_a$  represent the fold decrease in binding affinity relative to the highest affinity binder for the row. Bracketed values to the right of each  $K_a$  represent the fold decrease in binding affinity relative to the highest affinity binder in the column. Assays were performed at 25 °C.

		III	IV	V	VI
Number	r Polyamide	AAGAAGAAG	AAAAGGAAG	AAGAAAAAG	AAAAAGAAG
1	+)♦●♦○●♦♦○●	$\begin{array}{cc} 1.1 \ (\pm \ 0.05) \times 10^{10} & [1] \\ [1] \end{array}$	$\leq 1 \times 10^7$ [ $\geq$ 24] [ $\geq$ 1,100]	$\begin{array}{c} 9.4\ (\pm\ 0.5)\times10^9 \qquad \textbf{[4.6]}\\ \textbf{[1.2]}\end{array}$	$\begin{array}{c} 2.7\ (\pm\ 0.3)\times10^9  \textbf{[1.6]}\\ \textbf{[4.1]} \end{array}$
2	€000€	$\leq 3.3 \times 10^5$ [ $\geq 33,000$ ] [ $\geq 730$ ]	$\begin{array}{c} 2.4\ (\pm\ 0.4)\times10^8 \qquad [1]\\ [1]\end{array}$	≤ 3.3 × 10 <sup>5</sup> [≥130,000] [≥730]	$\leq 1 \times 10^{6}$ [≥4,200] [≥240]
7	+)	$\begin{array}{cc} 4.9 \; (\pm \; 0.9) \times 10^8 & \textbf{[22]} \\ \textbf{[88]} \end{array}$	$\begin{array}{c} 2.7\ (\pm\ 1.5)\times10^7  \textbf{[8.9]}\\ \textbf{[1,600]} \end{array}$	$\begin{array}{ll} 4.3 \ (\pm \ 0.4) \times 10^{10} & [1] \\ [1] \end{array}$	$\begin{array}{ccc} 1.3 \ (\pm \ 0.1) \times 10^9 & \textbf{[3.2]} \\ \textbf{[33]} \end{array}$
8	+)♦0♦0♦0€	7.5 $(\pm 0.9) \times 10^7$ [150] [56]	$\begin{array}{c} 1.3 \ (\pm \ 0.2) \times 10^8  \textbf{[1.9]} \\ \textbf{[32]} \end{array}$	$1.3 (\pm 0.2) \times 10^8$ [330] [32]	$\begin{array}{cc} 4.2 \ (\pm \ 0.9) \times 10^9 & [1] \\ [1] \end{array}$

transition (A to G, site III) reduces binding affinity 88-fold for polyamide 7, a double transition (site VI) only 33-fold, and a triple transition (site IV) 1600-fold. Polyamide 8 harbors similar levels of specificity as polyamide 7. It binds both single and double transition mismatch sites IV and V 32-fold more poorly than its match site. A second single transition mismatch site (site III) lowers the binding affinity 56-fold.

### **Polyamide Dye Conjugate**

Polyamide dye conjugates have functioned to identify the nuclear localization of a polyamide.<sup>23–25</sup> One polyamide-fluorophore conjugate was synthesized (Figure 2.9, compound **10**) to study the energetic penalty resulting from the attachment of a fluorescent BO-DIPY FL dye as compared to the energetic baseline provided by polyamide **9**. The DNase I footprinting gels and binding isotherms of polyamides **9–10** are shown in Figure 2.15. The corresponding binding affinities to sites **VII–X** are shown in Table 2.7.

Polyamide 9 binds site VII on pJWP5 with  $K_a = 1.5 \times 10^{10}$  M<sup>-1</sup>. It binds degenerately to sites IX and X with roughly sixty percent the affinity it has to site VII. It binds with 63-fold lower affinity to site VIII. Attaching BODIPY FL to polyamide 9 yields polyamide 10, which harbors 21-fold lower affinity to binding site VII ( $K_a = 7.3 \times 10^8$  M<sup>-1</sup>). The



*Figure 2.15.* Quantitative DNase I footprint titrations showing the effects of BODIPY FL conjugation on binding preferences and affinities. Lane 1 contains intact DNA from pJWP5; lanes 2 and 3 contain G and A chemical sequencing reactions; and lane 4 contains DNase I digested DNA. Lanes 5-15 contain 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 9 or 10, respectively for A and B. DNase I digestion of lanes 5–15 proceeds under identical conditions to lane 4. Binding sites from pJWP5 are identified as roman numerals to the right of each gel. Binding isotherms are shown directly beneath each gel. Hollow triangles represent binding site VII, filled triangles binding site VIII, hollow circles binding site IX, and filled circles binding site X.

