Chemistry & Biology, Vol. 10, 859–867, September, 2003, ©2003 Elsevier Science Ltd. All rights reserved. DOI 10.1016/j.chembiol.2003.09.001

Accessibility of Nuclear Chromatin by DNA Binding Polyamides

Brigitte Dudouet,[1,4] Ryan Burnett,[1] Liliane A. Dickinson,[1] Malcolm R. Wood,[1] Christian Melander,[1,2] Jason M. Belitsky,[2] Benjamin Edelson,[2] Nicholas Wurtz,[2] Christoph Briehn,[2] Peter B. Dervan,[2,3] and Joel M. Gottesfeld[1,3]

[1]Department of Molecular Biology
The Scripps Research Institute
La Jolla, California 92037

[2]Division of Chemistry and Chemical Engineering
California Institute of Technology
Pasadena, California 91125

Summary

Pyrrole-imidazole polyamides bind DNA with affinities comparable to those of transcriptional regulatory proteins and inhibit the DNA binding activities of components of the transcription apparatus. If polyamides are to be useful for the regulation of gene expression in cell culture experiments, one pivotal issue is accessibility of specific sites in nuclear chromatin. We first determined the kinetics of uptake and subcellular distribution of polyamides in lymphoid and myeloid cells using fluorescent polyamide-bodipy conjugates and deconvolution microscopy. Then cells were incubated with a polyamide-chlorambucil conjugate, and the sites of specific DNA cleavage in the nuclear chromatin were assayed by ligation-mediated PCR. In addition, DNA microarray analysis revealed that two different polyamides generated distinct transcription profiles. Remarkably, the polyamides affected only a limited number of genes.

Introduction

The development of chemical approaches for the regulation of gene expression in cell culture requires that sequence-specific DNA binding small molecules be cell permeable, transit to the nucleus, and access specific target sites in chromatin [1, 2]. Cell permeability and nuclear localization of pyrrole-imidazole polyamides have been demonstrated for several cell lines in culture [3, 4]. Although several gene regulation studies using polyamides in cell culture (and even in *Drosophila*) provide indirect evidence for access to nuclear chromatin [5–10], direct evidence that polyamides target sequences in nuclear DNA has been lacking. Polyamides bind specific DNA sites in a model nucleosome substrate with little loss in affinity and specificity [11, 12]. However, it is unclear how the sequence-specific binding of polyamides is affected by both higher-order chromatin structures and the large excess of competing genomic DNA sites in the cell nucleus. In the present study, we address several key issues for the use of polyamides as small molecule regulators of gene expression. Deconvolution microscopy is used to monitor subcellular localization, kinetics of uptake, and apoptotic effects of polyamides in various cell lines. This allows us to set parameters for time of incubation and concentration ranges for subsequent cell culture experiments. We assess sequence-specific DNA occupancy in nuclear chromatin by use of a polyamide-alkylator conjugate. In addition, the effects of polyamides on genomic transcription are examined in lymphoid cells by DNA microarray analysis.

[3]Correspondence: joelg@scripps.edu (J.M.G.), dervan@caltech.edu (P.B.D.)

[4]Present address: Sidney Kimmel Cancer Center, La Jolla, California 92037.

Results

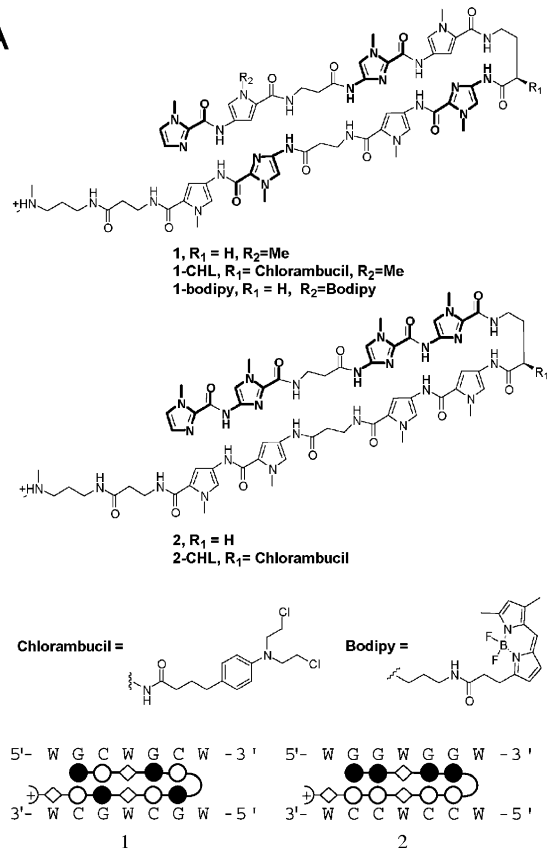
Polyamides and Polyamide Conjugates

Hairpin polyamide **1** and its chlorambucil (**1**-CHL) and bodipy (**1**-bodipy) conjugates used in this study have been described previously [3, 6, 13, 14]. Based on the pairing rules for DNA recognition, polyamide **1** binds sequences of the type 5′-WGCWGCW-3′ (where W = A or T), and polyamide **2** binds sequences of the type 5′-WGGWGGW-3′ (Figure 1A). A K_d of 0.1 nM was determined for polyamide **1** by quantitative DNase I footprinting [15] using a PCR fragment derived from the HIV-1 promoter and enhancer (Table 1 and Figure 1B). No sites for polyamide **2** are present in the region of this DNA that can be resolved by gel electrophoresis; however, a K_d of 0.2 nM was obtained for this polyamide on a different DNA fragment [15]. For the **1**-bodipy conjugate, a modest loss in binding affinity is observed [3] without loss in sequence specificity compared to the parent compound, whereas **1**-CHL binds its target sequences with no loss in affinity compared to the parent compound [13].

Subcellular Localization and Kinetics of Polyamide Uptake in Lymphoid Cells

The kinetics of polyamide **1**-bodipy uptake in lymphoid cells were examined by deconvolution fluorescence microscopy (Figure 2). Previous studies indicated that this polyamide-bodipy conjugate localizes in the nucleus of primary human lymphocytes and T cell lines. Similarly, we find that **1**-bodipy localizes in the nucleus of cultured myeloid and T cell lines including the myeloid-cell line KYO1 (Figure 2A) and the T cell lines CEM (Figure 2B), PM1 (Figure 2C), and MT2 (Figure 2D). In these experiments, cells were incubated in standard culture medium with the fluorescent **1**-bodipy conjugate for 16 hr prior to visualization. Immediately prior to microscopy, the cells were stained with MitoTracker (Molecular Probes) to visualize mitochondria in the cytoplasm. Note the presence of a dividing cell in Figure 2A with bright polyamide fluorescence staining of the chromosomes.

