Sequence specific alkylation of DNA by hairpin pyrrole–imidazole polyamide conjugates

Nicholas R Wurtz and Peter B Dervan

Background: Pyrrole–imidazole polyamides are synthetic ligands that recognize predetermined sequences in the minor groove of DNA with affinities and specificities comparable to those of DNA-binding proteins. As a result of their DNA-binding properties, polyamides could deliver reactive moieties for covalent reaction at specific DNA sequences and thereby inhibit DNA–protein interactions. Site-specific alkylation of DNA could be a useful tool for regulating gene expression. As a minimal first step, we set out to design and synthesize a class of hairpin polyamides equipped with DNA alkylating agents and characterize the specificity and yield of covalent modification.

Results: Bis(dichloroethylamino)benzene derivatives of the well-characterized chlorambucil (CHL) were attached to the γ turn of an eight-ring hairpin polyamide targeted to the HIV-1 promoter. We found that a hairpin polyamide–CHL conjugate binds and selectively alkylates predetermined sites in the HIV promoter at subnanomolar concentrations. Cleavage sites were determined on both strands of a restriction fragment containing the HIV-1 promoter, revealing good specificity and a high yield of alkylation.

Conclusions: The ability of polyamide–CHL conjugates to sequence specifically alkylate double-stranded DNA in high yield and at low concentrations sets the stage for testing their use as regulators of gene expression in cell culture and ultimately in complex organisms.

Address: Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology, Pasadena, CA 91125, USA.

Correspondence: Peter B Dervan
E-mail: dervan@its.caltech.edu

Key words: alkylation, DNA, minor groove, molecular recognition

Received: **14 December 1999**
Accepted: **6 January 2000**

Published: **14 February 2000**

Chemistry & Biology 2000, **7**:153–161

1074-5521/00/$ – see front matter

Introduction

Small molecules specifically targeted to any predetermined DNA sequence would be useful tools in molecular biology and potentially in human medicine. Polyamides containing pyrrole (Py), and imidazole (Im) amino acids are synthetic ligands that have an affinity and specificity for DNA comparable to naturally occurring DNA-binding proteins [1–5]. DNA recognition depends on side-by-side amino acid pairings in the minor groove. An antiparallel pairing of Im opposite Py (Im/Py) distinguishes G•C from C•G and both of these from A•T/T•A base pairs [6–9]. A Hp/Py pair specifies T•A from A•T and both of these from C•G/G•C [10,11]. A linker amino acid, γ-aminobutyric acid (γ), connects polyamide subunits in a carboxy-to-amino direction in a 'hairpin motif', and these ligands bind to predetermined target sites with > 100-fold enhanced affinity relative to unlinked dimers [1–5]. Paired β-alanine residues (β/β), restore the curvature of the dimer to enable recognition of larger binding sites and in addition, code for A•T/T•A base pairs [12–14]. Eight-ring hairpin polyamides can permeate eukaryotic cells and have been shown to modulate transcription by targeting promoters of specific genes [15,16].

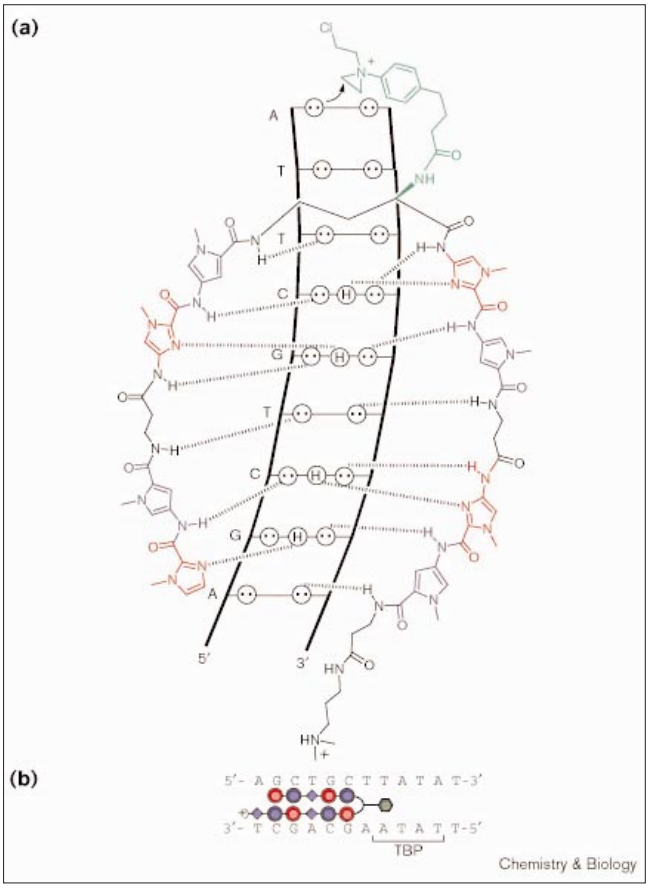
Sequence-specific covalent attachment in the DNA minor groove by small molecules

Hairpin polyamides do not appear to inhibit gene expression when bound to the coding region of genes (J.M. Gottesfeld and P.B.D., unpublished observations). In order to inhibit elongation by the RNA polymerase enzymes, it may be that one needs either a higher affinity class of DNA binders or alternatively bifunctional molecules that covalently attach to predetermined sequences in the minor groove of DNA. In this paper we address this latter class of molecules, the design of hairpin polyamides with separate domains for DNA binding and covalent attachment to DNA. This is a minimal first step towards creating new research tools for the field of functional genomics.