**Table 2.7.** Polyamide dye conjugate series. Polyamide ball-and-stick structures and their corresponding numbers are identified in columns to the left of the reported  $K_a$  values (M<sup>-1</sup>). Values reported are the mean values from at least three DNase I footprinting titration experiments, with the standard deviation given in parentheses. Bracketed values underneath each  $K_a$  represent the fold decrease in binding affinity relative to this highest affinity binder for the row. Bracketed values to the right of each  $K_a$  represent the fold decrease in binding affinity relative to the highest affinity binder in the column. Assays were performed at 25 °C.

		VII	VIII	IX	Х
Numb	er Polyamide	AAGAAGAAGAAG	AAGAAAAGGAAG	AAGAAGAAAAAG	AAGAAAAAGAAG
9	++)	$\begin{array}{c} 1.5 \ (\pm \ 0.1) \times 10^{10}  [1] \\ [1] \end{array}$	$\begin{array}{cc} 2.4 \ (\pm \ 0.8) \times 10^8 & \textbf{[1]} \\ \textbf{[63]} \end{array}$	$\begin{array}{c} 9.3 \ (\pm \ 0.7) \times 10^9  \textbf{[1]} \\ \textbf{[1.6]} \end{array}$	9.6 $(\pm 1.2) \times 10^9$ [1] [1.6]
10		$\begin{array}{c} 7.3 \ (\pm \ 1.4) \times 10^8  \textbf{[21]} \\ \textbf{[1]} \end{array}$	$1.3 (\pm 0.5) \times 10^7$ [19] [56]	$\begin{array}{c} 6.1 \ (\pm \ 0.2) \times 10^8  \textbf{[15]} \\ \textbf{[1.2]} \end{array}$	$\begin{array}{c} 4.3 \; (\pm \; 0.5) \times 10^8  \textbf{[22]} \\ \textbf{[1.7]} \end{array}$

specificity of binding in polyamide **10** remains roughly comparable to that seen without the dye.

### Discussion

From the array of footprinting data, one can observe three trends: a minimum polyamide length required for high binding affinity, an energetic penalty associated with increased molecular weight in polyamide-dye conjugate, and a confirmation of the pairing rules originally established for 1:1 binding  $\beta$ -linked linear polyamides, including binding orientation and preferred binding stoichiometry.<sup>20</sup> For the 5'-AAGAAGAAG-3' tract of DNA, a minimum of seven monomers (present as Im, Py, or  $\beta$ ) facilitates reasonable polyamide binding affinity. The appendage of BODIPY FL for tracking nuclear localization invokes a moderate energetic penalty on the binding affinity with minimal change to binding specificity.

The original sequence specificity studies on linear  $\beta$ -linked polyamides observed a tendency for pyrrole to pair adjacent to A·T or T·A base pairs and for imidazole to have no discernible preference for any of the base pairs.<sup>20</sup> While the current reported footprinting studies do not exhaustively place A·T, T·A, C·G, and G·C base pairs over an imidazole, a pyrrole, or a  $\beta$ -alanine, they both confirm and relax some of the prior observations. Four core polyamide sequences (containing Im, Py, or  $\beta$  monomers) have been studied in this footprinting series—Im-Py- $\beta$ -Im-Py- $\beta$ -Im- $\beta$  (as per polyamides **1**, **6**, and **9–10**), Im-Py- $\beta$ -Im-Im- $\beta$ -Py- $\beta$  (as per polyamide **2**), Im-Py- $\beta$ -Py- $\beta$ -Im- $\beta$  (as per polyamide **7**), and Im-Py- $\beta$ -Im-Py- $\beta$ -Im-Py- $\beta$  (as per polyamide **8**)—and enable the further elucidation of pairing rules in polypurine tracts of DNA.

Examining the sequence specificity for pyrrole (polyamides 7 and 8, binding sites III–VI), one confirms a general preference for A·T base pairs over G·C. However, there is insufficient data to correlate the location of the pyrrole monomer within the polyamide to the level of sequence specificity observed. Polyamide 7 prefers to place an internal pyrrole monomer adjacent to A·T versus G·C with 88-fold specificity (binding site V versus binding

site III). Polyamide 8 places a pyrrole closer to the C-terminus adjacent to A·T versus G·C with 56-fold specificity (binding site VI versus binding site III). Placing an internally located pyrrole of polyamide 8 adjacent to G·C reduces binding affinity 32-fold (binding site IV versus binding site VI).