We next examined the time course of polyamide uptake in unfixed lymphoid cells, again using deconvolution microscopy. Cells were first stained with Mito-



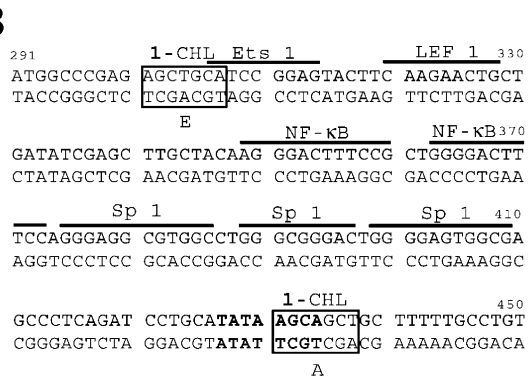


Figure 1. Structures and DNA Binding Models for Synthetic Polyamides

(A) Structures of the polyamides ImPy-β-ImPy-γ-ImPy-β-ImPy-β-Dp (**1**) and ImIm-β-ImIm-γ-PyPy-β-PyPy-β-Dp (**2**) (Im, imidazole; Py, pyrrole; γ, γ-aminobutyric acid; β, β-alanine; Dp, dimethylaminopropylamine). Imidazole rings are shown in bold. Polyamides were also synthesized with bodipy at the N-methyl position of the indicated pyrrole ring (yielding polyamide **1**-Bodipy) and with chlorambucil (CHL) at the α-position of the turn amino acid (yielding polyamide **1**-CHL and **2**-CHL). Polyamides are schematically represented at their respective DNA binding sites, where W = A or T. The solid and open circles represent Im and Py rings, respectively; the hairpin junction formed with γ-aminobutyric acid is shown as a curved line, and the diamond represents β-alanine. Base-sequence specificity depends on side-by-side pairing of Py and Im rings in the minor groove of DNA [1]. A pairing of an Im opposite a Py targets a G·C base pair, whereas a Py opposite an Im targets a C·G base pair. The Py/Py and β/β pairs are degenerate and target both A·T and T·A base pairs.

(B) Sequence of the HIV-1 LTR in isolate HXB2 [16] showing the location of the TATA box and binding sites for transcriptional activators Ets-1, LEF-1, NF-κB, and Sp1, along with the match binding sites for polyamide **1** and **1**-CHL (open boxes, sites A and E).

Table 1. Effect of Competing DNA on Polyamide 1 Binding Affinity

Competitor DNA (μg/reaction)	Polyamide Concentration Required for 50% Binding (nM)[a]
none	0.10 ± 0.02[b]
0.25 μg genomic DNA	1.04 (0.15)[c]
0.50 μg genomic DNA	2.48 (0.14)[c]
0.25 μg poly dG-dC	0.09 (0.01)[c]
2.5 μg poly dG-dC	0.22 (0.05)[c]

[a] Determined by DNase I footprinting and phosphorimage analysis for binding at site A (Figure 1B and data not shown).
[b] Mean and standard deviation for four determinations, with a range of 0.065 to 0.12 nM.
[c] Error values for each determination are given in parentheses.

Tracker, and then **1**-bodipy was added to a final concentration of 2 μM in the culture medium. Images were acquired every 2 min for 2 hr, then every 30 min for the next 10 hr. Nuclear localization in MT2 cells is evident after 1 hr, with intense nuclear staining after 2 hr. Video documentation of the experiment is available as Supplemental Data at http://www.chembiol.com/cgi/content/full/10/9/859/DC1. Note that some cells in the field do not appear to be stained with the polyamide: this is because the focal plane shown in this movie does not include the nuclei of all the cells in the microscopic field. In contrast, all of the nuclei are in the focal plane in the microscopic images shown in Figure 2, and all of the nuclei show bright polyamide fluorescence.

Polyamide Effects on Cell Viability, Growth, and Apoptosis

The stability of the bodipy-polyamide in lymphoid cells in culture was examined after 5-day incubations in standard culture media. Cells were incubated with polyamide, then stained with MitoTracker prior to deconvolution microscopy. Figure 2E demonstrates nuclear localization in MT2 cells under these conditions. Measurements of cell growth rates and viability (as measured by trypan blue exclusion) indicate that polyamides at 1–2 μM in standard culture medium have no deleterious effects on various cell lines, including CEM and MT2. We next examined whether polyamides induce apoptosis in MT2 and KYO1 cells. Cells were incubated in culture medium with either no polyamide or various concentrations of polyamide **1**-bodipy for 16 hr prior to staining with Alexa Fluor 594-annexin V. Annexin V binds to phosphotidyl serine on the surface of apoptotic cells, and such binding is an early marker of apoptosis. Polyamide **1**-bodipy at 2 μM (Figure 3B) does not induce annexin binding above the level observed in the absence of polyamide (Figure 3A), whereas polyamide concentrations above 10 μM induce apoptosis (Figures 3C and 3D). As a positive control for annexin V binding to apoptotic cells, cells

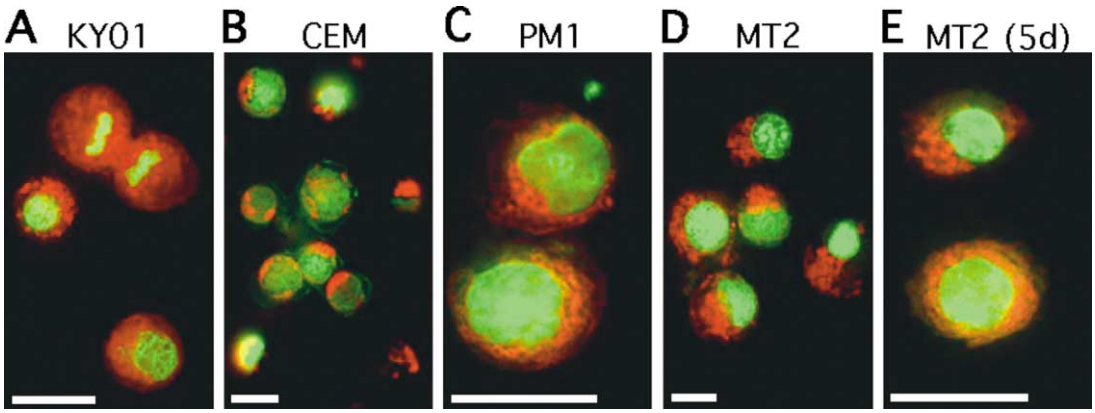


Figure 2. Deconvolution Microscopy of Unfixed Cells

Microscopic images of KYO1 myeloid cells (A), CEM (B), PM1 (C), and MT2 (D) lymphoid cells in culture incubated with bodipy-labeled polyamide (green) and MitoTracker Red 580 (red, mitochondrial dye). Cells were incubated with 2 μM polyamide for 16 hr prior to visualization. In (E), MT2 cells were incubated with 2 μM polyamide for 5 days prior to staining with MitoTracker and visualization by deconvolution microscopy. Images were acquired at different magnifications (ranging from 400 to 600×). The bar at the bottom of each panel corresponds to 10 microns.

were first incubated with 1-bodipy (2 μM) for 16 hr, then with 10 μM camptothecin for 4 hr prior to staining with annexin V (Figure 3E). Camptothecin induces apoptosis through generation of DNA breaks by inhibition of DNA topoisomerase I, and bright Alexa Fluor 594 fluorescence is observed under these conditions. Cells were also incubated with 1-bodipy (at 2 μM) for 48 hr prior to staining with annexin V (Figure 3F). Again, no evidence for apoptosis was obtained. Similar results were obtained with KYO1 cells (data not shown).