Hairpin polyamide-bis(dichloroethylamino)benzene conjugates

Nitrogen mustards are well characterized DNA alkylators with little sequence specificity [17]. A freely diffusing bis(dichloroethylamino)benzene moiety such as chlorambucil (CHL) reacts with the more nucleophilic N7 of guanine in the major groove, but will alkylate the N3 of adenine in the minor groove when attached to minor-groove DNA-binding ligands [18,19]. The combination

Figure 1



(a) Hydrogen-bonding model and alkylation mechanism of polyamide–CHL conjugate ImPy-β-ImPy-(*R*)^{CHL}γ-ImPy-β-ImPy-β-Dp (**2**) bound to the minor groove of 5′-AGCTGCT-3′. Circles with two dots represent the lone pairs of N3 purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogens of guanines. Putative hydrogen bonds are illustrated by dotted lines. Py and Im rings are represented as blue and red rings, respectively. The putative alkylating intermediate is green. **(b)** Model of polyamide conjugate **2** bound to the match site 5′-AGCTGCT-3′. Red and blue circles represent imidazole (Im) and pyrrole (Py) polyamide rings, respectively. Blue diamonds and green hexagons represent β alanine (β) and CHL, respectively. (*R*)-2,4-diaminobutyric acid ((R)γ) and dimethylamino-propylamide (Dp) are depicted as a curved line and a plus sign, respectively. TBP, TATA box binding region.

of subnanomolar binding hairpin polyamides and bis(dichloroethylamino)benzene derivatives creates a class of bifunctional agents that could potentially bind to predetermined DNA sequences with high affinity and specificity for subsequent covalent reactions at N3 of adenine. Key design issues are sites of attachment of the functional domains and the choice of linker length and flexibility. In order to target the adenines adjacent to polyamide-binding sites, we elected to attach bis(dichloroethylamino)benzene derivatives to the chiral α-amino group via a γ turn of a pyrrole/imidazole hairpin polyamide (Figure 1).

Polyamide **1** (Figure 2) has been shown to bind the sequence 5′-WGCWGCW-3′ (where W is A or T) adjacent to both sides of the TATA box in the HIV promoter (Figure 3) with high affinity ($K_a = 2.0 \times 10^{10}\,M^{-1}$) [15]. Substitutions at the α position of the γ turn using (*R*)-2,4-diaminobutyric acid potentially have only a modest effect on DNA binding affinity or specificity [20]. We chose this as the point of attachment for a nondiffusible electrophile that would react at the N3 of adenine adjacent to the polyamide-binding site. The reagent CHL, 4-[bis(2-chloroethyl)amino]benzene butanoic acid, contains a simple flexible linker and this proved adequate for the initial design–synthesis effort. Because alkylation by the electrophile results in irreversible binding, a nonalkylating conjugate **3** (Figure 2), in which the chlorides have been replaced with hydroxyl, was also synthesized as a control to investigate the energetic penalty (or lack thereof) on polyamide binding affinity and specificity for the target sequence. In addition, the CHL derivative **4** (Figure 2) will serve as a control to compare alkylation of the alkylating moiety unlinked to a polyamide.

Here, we report results with this first generation of hairpin polyamides that selectively alkylate and cleave DNA in the minor groove. Alkylation adducts and cleavage yields were determined using thermal cleavage assays on a 241 base pair HIV–LTR *Eco*RI–*Hind*III restriction fragment (Figure 3) and on 120-base pair synthetic oligonucleotides. DNA binding affinity and specificity were determined by DNase I footprinting of the nonreactive conjugate **3**.