For imidazoles, there may be a slight preference for binding G·C base pairs over A·T base pairs in homopurine tracts of DNA. When placing a centrally located imidazole adjacent to an A·T base pair (polyamide 1 adjacent to site V), there appears to be no energetic penalty. A fourfold penalty exists when placing a more C-terminal imidazole adjacent to an A·T base pair (polyamide 1 adjacent to site VI). Polyamide 6 exhibits at least 240-fold specificity for G·C over A·T adjacent to an internal imidazole (binding site IV versus binding site VI). Interestingly, placing imidazole adjacent to an A·T base pair either internally (polyamide 8, binding site V) or more C-terminally (polyamide 7, binding site VI) may temper the energetic penalty of placing a pyrrole adjacent to a G·C base pair. It is also possible that this tempering effect results from altered DNA microstructure that better complements the polyamide sequence.

Of important consequence from these experiments is the elucidation of a series of molecules that may harbor biologically significant effects in Friedreich Ataxia cell culture. A threshold has been established for minimum polyamide length to promote reasonable binding affinities. Because of its similar binding properties, a polyamide with an ethylmorpholino tail may improve the biological efficacy of polyamide **1**. Two new single amino acid mismatch control compounds have been characterized (polyamides **7** and **8**) that harbor high affinities for a polypurine tract of DNA but exhibit high specificity (56-fold to 88-fold) for their match site over a 5'-AAGAAGAAG-3' binding site. These new compounds can begin to answer whether a high binding affinity linear  $\beta$ -linked polyamide alone can effect increased frataxin expression.

While polyamide **1** has been observed to bind four other DNA sites beyond 5'-AAGAAGAAG-3', an exhaustive quantitative DNase I footprinting study of 9-base

pair DNA sequences is impractical, necessitating 131,072 unique binding sites. The global sequence preferences of Polyamide 1 will be studied using cognate site identifier (CSI) microarrays.<sup>28</sup>

### **MATERIALS AND METHODS**

### Chemicals

Polyamides 1–10 were synthesized by solid-phase methods on Boc- $\beta$ -ala-PAM resin (Peptides International, Louisville, KY),<sup>29</sup> on Kaiser oxime resin (Nova Biochem, Laufelfingen, Switzerland),<sup>30</sup> or on Marshall-Liener resin (Nova Biochem, Laufelfingen, Switzerland).<sup>31</sup> Each step of the polyamide synthesis reaction was monitored by analytical HPLC. The polyamide-BODIPY FL conjugate was prepared as previously described.<sup>24</sup> The polyamide-EDTA conjugate (**1E**) was prepared essentially as previously described,<sup>32</sup> although dropwise addition of polyamide **9** in NMP over 1 hour to the vigorously stirred solution of EDTA-dianhydride, Hünig's base, and 1:1 DMSO:NMP was required. The

<i>Table 2.8.</i> MALDI-TOF Mass Spectral Data for Polyamide	<i>Table 2.8.</i>	MALDI-TOF	Mass Spectral	l Data f	or Pol	lyamide	es
--	-------------------	-----------	---------------	----------	--------	---------	----

Compound	Calc'd [M+H] <sup>+</sup>	Obsv'd [M+H] <sup>+</sup>	$\lambda_{\max(nm)}$	ε (L·mol <sup>-1</sup> ·cm <sup>-1</sup> )
1	914.45	914.55	289	43,125
<b>1E</b>	1231.57	1231.50	289	43,125
2	914.45	914.46	289	43,125
3	843.41	843.37	289	43,125
4	649.33	649.72	285	34,500
5	598.32	598.30	283	25,875
6	942.44	942.54	289	43,125
7	913.45	913.44	289	43,125
8	913.45	913.40	289	43,125
9	957.49	957.50	289	43,125
10	1231.60	1231.52	289	43,125

purity and identification of the final products were verified by analytical HPLC and MALDI-TOF MS. Mass spectral data and UV maxima, along with their estimated extinction coefficients are included for each compound in Table 2.8. *N*-methylpyrrolidinone, di-isopropylethylamine, 3-(Dimethylamino)-1-propylamine, 3,3'-diamino-*N*-methyldipropylamine, 4-(2-aminoethyl)morpholine, dimethylformamide, dichloromethane, dimethylsulfoxide, and EDTA-dianhydride were purchased from Sigma-Aldrich. Trifluoroacetic acid was purchased from Halocarbon. Acetic anhydride was from EMD Biosciences and methanol from Fluka. HBTU and Boc- $\beta$ -Ala-OH were purchased from Peptides International. The Boc- $\beta$ -Im-OH dimer was utilized in place of coupling  $\beta$ -alanine to *N*-methylimidazole, increasing the polyamide synthesis yield. BODIPY FL succinimidyl ester was purchased from Invitrogen. All commercially available reagents were used as received without additional purification.