Effect of Competing Genomic DNA on Polyamide Binding Affinities

For polyamides to be useful for regulation of gene expression in cells, these molecules must be able to access their desired target sites in the context of genomic DNA where a large excess of competing sites will be present. To assess the effect of competing sites on sequence-specific binding, we performed footprint titrations in the presence of competing genomic DNA isolated from a human T cell line (line 5.25). This competition experiment also allows us to estimate the number of binding sites for this polyamide in human genomic DNA. A radiolabeled PCR fragment (Figure 1B) was mixed with unlabeled genomic DNA prior to addition of the polyamide and incubated for 18 hr prior to digestion with DNase I and gel analysis. We expect that the concentration of polyamide required to bind the radiolabeled DNA will increase in the presence of competing genomic DNA. In the absence of competing DNA, 50% binding is observed at a polyamide concentration of 0.1 nM (Table 1; representative footprints available in Supplemental Figure S1 at *Chemistry & Biology*'s website). In reactions containing 500 ng of competing DNA (a ~1700-fold mass excess of genomic DNA over the radiolabeled DNA), 50% occupancy of the radiolabeled DNA is observed at a polyamide concentration of 2.5 nM (Table 1). In the presence of 250 ng of competing DNA, 50% occupancy of the radiolabeled DNA is observed at a polyamide concentration of 1 nM. Based on

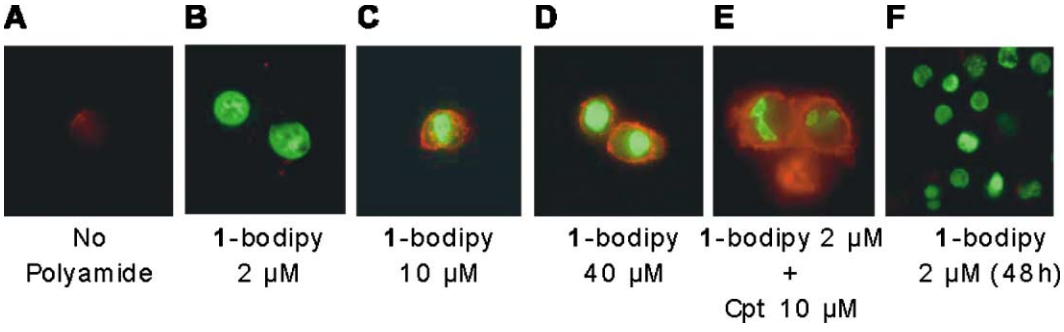


Figure 3. Effect of Polyamide Concentration on Cell Viability

Deconvoluted microscopic images of MT2 lymphoid cells in culture incubated with bodipy-labeled polyamide (green) and annexin V-Alexa Fluor 594 (red). Cells were incubated with polyamide 1-bodipy for 16 hr prior to staining with annexin V. Polyamide concentrations: no polyamide (A), 2 μM (B), 10 μM (C), and 40 μM (D). In (E), cells were incubated with 2 μM polyamide for 16 hr prior to adding 10 μM camptothecin to induce apoptosis. Four hours later, cells were stained with annexin V-Alexa Fluor 594 and visualized. In (F), cells were incubated for 48 hr with 2 μM bodipy-polyamide prior to annexin V-Alexa Fluor 594 staining and visualization.

the number of base pairs present in these amounts of competing genomic DNA and the concentration of polyamide required for 50% occupancy of the radiolabeled DNA (averaged from three experiments), calculation suggests that one match site for the polyamide occurs every ~1900 base pairs, or 1.3 million sites per haploid genome. If a sequence of the form 5′-WGCWGCW-3′ occurred at random in genomic DNA, we would expect such a sequence to be present once every 2048 bp ($= 4^4 \times 2^3$). The similarity between the calculated and experimentally determined frequencies of occurrence of binding sites in genomic DNA suggests that only match sites in genomic DNA are competing for binding of polyamide **1** to the radiolabeled DNA. Thus, the large excess of nontarget sites in genomic DNA does not affect polyamide binding to match sites. As a control for these experiments, we measured the binding affinity of the polyamide in the presence of either 250 ng or 2.5 μg of poly(dG-dC) nontarget DNA—that is, DNA that does not contain polyamide binding sites—providing either an ~850- or ~8500-fold mass excess over the radiolabeled DNA. The lower nontarget DNA concentration had no effect on the binding affinity of the polyamide, whereas the higher concentration caused only a modest 2-fold increase in the concentration of polyamide required for 50% binding compared to reactions lacking competitor DNA (Table 1), indicating that nonspecific DNA is not an effective competitor for sequence-specific polyamide binding.

Specific DNA Alkylation by a Polyamide-CHL Conjugate in Nuclear Chromatin

We used the polyamide-DNA alkylator conjugate **1**-CHL to assess polyamide binding in the cell nucleus. In these studies, the bis(dichloroethylamino)benzene moiety of CHL was covalently linked to a hairpin polyamide at the α-amino position of the turn amino acid (Figure 1A). Free CHL alkylates the N7 position of guanine in the major groove, but when attached to minor groove binding ligands, alkylation is observed at the nucleophilic N3 positions of adenine and guanine, at sites adjacent to the polyamide binding site [13, 14]. Sequence-specific alkylation of purified DNA with polyamide **1**-CHL has been observed at sites adjacent to the polyamide binding sites in the HIV-1 promoter and enhancer [13] and in SV40 DNA [14]. To determine whether the **1**-CHL conjugate binds its nuclear target sequences in cellular chromatin, we used a modified genomic DNA-footprinting technique, in which sites of DNA alkylation by the **1**-CHL conjugate are mapped by ligation-mediated PCR using HIV-1 LTR-specific primers and a CEM-derived T cell line (5.25) that contains a stably integrated copy of the HIV-1 LTR [16]. Specific alkylation of adenine or guanine residues adjacent to a polyamide binding site will demonstrate specific binding of the polyamide to this site. Cells were incubated with either polyamide **1**-CHL or polyamide **2**-CHL for 24 hr in standard culture medium. These incubations (at polyamide concentrations up to 2 μM) had no effect on cell viability, as measured by trypan blue exclusion, with approximately 95% of the cells remaining viable (data not shown). Genomic DNA was extracted from treated cells, cleaved at modified residues with piperidine, and subjected to ligation-mediated PCR (as described in Experimental Procedures). The PCR products were then subjected to electrophoresis on a denaturing polyacrylamide gel alongside a chemical sequencing reaction of unmodified DNA. Figure 4A shows that treatment of either purified genomic DNA ("DNA," lane 2) or intact cells ("cells," lane 4) with polyamide **1**-CHL results in alkylation of A_{308}, located immediately adjacent to the polyamide **1** binding site and overlapping the Ets-1 recognition site (site E; see Figure 1B). Polyamide **2**-CHL, a mismatch control, does not give rise to a specific band at this position (Figure 4A, lanes 3 and 5), indicating a lack of binding to this site. Additionally, alkylation with **1**-CHL is observed at A_{341}, corresponding to a single nucleotide mismatch site (Figure 4A). It is noteworthy that the band corresponding to A_{308} is fainter in the "cells" sample (compare lanes 2 and 4 in Figure 4A), indicating that access to this site by polyamide is hindered in the cell nucleus, possibly by Ets-1 or another ets-like protein bound to its recognition site. This finding is consistent with previous experiments in which inhibition of Ets-1 binding to a DNA fragment was observed only when polyamide was added to the DNA prior to Ets-1 protein, suggesting that the polyamide could not displace Ets-1 protein once bound to its recognition site [17]. Alternatively, chromatin structure might preclude polyamide binding at this site [11, 12]. Figure 4B shows alkylation at the TATA box by polyamide **1**-CHL both in purified genomic DNA (lane 2) and in cells (lane 4). The bands correspond to positions A_{428} and A_{430}, with A_{430} showing a slightly stronger intensity of alkylation, likely due to the proximity of A_{430} to the CHL moiety of the polyamide (Figure 4B). The extent of alkylation at this site is similar in both the in vitro and cell culture experiments. The mismatch polyamide **2**-CHL does not give rise to specific cleavage products over a region where no match sites for this polyamide are present in the DNA sequence (lanes 3 and 5). The specific alkylation sites detected here correspond to the same sites that were described previously with plasmid DNA containing the HIV-1 LTR [13] and with PCR products derived from the same cellular DNA sequence (see Supplemental Figure S2 at *Chemistry & Biology*'s website).