Results and discussion

Synthesis of conjugates

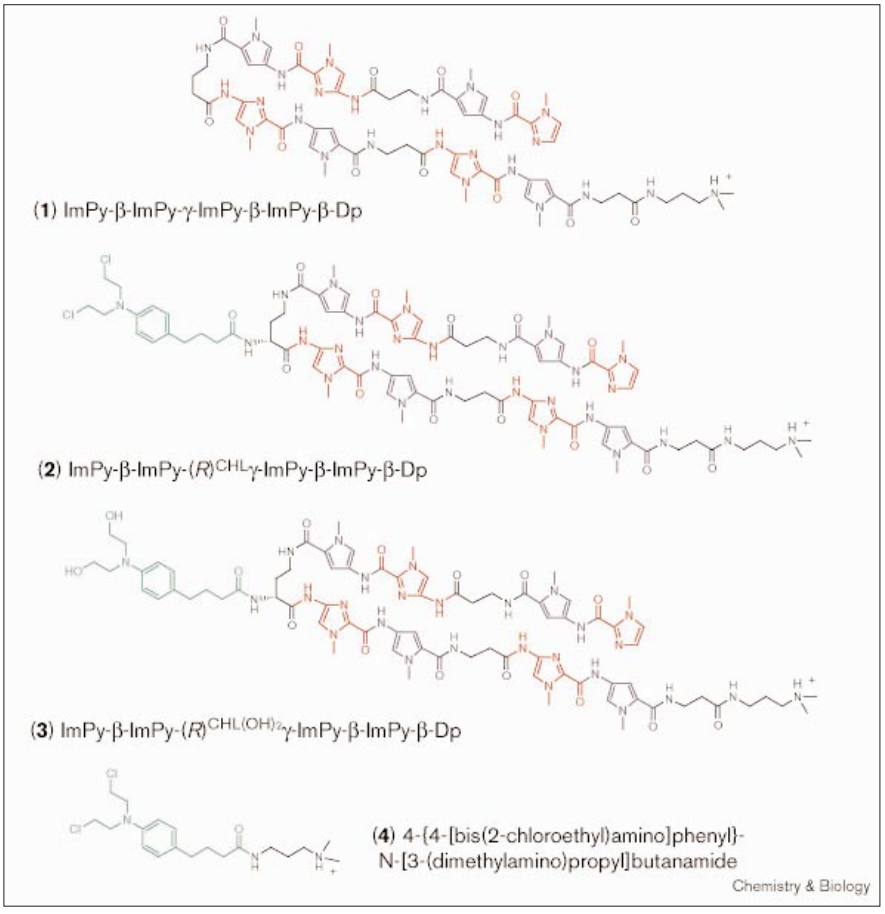
Polyamide **5**, ImPy-β-ImPy-(*R*)^{H_2N}γ-ImPy-β-ImPy-β-Dp, was prepared by manual solid-phase polyamide synthesis (Figure 4) [21]. After purification by reverse-phase high-performance liquid chromatography (HPLC), the appropriate activated carboxylate derivative of bis(dichloroethylamino)benzene was coupled to the α-amino group on the γ turn, using standard DCC/HOBt (dicyclohexylcarbodiimide/1-hydroxybenzotriazole hydrate) conditions, to yield polyamide–nitrogen mustard conjugate **2** (Figure 4). Control conjugate **3** was synthesized by allowing **2** to react with 0.1 M NaOH, followed by neutralization and lyophilization. Control compound **4** was synthesized by coupling CHL to (dimethylamino)-propylamine (Dp). All compounds were purified by reverse-phase HPLC. Matrix-assisted laser desorption/ionization time-of-flight (MALDI–TOF) mass spectrometry analysis of each compound was consistent with the calculated masses of the compounds.

Polyamide binding affinity and specificity is unaltered by substitution

Quantitative DNase I footprint titration experiments were performed on compound **3** to measure the equilibrium association constant for the match binding sites

Figure 2

Structures of polyamides **1**–**3** and a nonspecific alkylator **4**.

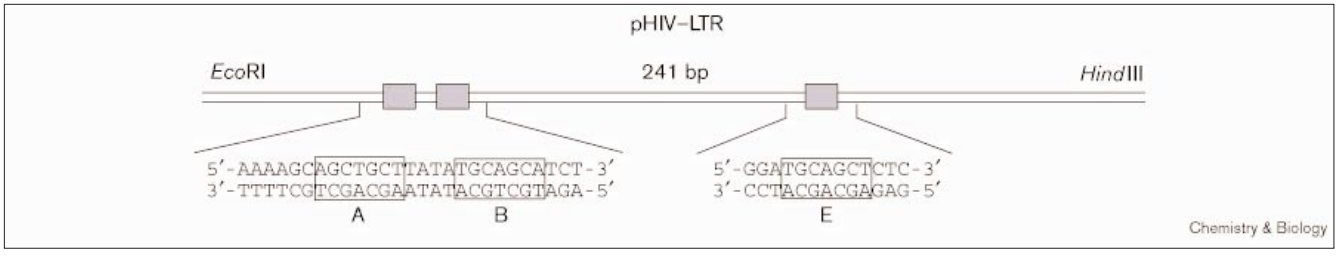


(5′-WGCWGCW-3′) on a 241 bp restriction fragment derived from the HIV-1 promoter region (Figure 5). The conjugate binds the match sites 5′-AGCTGCT-3′ and 5′-TGCAGCA-3′ (Figure 5b, sites A and B) with equilibrium association constants of $K_a = 1.6 \pm 0.7 \times 10^{10}\ M^{-1}$ and $1.3 \pm 0.7 \times 10^{10}\ M^{-1}$, respectively. This is remarkably similar to unsubstituted polyamide **1** ($K_a = 2.0 \times 10^{10}\ M^{-1}$) [16]. Affinities for the double base pair mismatch sites, 5′-TGTAGAA-3′ and 5′-AGCAGTC-3′ (Figure 6, sites C and D), were determined to be > 100-fold lower than those for the match sites. Attachment of the benzene moiety to the γ turn does not appear to affect the DNA binding affinity or specificity of the polyamide.