### **Plasmid Preparation**

Plasmids were constructed by ligating (Rapid DNA Ligation Kit, Roche) the following hybridized inserts (Integrated DNA Technologies, Coralville, IA) into the *Bam*HI / *Hin*DIII (Roche Biosciences) polycloning site in pUC19 (Sigma):

*pJWP2*. 5'-GATCCGGCCAAGAAGAAGCCGGGGGCCAAAAGGAAGCCGGGG CCAAGAAAAAGCCGGGGGCCAAAAAGAAGCCGGAT-3'·5'-AGCTATCCGG CTTCTTTTTGGCCCCGGCTTTTTCTTGGCCCCGGCTTCCTTTTGGCCCCG-GCTTCTTCTTGGCCG-3'.

The ligated plasmid was then transformed into JM109 subcompetent cells (Promega).

Colonies were selected for a-complementation on 25 mL Luria-Bertani agar plates containing 50 mg/L ampicillin (Sigma), 120 mg/L IPTG (ICN Biomedicals), and 40 mg/L X-gal (Invitrogen) after overnight growth at 37 °C. Cells were harvested after 16 h growth at 37 °C in LB medium containing 50 mg/L ampicillin. Plasmid purification was performed using WizardPlus Midi Preps (Promega). The presence of the desired inserts was determined by capillary electrophoresis dideoxy sequencing methods.

### **Preparation of 5'-End-Labeled Fragments**

Two 21 base-pair primer oligonucleotides, 5'-AATTCGAGCTCGGTACCCGGG-3' (forward) and 5'-CTGGCACGACAGGTTTCCCGA-3' (reverse) were constructed for PCR amplification (Integrated DNA Technologies). The forward primer was radiolabeled using  $[\gamma^{-32}P]$ -dATP (MP Biomedicals) and polynucleotide kinase (Roche), followed by purification using ProbeQuant G-50 spin columns (GE Healthcare). The desired DNA segment was amplified as previously described.<sup>33</sup> The labeled fragment was loaded onto a 7% nondenaturing preparatory polyacrylamide gel (5% cross-link), and the desired 279 (pJWP2), 291 (pJWP5), or 248 (pJWP10) base-pair band was visualized by autoradiography and isolated. Chemical sequencing reactions were performed according to published protocols.<sup>33</sup>

# Quantitative DNase I Footprint Titrations, Affinity Cleavage, and Methidium Propyl EDTA (MPE) Footprinting Methods

All reactions were carried out in a volume of 400 mL according to published protocols.<sup>33</sup> Quantitation by storage phosphor autoradiography and determination of equilibrium association constants were as previously described.<sup>33</sup>

We thank the Caltech Sequence and Structure Analysis Facility for sequencing the insert regions of the plasmids. Funding was provided by the Friedreich Ataxia Research Alliance and the National Institutes of Health (GM27681).