Because the time required for DNA extraction and LM-PCR spans several hours, it is conceivable that genomic DNA alkylation by the polyamide conjugate did not take place in the cell nucleus but occurred after cell lysis during the work-up of the sample. To assess this possibility, we added polyamide **1**-CHL to radiolabeled DNA simultaneously with cell lysis buffer and incubated these samples for 22 hr at ambient temperature prior to thermal cleavage and gel analysis (data not shown). DNA alkylation was not observed under these conditions. This finding suggests that alkylation of genomic DNA took place in the cell nucleus rather than after cell lysis. We conclude that polyamide **1**-CHL specifically binds to its target site in nuclear chromatin.

Effects of Polyamides on Genomic Transcription

The effects of polyamide treatment on nuclear transcription were monitored by DNA microarray analysis using Affymetrix high-density U133A arrays, which contain oli-

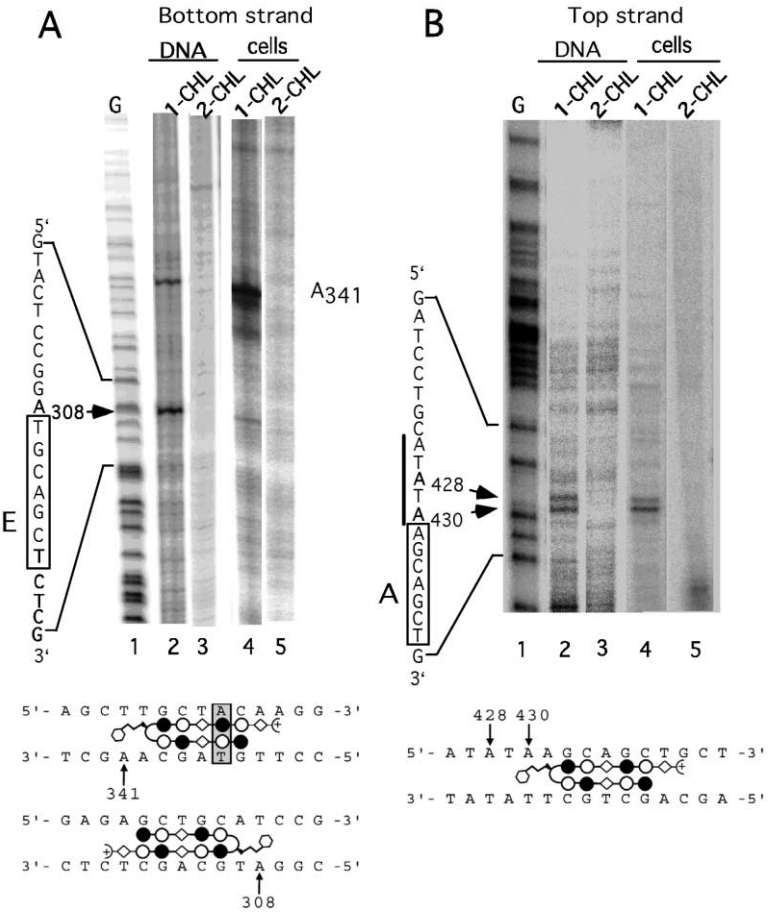


Figure 4. Site-Specific Alkylation by Polyamide **1**-CHL in Genomic DNA and in a Stably Transfected Cell Line

Phosphorimages showing the results of polyamide-CHL induced alkylation with intact cells ("cells") and with purified genomic DNA ("DNA"). The polyamide concentrations were 125 nM in (A) and 500 nM in (B), respectively. In (A), the upstream region of the HIV-1 LTR encompassing polyamide binding site E (Figure 1B) and a mismatch site is analyzed, whereas in (B) the downstream region of the LTR encompassing the TATA box and polyamide site A (Figure 1B) is analyzed. Polyamide binding sites are boxed. Alkylated residues are indicated with horizontal arrows alongside the gels. Nucleotide positions are relative to the start of the 5′ LTR in strain HXB2. Lane 1 shows a G-ladder (denoted G). The position of the TATA box is indicated with a solid vertical line. One match site for polyamide **2**-CHL is present in the HIV-1 LTR at nucleotide position 258 (Figure 1B); however, this site is distal to the region analyzed in this experiment (A), and hence no DNA alkylation by this polyamide is observed. At the bottom of each panel, the DNA sequences encompassing each of the polyamide binding sites are shown, along with the sites of alkylation, indicated with vertical arrows and nucleotide positions. Polyamide binding models are as in Figure 1A, with CHL denoted with a hexagon; the single nucleotide mismatch is boxed.

gonucleotides representing ~18,000 genes. MT2 cells were treated (in triplicate) with no polyamide, polyamide **1**, or polyamide **2** at a concentration of 2 μM in culture medium for 48 hr. These polyamides were chosen for this study because they contain the same number of Py and Im rings but bind different sequences. Hence, we expected that different sets of genes would be affected by the two polyamides. As before, no obvious cytotoxicity was observed, and the 48 hr time point was chosen such that the effects of the polyamides on new transcription could be assessed. Based on the half-lives of mammalian mRNAs, the majority of mRNA molecules that were present prior to polyamide treatment would be expected to turn over during this incubation period [18]. Total RNA was isolated and converted into fluorescent cRNA, which was hybridized to the oligonucleotide microarrays. Similar to the findings of Supekova et al. [10], the transcription levels of a surprisingly limited number of genes were affected by polyamide treatment (listed in Table 2). Using significance analysis of microarrays software (SAM) [19], 21 genes were identified as downregulated for cells treated with either polyamide **1** or **2**, with 10 common genes between the two data sets (at false discovery rates of 4.2% and 4.3% for polyamides **1** and **2**, respectively; Figure 5). The largest fold changes observed were for the genes encoding heat shock proteins hsp70(1B) and hsp70(6) (2.2- and 2.4-fold downregulated, respectively). This effect was validated by real-time RT-PCR of hsp70(1B) mRNA levels, where 2.7-fold downregulation was observed for each polyamide. Table 2 also lists the frequency of occurrence of match binding sites for the polyamides in 3 kb of DNA sequence upstream from the transcription start sites and 1 kb of sequence downstream. In each instance, match sites for the inhibitory polyamide are found in the 5′ flanking sequences of these genes. Remarkably, match sites for the noninhibitory polyamide are also found in most of the affected genes (Table 2). Ongoing efforts are aimed at assessing the mechanisms whereby these polyamides affect transcription in cell culture.