DNA alkylation by the polyamide conjugate proceeds in high yield

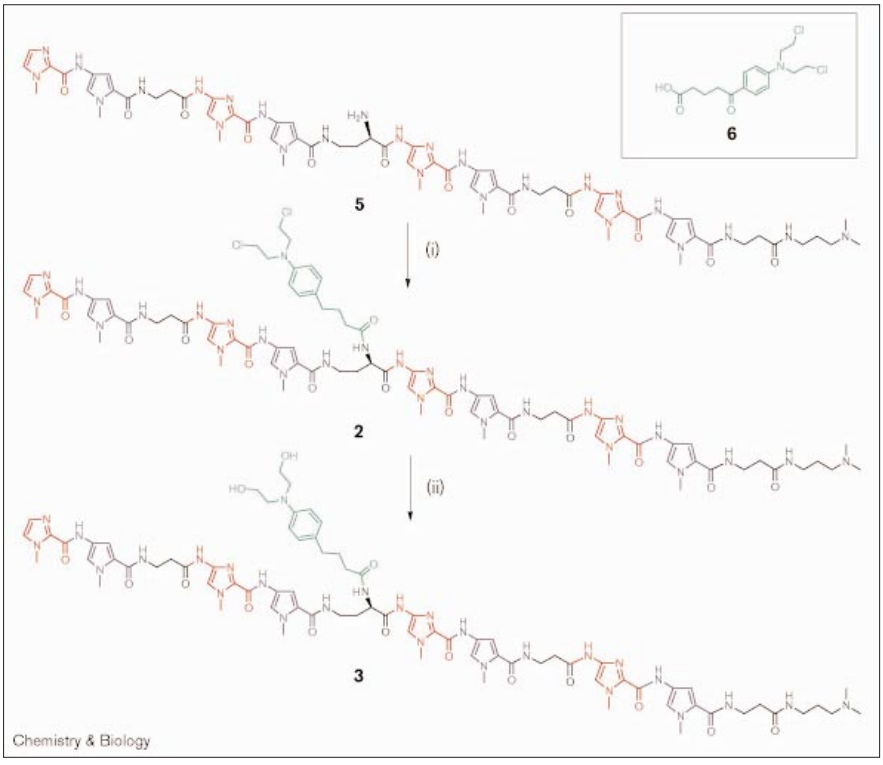
In order to measure covalent attachment to DNA, thermal cleavage assays were performed on 3′ and 5′

Figure 3



The 241 base pair pHIV–LTR *EcoRI*–*Hind*III restriction fragment with the position of the sequences indicated. The binding sites 5′-AGCTGCT-3′, 5′-TGCAGCA-3′ and 5′-TGCAGCT-3′ are highlighted in light purple.

Figure 4



Synthetic scheme for preparation of polyamide conjugate ImPy-β-ImPy-(*R*)^{CHL}γ-ImPy-β-ImPy-β-Dp (**2**). The HOBt-activated nitrogen mustard **6** is coupled with standard DCC/HOBt activation to polyamide **5** and purified by reverse-phase HPLC (step i). The hydrolyzed conjugate **3** is prepared by addition of 0.1 M NaOH (step ii).

^{32}P-labeled 241-base pair restriction fragments containing the HIV-1 promoter (Figure 6) [22]. Alkylation in high yield was observed at subnanomolar concentrations for conjugate **2**. In fact quantitative cleavage by conjugate **2** at 10 nM was observed on one strand (Figure 6). In contrast, no cleavage was observed at 10 μM of the freely diffusing CHL–(dimethylamino) propylamine derivative **4**.

The polyamide conjugate specifically alkylates adenines adjacent to binding sites

Upon heat-induced cleavage, all alkylation sites observed for conjugate **2** at subnanomolar concentrations were adjacent to target polyamide-binding sites (Figure 7). Because of the symmetry of the 5′-WGCWGCW-3′ binding site, the polyamide binds in two orientations, alkylating adenine residues adjacent on either side of the polyamide-binding site (Figure 7b). At concentrations greater than 1 nM, minor cleavage was observed proximal to double base pair mismatch sites as well. Specificity for the reaction at match sites over mismatch sites is ~20-fold.

In the thermal-cleavage assays, alkylation at all observed sites occurred at adenine, with one interesting exception, namely cleavage of guanine at one of the mismatch sites. Reaction at N3 of guanines has previously been observed with tallimustine, a Py_3 conjugate with a bis(dichloroethylamino)benzene derivative [23,24].

Alkylation of the exocyclic amine of guanine by polyamide conjugates

When the thermal-cleavage assays were carried out on the 5′-end labeled restriction fragment, a slower mobility fragment appeared in the gel electrophoresis assay above the intact DNA, indicating the presence of a stable adduct retarding the full-length fragment (see Figure 6a). The stability of the adduct to heat and piperidine workups suggested that the adduct was the result of alkylation of the exocyclic amine of guanine in the minor groove. To test this, inosine substitution at that position was employed.