### References

- Campuzano, V.; Montermini, L.; Molto, M. D.; Pianese, L.; Cossee, M.; Cavalcanti,
   F.; Monros, E.; Rodius, F.; Duclos, F.; Monticelli, A.; Zara, F.; Canizares, J.;
   Koutnikova, H.; Bidichandani, S. I.; Gellera, C.; Brice, A.; Trouillas, P.; DeMichele,
   G.; Filla, A.; DeFrutos, R.; Palau, F.; Patel, P. I.; DiDonato, S.; Mandel, J. L.;
   Cocozza, S.; Koenig, M.; Pandolfo, M. *Science* 1996, *271*, 1423–1427.
- (2) Sakamoto, N.; Chastain, P. D.; Parniewski, P.; Ohshima, K.; Pandolfo, M.; Griffith, J. D.; Wells, R. D. *Mol. Cell.* 1999, *3*, 465–475.
- (3) Vetcher, A. A.; Napierala, M.; Iyer, R. R.; Chastain, P. D.; Griffith, J. D.; Wells, R. D. J. Biol. Chem. 2002, 277, 39217–39227.
- (4) Vetcher, A. A.; Napierala, M.; Wells, R. D. J. Biol. Chem. 2002, 277, 39228– 39234.
- (5) Vetcher, A. A.; Wells, R. D. J. Biol. Chem. 2004, 279, 6434–6443.
- (6) Napierala, M.; Dere, R.; Vetcher, A.; Wells, R. D. J. Biol. Chem. 2004, 279, 6444–6454.
- Saveliev, A.; Everett, C.; Sharpe, T.; Webster, Z.; Festenstein, R. *Nature* 2003, 422, 909–913.
- Bidichandani, S. I.; Ashizawa, T.; Patel, P. I. Am. J. Hum. Genet. 1998, 62, 111–121.
- (9) Ohshima, K.; Montermini, L.; Wells, R. D.; Pandolfo, M. J. Biol. Chem. 1998, 273, 14588–14595.
- (10) Grabczyk, E.; Usdin, K. *Nucleic Acids Res.* **2000**, *28*, 2815–2822.
- (11) Janssen, S.; Durussel, T.; Laemmli, U. K. Mol. Cell 2000, 6, 999–1011.
- Burnett, R.; Melander, C.; Puckett, J. W.; Son, L. S.; Wells, R. D.; Dervan, P. B.;
  Gottesfeld, J. M. *Proc. Natl. Acad. Sci. U. S. A.* 2006, *103*, 11497–11502.
- (13) Dervan, P. B.; Edelson, B. S. Curr. Opin. Struct. Biol. 2003, 13, 284–299.
- (14) Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B. Bioorg. Med. Chem. 2003,

11, 4333–4340.

- (15) Edayathumangalam, R. S.; Weyermann, P.; Gottesfeld, J. M.; Dervan, P. B.; Luger,
   K. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101*, 6864–6869.
- (16) Doss, R. M.; Marques, M. A.; Foister, S.; Dervan, P. B. Chem. Biodivers. 2004, 1, 886–899.
- (17) Marques, M. A.; Doss, R. M.; Foister, S.; Dervan, P. B. J. Am. Chem. Soc. 2004, 126, 10339–10349.
- (18) Janssen, S.; Cuvier, O.; Muller, M.; Laemmli, U. K. Mol. Cell 2000, 6, 1013– 1024.
- (19) Urbach, A. R.; Love, J. J.; Ross, S. A.; Dervan, P. B. J. Mol. Biol. 2002, 320, 55–71.
- (20) Urbach, A. R.; Dervan, P. B. Proc. Natl. Acad. Sci. U. S. A. 2001, 98, 4343–4348.
- (21) Dervan, P. B.; Urbach, A. R. In Essays in Contemporary Chemistry: From Molecular Structure towards Biology; Quinkert, G., Kisakürek, M. V., Eds.; Verlag Helvetica Chimica Acta: Zürich, 2001, p 327–339.
- Miranda, C. J.; Santos, M. M.; Ohshima, K.; Smith, J.; Li, L. T.; Bunting, M.;
  Cossee, M.; Koenig, M.; Sequeiros, J.; Kaplan, J.; Pandolfo, M. *FEBS Lett.* 2002, 512, 291–297.
- (23) Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B. *Bioorg. Med. Chem.* 2002, *10*, 3313–3318.
- (24) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U. S.* A. 2003, 100, 12063–12068.
- (25) 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.
- (26) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* 2007, 35, 363–370.
- (27) Kaizerman, J. A.; Gross, M. L.; Ge, Y. G.; White, S.; Hu, W. H.; Duan, J. X.; Baird,

E. E.; Johnson, K. W.; Tanaka, R. D.; Moser, H. E.; Burli, R. W. J. Med. Chem. **2003**, *46*, 3914–3929.

- Warren, C. L.; Kratochvil, N. C. S.; Hauschild, K. E.; Foister, S.; Brezinski, M. L.;
  Dervan, P. B.; Phillips, G. N.; Ansari, A. Z. Proc. Natl. Acad. Sci. U. S. A. 2006, 103, 867–872.
- (29) Baird, E. E.; Dervan, P. B. J. Am. Chem. Soc. 1996, 118, 6141–6146.
- (30) Belitsky, J. M.; Nguyen, D. H.; Wurtz, N. R.; Dervan, P. B. *Bioorg. Med. Chem.*2002, 10, 2767–2774.
- (31) Stafford, R. L., Ph. D. Thesis. California Institute of Technology, 2008.
- (32) Herman, D. M.; Baird, E. E.; Dervan, P. B. J. Am. Chem. Soc. 1998, 120, 1382–1391.
- (33) Trauger, J. W.; Dervan, P. B. Methods Enzymol. 2001, 340, 450–466.