Discussion

Previous polyamide binding studies have utilized short (~200–400 bp) DNA fragments and equilibrium conditions to measure apparent association constants for match versus mismatch sites [1, 20]. For eight-ring hairpin polyamides, approximately 10-fold higher concentrations are required to bind single nucleotide mismatch sites compared to match sites. These differences in affinities for match versus mismatch sites are comparable to the binding specificities reported for natural transcription factors (reviewed in [20, 21]). We find that inclusion of genomic DNA in binding reactions increases the concentration of polyamide required to bind match sites in a radiolabeled DNA fragment by the statistically pre-

Table 2. Genechip Analysis of Expression Changes in MT2 Cells Incubated with **1** and **2**

Polyamide	Probe Set ID	Accession No.	Gene Description	Expression Ratio[a]		**1** match site WGCWGCW[b]		**2** match site WGGWGGW[b]	
						Upstream	Down	Upstream	Down
1	1 201296_s_at	NM_015626	SOCS box-containing WD protein SWiP-1 isoform 1	0.75		1	1	2	2
	2 202022_at	NM_005165	Aldolase C, fructose-bisphosphate	0.64		4	1	11	4
	3 202314_at	NM_000786	Homo sapiens cytochrome P450, 51 (lanosterol 14-alpha-demethylase)	0.81		0	4	3	0
	4 202364_at	NM_005962	MAX interacting protein 1	0.75		3	0	7	0
	5 202464_s_at	NM_004566	6-phosphofructo-2-kinase/fructose-2,6-biphosphatase 3	0.71		5	1	2	3
	6 209146_at	NM_006745	Sterol-C4-methyl oxidase-like	0.74		5	1	4	4
	7 209566_at	NM_016133	Homo sapiens insulin induced gene 2 (INSIG2)	0.66		0	0	*multiple match repeats*	
	8 218247_s_at	NM_016626	Hypothetical protein LOC51320	0.72		4	0	6	3
	9 219410_at	NM_018004	Hypothetical protein FLJ10134	0.66		3	1	1	1
	10 221479_s_at	NM_004331	BCL2/adenovirus E1B 19kD-interacting protein 3-like	0.56		5	1	2	1
	11 36711_at	AL021977	v-maf musculoaponeurotic fibrosarcoma oncogene	0.67		1	0	5	1
				1	**2**				
1 & 2	1 201849_at	NM_004052	BCL2/adenovirus E1B 19kD-interacting protein 3	0.56	0.57	0	1	3	2
	2 201968_s_at	NM_002633	Homo sapiens phosphoglucomutase 1 (PGM1)	0.75	0.76	1	1	3	4
	3 202581_at	NM_005346	Heat shock 70kD protein 1B	0.45	0.45	4	0	5	3
	4 202620_s_at	NM_000935	Procollagen-lysine, 2-oxoglutarate 5-dioxygenase (lysine hydroxylase) 2	0.61	0.71	1	3	3	3
	5 202887_s_at	NM_019058	Homo sapiens HIF-1 responsive RTP801	0.74	0.68	2	3	8	2
	6 212689_s_at	NM_018433	Homo sapiens zinc finger protein (TSGA)	0.70	0.74	2	0	1	1
	7 213418_at	NM_002155	Heat shock 70kD protein 6	0.41	0.43	3	0	7	2
	8 214567_s_at	NM_003175	Homo sapiens chemokine (C motif) ligand 2 (XCL2)	0.68	0.77	1	1	8	1
	9 218149_s_at	NM_017606	Hypothetical protein DKFZp434K1210	0.57	0.67	2	1	4	2
	10 221478_at	NM_004331	BCL2/adenovirus E1B 19kD-interacting protein 3-like	0.53	0.58	5	1	2	1
2	1 200664_s_at	NM_006145	DnaJ (Hsp40) homolog, subfamily B, member 1	0.61		1	0	6	1
	2 201848_s_at	NM_004052	BCL2/adenovirus E1B 19kD-interacting protein 3	0.64		0	1	3	2
	3 202336_s_at	NM_000919	Peptidylglycine alpha-amidating monooxygenase (PAM)	0.80		1	1	5	1
	4 202934_at	AI761561	Chromosome 2p12 near hexokinase II (Z46376)	0.74		*no searchable gene*			
	5 203946_s_at	NM_001172	Arginase, type II precursor	0.86		2	1	4	3
	6 204702_s_at	NM_004289	Nuclear factor (erythroid-derived 2)-like 3	0.88		1	5	7	1
	7 206140_at	NM_004789	LIM homeobox protein 2	0.75		0	0	10	2
	8 207543_s_at	NM_000917	Procollagen-proline, 2-oxoglutarate 4-dioxygenase (proline 4-hydroxylase)	0.78		1	3	3	1
	9 207574_s_at	NM_015675	Homo sapiens growth arrest and DNA-damage-inducible, beta (GADD45B)	0.80		1	0	9	2
	10 212399_s_at	XM_052386	KIAA0121 protein	0.85		4	1	2	3
	11 218750_at	NM_024116	Hypothetical protein MGC5306	0.84		2	1	3	3

[a] Genechip experiments were performed in triplicate. SAM analysis from cells incubated with polyamide **1** versus control and **2** versus control was adjusted to 0.89 false significant genes out of 21 identified for each condition. The expression ratio describes the expression of the significant gene in the treated sample compared to that of the control.

[b] Sequences were scanned 3 Kb upstream and 1 Kb downstream of transcription start site.