Inosine, which lacks the exocyclic NH_2 of guanosine, was incorporated into a 120-base-pair fragment at the putative guanine reaction site located adjacent to the match site. The strands with inosine substitutions, as well as a strand with no inosine substitutions, were 5′-end labeled and annealed to a complementary strand and thermal cleavage assays were performed (Figure 8). The results of the thermal cleavage assay of the 5′-labeled synthetic oligonucleotides without inosine substitutions were identical to alkylation of the restriction fragment (Figure 8a). Substitution of the guanosine in the sequence 5′-A*G*CAGCT-GCT-3′ with inosine eliminated the anomalous band which results from the stable adduct (Figure 8b). Thus, certain DNA sequences allow the covalent reaction of the exocyclic amine of guanine with the polyamide–bis(dichloroethylamino)benzene conjugate.

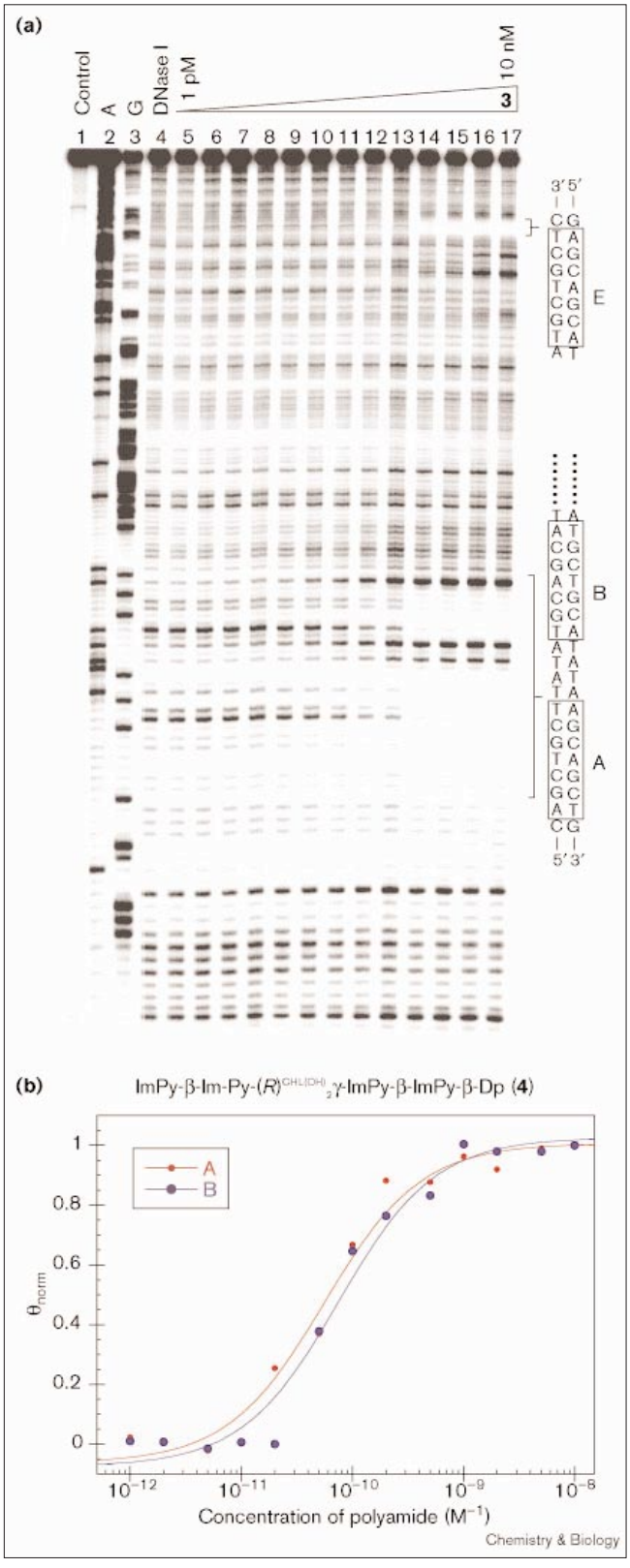


Figure 5

(a) Quantitative DNase I footprint titration experiments with ImPy-β-ImPy-(*R*)$^{CHL(OH)_2}$γ-ImPy-β-ImPy-β-Dp (**3**) on the 3′-^{32}P-end-labeled 241 base pair *Eco*RI–*Hind*III restriction fragment from plasmid pHIV–LTR [27]. Storage phosphor autoradiograms of 8% denaturing polyacrylamide gels used to separate the fragments generated by DNase I digestion. All reactions contained 20,000 cpm restriction fragment, 10 mM Tris–HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$ and were performed at 22°C. Lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lane 4, DNase I standard; lanes 5–17, 1 pM, 2 pM, 5 pM, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM and 10 nM, respectively. The 5′-AGCTGCT-3′ (A), 5′-TGCAGCA-3′ (B), and 5′-TGCTGCT-3′ (E) match sites are shown in boxes on the right-hand side of the autoradiogram. **(b)** Data from quantitative DNase I footprint titration experiments for ImPy-β-ImPy-(*R*)$^{CHL(OH)_2}$γ-ImPy-β-ImPy-β-Dp (**3**) binding to the 5′-AGCTGCT-3′ (A) and 5′-TGCAGCA-3′ sites (B). θ_{norm} points were obtained using storage phosphor autoradiography and processed as described in the Materials and methods section. The data for the binding of **3** to 5′-AGCTGCT-3′ is indicated by red circles and binding to 5′-TGCAGCA-3′ by larger blue circles. The solid curves are best-fit Langmuir binding titration isotherms obtained from a nonlinear least-squares algorithm equation 2 (n = 1; see the Materials and methods section).