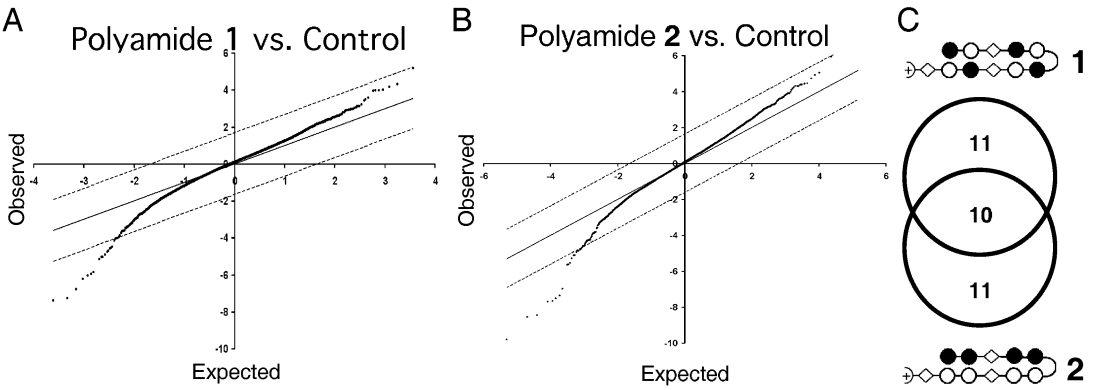


Figure 5. Significance Analysis of Microarray Data

(A and B) RNA from cells incubated for 48 hr with 2 μM polyamide **1** or **2** was used to probe Affymetrix U133A microarray chips. Resulting SAM analysis [19] reveals 21 significantly affected genes at a false discovery rate of 4.2% and 4.3% for **1** and **2** treated cells, respectively. Of the 21 genes, 10 were found to be common between both groups.

(C) Venn diagram of distribution of affected genes for cells treated with polyamides **1** and **2**. Models for polyamides structures are shown (as in Figure 1A).

dicted frequency of occurrence of match sites in genomic DNA. This suggests that only match sites in the genomic DNA compete for polyamide binding to the radiolabeled DNA, and the large excess of nontarget sites are without significant effect. Similar results have been reported for native and selected zinc finger proteins; binding affinities in the presence of competing genomic DNA suggest that these proteins have >20,000-fold higher affinities for match sites than random sequence DNA [21].

Using deconvolution microscopy, we demonstrate that a polyamide-bodipy conjugate is rapidly taken up in the nucleus of lymphoid and myeloid cells in culture [3]. Nuclear localization of polyamides in lymphoid cells is consistent with the inhibition of HIV-1 replication in primary peripheral blood mononuclear cells [6]. Assuming only match sites compete for polyamide binding in vivo, we can calculate the polyamide concentration required to saturate all potential sites within the genome in a eukaryotic cell nucleus. The frequency of occurrence of match sites for polyamide **1** in human DNA (one site every ~2 kb) corresponds to 5×10^{-18} moles of sites per nucleus or to a concentration of ~20 μM (based on a 7 micron diameter T cell nuclei, corresponding to a volume of 2×10^{-13} l). Results with the polyamide-bodipy conjugate (Figure 2 and data not shown) suggest that polyamides are concentrated by at least 20- to 40-fold in the nucleus over the input concentration in the culture medium. Thus, 0.5 to 1 μM polyamide should be sufficient to bind all potential match sites in living cells. It is unclear, however, whether all match sites are available for polyamide binding in the cell nucleus, because location of a polyamide binding site in nucleosomal DNA can determine polyamide accessibility in chromatin binding [11]. Additionally, the effects of higher-order chromatin structures on polyamide binding affinities remain unknown. In this regard, it is noteworthy that polyamide **1**-CHL was able to access only certain sites in cellular chromatin (Figure 4). Nevertheless, our ligation-mediated PCR mapping of alkylation by a polyamide-CHL conjugate in genomic DNA provides direct evidence for occupancy of a predicted target site in the cell nucleus by a polyamide at low micromolar to submicromolar concentration.

Significance

Our results provide evidence for sequence-specific targeting by hairpin polyamides in nuclear DNA of lymphoid cells. A polyamide-chlorambucil conjugate accesses a target site in the HIV-1 enhancer and promoter in the nuclei of cultured cells, as determined by ligation-mediated PCR. The polyamide concentrations required to observe promoter occupancy and to elicit transcriptional effects are below those concentrations that are toxic to the cell. Moreover, we find that the transcription profiles of only a limited number of genes are affected by polyamide treatment of cells in culture, despite the large number of potential binding sites for these polyamides in genomic DNA.

Experimental Procedures

Polyamides

Polyamides were synthesized by solid phase methods as described [22]. Polyamide-bodipy conjugates were synthesized by reacting the succinimidyl ester of Bodipy FL with a free amine of the precursor polyamide (at the N-position of a pyrrole ring, pointing out of the minor groove) [3] (Figure 1A). Polyamide-chlorambucil conjugates were also prepared as described [13]. Polyamides were characterized by matrix-assisted laser desorption ionization-time of flight mass spectrometry (MALDI-TOF MS), ^{1}H-NMR, and analytical high performance liquid chromatography (HPLC). Polyamides were purified by preparative HPLC, lyophilized, and stored at 4°C prior to being dissolved in a buffer containing 10 mM Tris-Cl (pH 7.6) and 20 mM NaCl. Concentrations were determined by UV absorbance using empirically determined extinction coefficients.

Cell Culture

The human T-lymphoblastic cell line CEM was purchased from the American Type Culture Collection (ATCC Number CCL-119); the human myeloid cell line KYO1 was the generous gift of Dr. P.K. Vogt (Scripps); the human T-lymphocyte (CD4$^+$) MT2 and PM1 cell lines were from Dr. D.E. Mosier (Scripps). Cells were cultured in RPMI 1640 medium containing 10% fetal bovine serum, 2 μM glutamine,

10 mM HEPES, 1 mM sodium pyruvate, 100 units penicillin, and 100 μg streptomycin sulfate per ml at 37°C in 5% CO_2-containing humidified air. The cell line 5.25 (kindly obtained from Dr. N. Landau, Salk Institute, La Jolla) was derived from a hybrid human CEM T cell and B-cell line and contained the HIV-1 LTR (from strain HXB2) stably integrated in the genome [16]. Cells were grown in RPMI-1640 medium supplemented with 10% fetal calf serum, 0.2 mg/ml G418, 0.5 μg/ml puromycin, and 1% penicillin/streptomycin/fungizone.

Deconvolution Microscopy

Cells were cultured at a density of 3×10^5 cells/ml in micro-dishes (Lab-Tek II, chamber #1.5 German coverglass system) for 24 hr prior to adding the bodipy-conjugated polyamide. Micro-dishes were precoated with poly-L-lysine (Sigma, P-8920) for 15 min at 4°C and washed twice in cold PBS prior to use. Cell viability was assayed by trypan blue exclusion. As a counterstain, mitochondria were labeled with MitoTracker Red 580 (M-22425, Molecular Probes). After incubation for 45 min, cells were washed twice in prewarmed growth medium and visualized with a wide-field deconvolution microscope (DeltaVision, API, Issaquah, WA). Images were documented using a Photometrics CH350L liquid cooled CCD camera attached to an Olympus IX-70 inverted microscope with a 60× oil immersion objective lens (NA 1.4) and specific filter sets. For the time-lapse study, images were initially collected every 2 min for 2 hr and subsequently every 30 min for 12 hr. All images were deconvolved using constrained iterative algorithms (10 iterations) of DeltaVision software (softWoRx, v2.5). This software was later used to prepare movies to depict the progressive changes in fluorescence labeling. Adobe Premier software was used to annotate the movie.