Time dependence of alkylation

The half-life for hydrolysis of the bis(dichloroethylamino)benzene moiety of CHL in water at pH 7.5 and 37°C is ~1.3 h [25]. One would anticipate that at very low concentrations of polyamide conjugates there will be a competition between inactivation of the conjugate by hydrolysis and binding/covalent reaction with DNA. The time dependence of alkylation reactions with the polyamide conjugate **2** at 0.5 nM were therefore monitored for 24 h (pH 7.0, 37°C; Figure 9). Under these conditions, alkylation was half completed at match sites in 2.2 h and the final cleavage yield on the bottom strand was 45%. Cleavage was detectable as early as 10 min and the yield did not increase after 24 h.

Significance

Hairpin polyamides are a class of synthetic ligands that recognize specific DNA sequences with affinities comparable to those of DNA-binding proteins. Because of this property, we were interested in determining whether hairpin polyamides could be used to deliver reactive moieties for covalent reaction at specific DNA sequences [26]. DNA–protein interactions could be disrupted if the target DNA sequence has been alkylated, for example. We wanted to synthesize DNA alkylating agents attached to hairpin polyamides, and then characterize the specificity and yield of covalent modification.

At subnanomolar concentrations the hairpin polyamide delivers the electrophilic bis(dichloroethylamino)benzene to the predetermined sequence and produces a specific covalent reaction in high yield in the minor groove of DNA. The attachment of the chlorambucil moiety at the γ turn results in no significant alteration of the DNA-binding affinity of the polyamide. The parent eight-ring hairpin polyamide is known to be cell permeable, and the next technical hurdle is to determine whether eukaryotic cells will be permeable to the new polyamide–bis(dichloro-

ethylamino)benzene conjugates. The parent hairpin polyamide 1 has been shown to inhibit HIV replication in cell culture, and this will form the basis for our next experiment. Conjugate 2 targeted to the same HIV-1 promoter can now be tested in cell culture, to determine its ability to inhibit transcription. If all goes according to plan, this should set the stage for site specific alkylation in the 'coding region' of genes in order to inhibit transcription during the elongation phase. Sequence-specific DNA cleavage by multiple polyamide conjugates could provide a basis for 'genetic microsurgery', whereby undesired gene segments or integrated viral DNAs could be selectively removed from a host genome *in vivo*.

Materials and methods

Materials

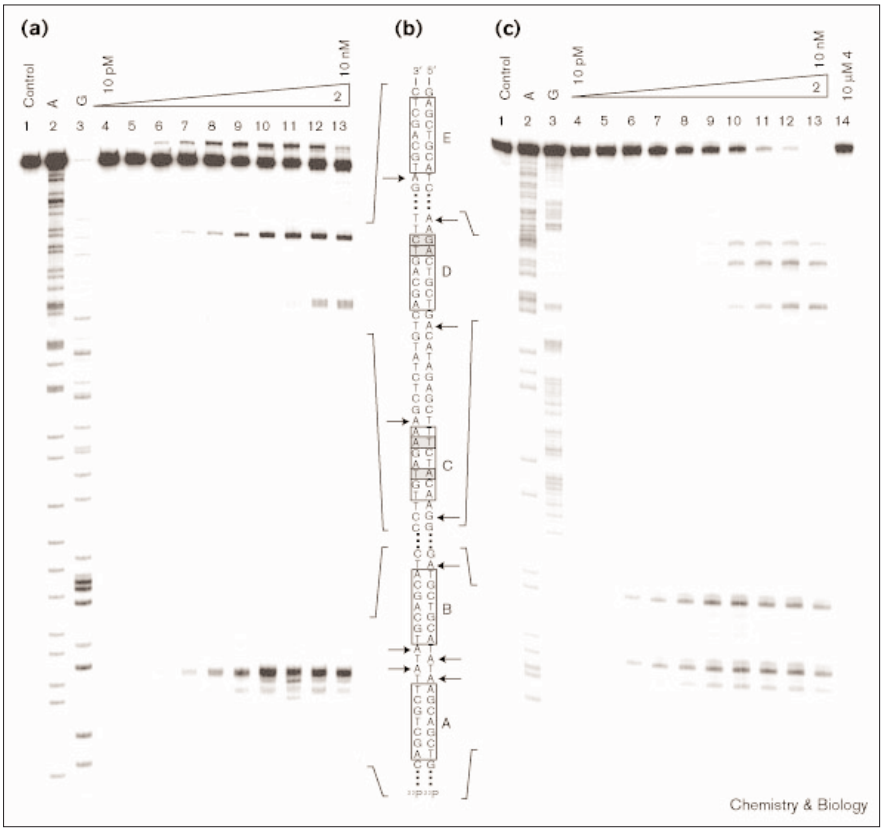
^{1}H NMR spectra were recorded on a General Electric-QE NMR spectrometer at 300 MHz with chemical shifts reported in parts per million relative to residual solvent. UV spectra were measured in water on a Hewlett-Packard Model 8452A diode array spectrometer. MALDI–TOF mass spectrometry was performed at the Protein and Peptide Microanalytical Facility of the California Institute of Technology. HPLC analysis was performed using a Beckman Gold Nouveau system using a Rainen C18, Microsorb MV, 5 μm, 300 × 4.6 mm reversed phase column on 0.1% (wt/v) TFA with acetonitrile as eluent and a flow rate of 1.0 ml/min, gradient elution 1.25% acetonitrile/min. Preparatory reverse phase HPLC was performed on a Beckman HPLC with a Waters DeltaPak 25 × 100 mm, 100 μm C18 column equipped with a guard, 0.1% (wt/v) TFA, 0.25% acetonitrile/min. Water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2 μm filtered. Reagent-grade chemicals were used unless otherwise stated. Oligonucleotides were synthesized by the Caltech Biopolymer Synthesis and Analysis Resource Center. Restriction enzymes were purchased from either New England Biolabs or Boeringer-Mannhein and used accorded to the manufacturer's protocols. [γ-^{32}P]-Thymidine-5′-triphosphate (≥ 3000 Ci/ mmol) and [γ-^{32}P]-deoxyadenosine-5′-triphosphate (≥ 6000 Ci/mmol) were purchased from DuPont/NEN. [γ-^{32}P]-Adenosine-5′-triphosphate (≥ 7000 Ci/ mmol) was obtained from ICN.

Synthesis of polyamide–bis(dichloroethylamino)benzene conjugate

*ImPy-β-ImPy-(*R*)^{H_2N}γ-ImPy-β-ImPy-β-Dp (**5**).* ImPy-β-ImPy-(R)^{H_2N}γ-ImPy-β-ImPy-β-Pam-resin was synthesized in a stepwise fashion by Boc-chemistry manual solid-phase protocols [20,21]. A sample of the resin was treated with neat (dimethylamino)-propylamine (2 ml), heated (55°C, 48 h) and purified by reversed-phase HPLC. ImPy-β-ImPy-(R)^{H_2N}γ-ImPy-β-ImPy-β-Dp was recovered as a white powder upon lyophilization of the appropriate fraction (18 mg, 10.4% recovery). UV (H_2O) λ_{max} (ε), 306 (69,500); ^{1}H NMR (DMSO-d_6): δ 10.97 (s, 1 H), 10.39 (s, 1 H), 10.25 (s, 1 H), 10.00 (s, 1 H), 9.97 (s, 1 H), 9.90 (s, 1 H), 9.24 (bs, 1 H), 8.15 (m, 5 H), 8.03 (bd, 3 H, J = 6.1 Hz), 7.47 (s, 1 H), 7.42 (s, 1 H), 7.40 (s, 1 H), 7.35 (s, 1 H), 7.18 (m, 4 H), 7.01 (s, 1 H), 6.97 (s, 1 H), 6.92 (s, 1 H), 6.90 (m, 2 H), 3.91 (m, 13 H), 3.75 (s, 12 H), 3.33 (m, 6 H), 3.24 (m, 2 H), 3.06 (q, 2 H, J = 6.0), 2.95 (quintet, 2 H, J = 5.0), 2.54 (t, 4 H, J = 6.8 Hz), 2.30 (t, 2 H, J = 7.0), 1.94 (m, 2 H), 1.68 (quintet, 2 H, J = 7.3 Hz); MALDI–TOF MS (monoisotopic) [M + H] 1381.6 (1381.6 calc'd for $C_{62}H_{81}N_{26}O_{12}$).