Apoptosis Assays

The effect of polyamide concentration on cell viability was assessed using annexin V-Alexa Fluor 594 conjugate (Molecular Probes). MT2 cells grown in poly-L-lysine precoated micro-dishes (at a density of 6×10^5 cells/ml) were incubated for various times in growth medium containing different bodipy-labeled polyamide concentrations. As a positive control for apoptosis, cells were treated with camptothecin (10 μM, Sigma) for 4 hr prior to the experiment. As a negative control, cells were incubated in growth medium without polyamide. After incubation, cells were washed twice in cold PBS. Five microliters of the annexin V conjugate per 100 μl of annexin binding buffer (Molecular Probes) were added, and the cells were incubated for 5 min at ambient temperature, washed in annexin binding buffer, and examined by deconvolution microscopy (as above). Healthy cells had either no staining or weak staining of the cellular membrane, whereas apoptotic cells exhibited a high degree of surface labeling.

DNA Templates for Footprinting and DNA Cleavage Reactions

For DNase I footprinting and alkylation reactions, singly end-labeled PCR products were derived using genomic DNA from the 5.25 cell line as template. Genomic DNA was extracted from cells using a Qiagen genomic extraction kit, and 100 ng of genomic DNA was used along with 100 ng of the primers listed below in PCR reactions using High Fidelity PCR Master (Roche). The upstream primer 5′-GCTTGTTACACCCTGTGAGCCTGCATGG-3′ (corresponding to nucleotide positions 199–227 of the HIV-1 HXB2 sequence in GenBank accession number K03455) and the downstream primer 5′-GCCA GAGAGCTCCCAGGCTCAGATCTG-3′ (corresponding to nucleotide positions 472–498) were both obtained from Sigma Genosys (The Woodlands, TX). To generate radiolabeled PCR products, primers were 5′ end labeled with T4 polynucleotide kinase and γ-^{32}P-ATP and used along with the respective unlabeled primer in separate PCR reactions. PCR products were separated from unincorporated primer and purified using a PCR extraction kit from Qiagen. The labeled 300 bp fragment was recovered from a nondenaturing polyacrylamide gel and purified by Elutip-D chromatography (Schleicher and Schuell). The identity of the PCR products was confirmed by Maxam-Gilbert chemical sequencing [23].

DNase I Footprinting and DNA Alkylation Reactions

Quantitative footprinting was performed as described [15, 24] with modifications noted below. The labeled DNA (~10,000 cpm or ~3 fmol) was incubated with polyamides for 18 to 22 hr at ambient temperature in a buffer containing 10 mM Tris-Cl (pH 7.6), 10 mM NaCl, 5 mM $MgCl_2$, and 2.5 mM $CaCl_2$ in a volume of 200 μl, yielding a DNA concentration of ~15 pM. In some experiments, genomic DNA (isolated from the 5.25 cell line) or poly(dG-dC) (Amersham Biosciences) was included in the incubations in a total volume of 100 μl. Digestion was allowed to proceed for 3 to 4 min at 23°C with 5×10^{-4} units of DNase I (Roche) diluted in 10 mM Tris-HCl (pH 8), 10 mM dithiothreitol. Reactions were stopped by the addition of 100 μl of a buffer consisting of 2 M NaCl, 10 mM Tris-Cl (pH 7.6), 25 mM EDTA. Samples were subjected to ethanol precipitation using 20 μg of glycogen per reaction as a carrier. For polyamide-CHL conjugates, sites of DNA alkylation were determined by incubating the conjugated polyamide with singly end-labeled PCR products for 22 hr at 37°C. Reactions were stopped by the addition of cell lysis buffer (Qiagen) followed by ethanol precipitation as described [13]. Samples were dissolved in 40 μl of 10 mM sodium citrate buffer (pH 7.2) and incubated for 30 min at 95°C followed by ethanol precipitation. Formic acid (0.3% for 25 min at 37°C) was used to generate A+G sequence markers, and dimethylsulfate (2% for 2 min at 23°C) was used for the G-only reaction [23]. Both footprinting and cleavage reactions were analyzed by electrophoresis on 6% sequencing polyacrylamide gel containing 8.3 M urea and 88 mM Tris-borate (pH 8.3), 2 mM EDTA. The dried gels were exposed to Kodak Bio-Max film at ambient temperature. Quantitation of the footprint titrations was by storage phosphor analysis utilizing Kodak Storage Phosphor Screens (SO 230) and a Molecular Dynamics SF PhosphorImager. The data were analyzed using ImageQuant software from Molecular Dynamics and fit to the Langmuir isotherm as described [13]. Nonlinear least square fits were performed with KaleidaGraph software and gave correlation coefficients greater than 0.98.

Ligation-Mediated PCR

Polyamides were added directly to approximately 2×10^7 5.25 cells and incubated in a CO_2 incubator at 37°C for 24 hr. Genomic DNA was extracted using a Qiagen genomic extraction kit. DNA (50 μg) was cleaved with 1 M piperidine in 100 μl for 30 min at 95°C, precipitated with ethanol twice, and used in ligation-mediated PCR with nested primers according to Garrity and Wold [25]. The upstream primer set (top strand) consisted of two primers spanning positions 150–177 and 199–227, respectively (numbering according to GenBank accession number K03455). The downstream primers (bottom strand) spanned positions 520–547 and 481–510. The linker sequence was as described [25]. The upstream primer set was used to map alkylation near the Ets-1 site (Figure 1B), and the downstream primer set was used to map alkylation at the TATA box. The LM-PCR products were visualized by inclusion of end-labeled primer in the last two rounds of the PCR. Purified genomic DNA from 5.25 cells was also incubated with polyamide-CHL, cleaved with piperidine, and analyzed by LM-PCR as described for the cell culture experiments. The G-ladder was obtained by treating unmodified genomic DNA (35 μg) with dimethylsulfate followed by piperidine cleavage according to standard protocols and LM-PCR.

Affymetrix Oligonucleotide Array

Total RNA from MT2 cells from triplicate experiments was isolated using a Qiagen RNeasy Midi Kit according to the manufacturer's instructions. Microarray experiments were performed at the DNA Array Core Facility of The Scripps Research Institute using Affymetrix Genechip Human Genome U133A chips. Genechip data were analyzed with Affymetrix MicroArray Suite (MAS 5.0) software. RMA values for probe sets were analyzed using Significance Analysis of Microarrays (SAM) 1.21.

Real-Time Quantitative RT-PCR

Real-time quantitative RT-PCR analysis was performed essentially as previously described [26]. However, the forward primer 5′-AGGC GGACAAGAAGAAGGTGC-3′ and the reverse primer 5′-TGGTAC AGTCCGCTGATGATGG-3′ were used to amplify a 144 bp fragment from the 3′ translated region of HSP (1B) (GenBank accession number NM 005346). RNA was standardized by quantification of the general housekeeping gene glyceraldehyde-3-phosphate dehydro-

genase (GAPDH) using the forward primer 5′-TGCACCACCAACTG CTTAGC-3′ and the reverse primer 5′-GGCATGGACTGTGGTCAT GAG-3′, as described [27]. Quantitative real-time RT-PCR was performed using Quantitect SYBR Green RT-PCR (Qiagen) under the following conditions: 1 cycle at 50°C for 30 min, 1 cycle at 95°C for 15 min, then 45 cycles of 94°C for 15 s, 60°C for 30 s, and 72°C for 30 s. Temperature cycling and detection of the SYBR Green emission were performed with a Cepheid SmartCycler II. Statistical analysis was performed on three independent quantitative RT-PCR experiments for each RNA sample.

Acknowledgments

This work was supported by National Institutes of Health grants CA84192 (to J.M.G.), GM57148 (to J.M.G. and P.B.D.), and GM51747 (P.B.D). B.E. was supported by a Howard Hughes Medical Institute Predoctoral Fellowship, and R.B. was supported by NIH training grant T32 AI07354. We wish to thank Steve Head and members of The Scripps Research Institute DNA core facility for help with microarray analysis.

Received: June 11, 2003
Revised: July 9, 2003
Accepted: July 24, 2003
Published: September 19, 2003

References

1. Dervan, P.B. (2001). Molecular recognition of DNA by small molecules. Bioorg. Med. Chem. *9*, 2215–2235.
2. Gottesfeld, J.M., Turner, J.M., and Dervan, P.B. (2000). Chemical approaches to control gene expression. Gene Expr. *9*, 77–91.
3. Belitsky, J.M., Leslie, S.J., Arora, P.S., Beerman, T.A., and Dervan, P.B. (2002). Cellular uptake of N-methylpyrrole/N-methylimidazole polyamide-dye conjugates. Bioorg. Med. Chem. *10*, 3313–3318.
4. Crowley, K.S., Phillion, D.P., Woodard, S.S., Schweitzer, B.A., Singh, M., Shabany, H., Burnette, B., Hippenmeyer, P., Heitmeier, M., and Bashkin, J.K. (2003). Controlling the intracellular localization of fluorescent polyamide analogues in cultured cells. Bioorg. Med. Chem. Lett. *13*, 1565–1570.
5. Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E., and Dervan, P.B. (1997). Regulation of gene expression by small molecules. Nature *387*, 202–205.
6. Dickinson, L.A., Gulizia, R.J., Trauger, J.W., Baird, E.E., Mosier, D.E., Gottesfeld, J.M., and Dervan, P.B. (1998). Inhibition of RNA polymerase II transcription in human cells by synthetic DNA-binding ligands. Proc. Natl. Acad. Sci. USA *95*, 12890–12895.
7. Janssen, S., Cuvier, O., Müller, M., and Laemmli, U.K. (2000). Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to *Drosophila melanogaster*. Mol. Cell *6*, 1013–1024.
8. Janssen, S., Durussel, T., and Laemmli, U.K. (2000). Chromatin opening of DNA satellites by targeted sequence-specific drugs. Mol. Cell *6*, 999–1011.
9. Coull, J.J., He, G., Melander, C., Rucker, V.C., Dervan, P.B., and Margolis, D.M. (2002). Targeted derepression of the human immunodeficiency virus type 1 long terminal repeat by pyrrole-imidazole polyamides. J. Virol. *76*, 12349–12354.
10. Supekova, L., Pezacki, J.P., Su, A.I., Loweth, C.J., Riedl, R., Geierstanger, B., Schultz, P.G., and Wemmer, D.E. (2002). Genomic effects of polyamide/DNA interactions on mRNA expression. Chem. Biol. *9*, 821–827.
11. Gottesfeld, J.M., Melander, C., Suto, R.K., Raviol, H., Luger, K., and Dervan, P.B. (2001). Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides. J. Mol. Biol. *309*, 615–629.
12. Suto, R.K., Edayathumangalam, R.S., White, C.L., Melander, C., Gottesfeld, J.M., Dervan, P.B., and Luger, K. (2003). Effects of ligand binding on structure and dynamics of the nucleosome core particle. J. Mol. Biol. *326*, 371–380.
13. Wurtz, N.R., and Dervan, P.B. (2000). Sequence specific alkylation of DNA by hairpin pyrrole-imidazole polyamide conjugates. Chem. Biol. *7*, 153–161.
14. Wang, Y.D., Dziegielewski, J., Wurtz, N.R., Dziegielewska, B., Dervan, P.B., and Beerman, T.A. (2003). DNA crosslinking and biological activity of a hairpin polyamide-chlorambucil conjugate. Nucleic Acids Res. *31*, 1208–1215.
15. Trauger, J.W., and Dervan, P.B. (2001). Footprinting methods for analysis of pyrrole-imidazole polyamide/DNA complexes. Methods Enzymol. *340*, 450–466.
16. Mariani, R., Rutter, G., Harris, M.E., Hope, T.J., Krausslich, H.G., and Landau, N.R. (2000). A block to human immunodeficiency virus type 1 assembly in murine cells. J. Virol. *74*, 3859–3870.
17. Dickinson, L.A., Trauger, J.W., Baird, E.E., Dervan, P.B., Graves, B.J., and Gottesfeld, J.M. (1999). Inhibition of Ets-1 DNA binding and ternary complex formation between Ets-1, NF-κB, and DNA by a designed DNA-binding ligand. J. Biol. Chem. *274*, 12765–12773.
18. Ross, J. (1995). mRNA stability in mammalian cells. Microbiol. Rev. *59*, 423–450.
19. Tusher, V.G., Tibshirani, R., and Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA *98*, 5116–5121.
20. Trauger, J.W., Baird, E.E., and Dervan, P.B. (1996). Subnanomolar sequence-specific recognition in the minor groove of DNA by designed ligands. Nature *382*, 559–561.
21. Greisman, H.A., and Pabo, C.O. (1997). A general strategy for selecting high-affinity zinc finger proteins for diverse DNA target sites. Science *275*, 657–661.
22. Baird, E.E., and Dervan, P.B. (1996). Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. J. Am. Chem. Soc. *118*, 6141–6146.
23. Maxam, A., and Gilbert, W. (1980). Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. *65*, 497–559.
24. White, S., Baird, E.E., and Dervan, P.B. (1997). On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. Chem. Biol. *4*, 569–578.
25. Garrity, P.A., and Wold, B.J. (1992). Effects of different DNA polymerases in ligation-mediated PCR: enhanced genomic sequencing and in vivo footprinting. Proc. Natl. Acad. Sci. USA *89*, 1021–1025.
26. Chuma, M., Sakamoto, M., Yamazi, K., Ohta, T., Ohki, M., Asaka, M., and Hirohashi, S. (2003). Expression profiling in multistage hepatocarcinogenesis: identification of HSP70 as a molecular marker of early hepatocellular carcinoma. Hepatology *37*, 198–207.
27. Pattyn, F., Speleman, F., De Paepe, A., and Vandesompele, J. (2003). RTPrimerDB: the real-time PCR primer and probe database. Nucleic Acids Res. *31*, 122–123.