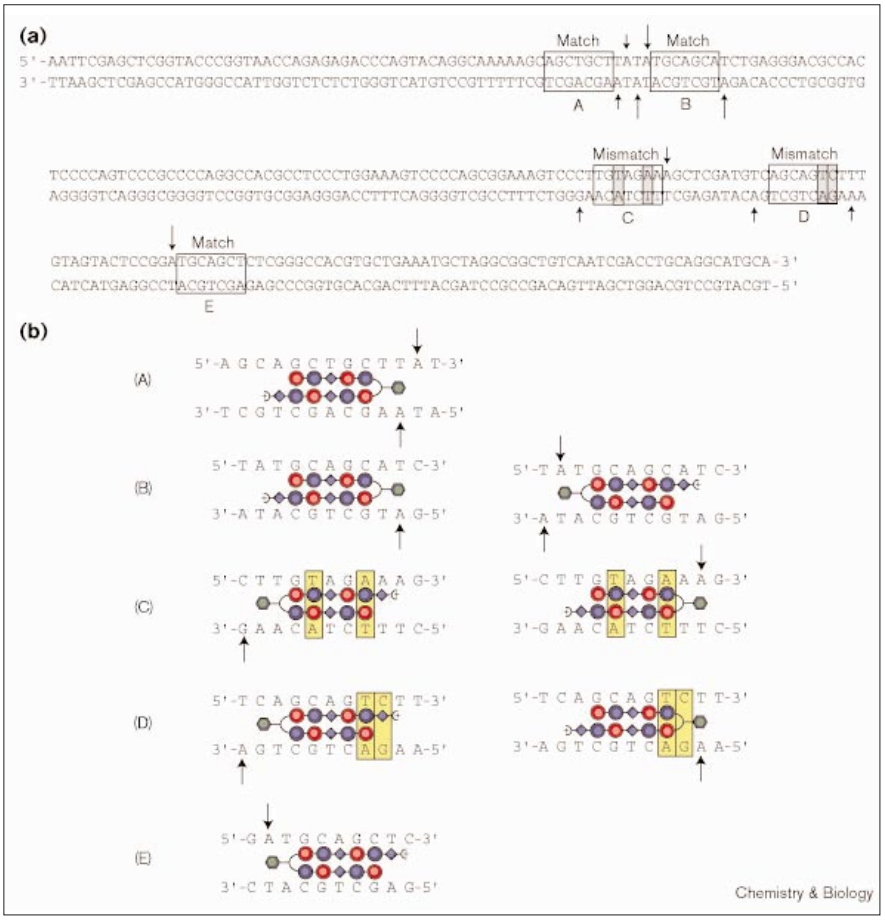
*ImPy-β-ImPy-(*R*)^{CHL}γ-ImPy-β-ImPy-β-Dp (**2**).* To a solution of CHL (5.7 mg, Fluka) in 20 μl DMF was added DCC (7.7 mg) and

Figure 6



Thermal cleavage assay experiments with ImPy-β-ImPy-(*R*)^{CHL}γ-ImPy-β-ImPy-β-Dp (**4**) on the 5′and 3′-^{32}P-end-labeled 241 base pair *Eco*RI–*Hind*III restriction fragment from plasmid pHIV–LTR [27]. Storage phosphor autoradiograms of 8% denaturing polyacrylamide gels used to separate the fragments generated by heat-induced DNA cleavage at alkylation sites. All reactions contained 10,000 cpm restriction fragment, 10 mM Tris–HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$ and were performed at 37°C. Following 24 h of equilibration, the DNA pellet was resuspended in sodium citrate buffer (pH 7.2) and heated to 90°C for 15 min to cleave thermally at sites of adenine- or guanine-N3 lesions. **(a)** 5′-^{32}P-end-labeled restriction fragment. **(c)** 3′-^{32}P-end-labeled restriction fragment. (a,c) Lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lanes 4–13, 10 pM, 20 pM, 50 pM, 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, respectively. **(b)** Lanes 1–13, see (a); lane 14, 10 μM CHL-Dp (**5**). **(b)** Alkylation sites on the DNA fragment, indicated by arrows. Sites 5′-AGCTGCT-3′ (A), 5′-TGCAGCA-3′ (B), and 5′-TGCAGCT-3′ (E) are match sites and 5′-TGTAGAA-3′ (C) and 5′-AGCAGTC-3′ (D) are double base pair mismatches (mismatches indicated by gray boxes).

Figure 7



(a) Sequence of the 241 base pair pHIV–LTR *Eco*RI–*Hind*III restriction fragment. The five binding sites are indicated by boxes. Base pair mismatches are indicated by gray boxes. Cleavage sites of adenine- or guanine-N3 lesions at 10 nM are indicated by arrows. Large and small arrows indicate major and minor alkylation sites, respectively. **(b)** Binding models for alkylation are shown. Models are colored as in Figure 2. Base pair mismatches are indicated by yellow boxes.

HOBt (2.7 mg). The solution was stirred for 1 h. The OBt ester solution was added to ImPy-β-ImPy-(*R*)^{H_2N}γ-ImPy-β-ImPy-β-Dp (4 mg) in 400 µl DMF followed by 200 µl DIEA. The reaction mixture was stirred for 2 h. TFA (150 µl) was added to the reaction mixture and the mixture was purified by reversed phase HPLC. ImPy-β-ImPy-(R)^{CHL}γ-ImPy-β-ImPy-β-Dp was recovered as a white powder upon lyophilization of the appropriate fraction (1 mg, 21% recovery). UV (H_2O) λ_{max} (ε) 306 (69,500); ; ^{1}H NMR (DMSO-d_6): δ 10.43 (s, 1 H), 10.31 (s, 1 H), 10.30 (s, 1 H), 10.25 (s, 1 H), 10.09 (s, 1 H), 9.96 (s, 1 H), 9.52 (s, 1 H), 8.09, (m, 8 H), 7.47 (s, 2 H), 7.45 (s, 1 H), 7.40 (s, 1 H), 7.25 (m, 4 H), 7.05 (s, 2 H), 7.02 (s, 1 H), 6.96 (m, 4 H), 6.67 (s, 1 H), 6.64 (s, 1 H), 3.98 (s, 3 H), 3.95 (s, 6 H), 3.81 (s, 6 H), 3.79 (s, 1 H), 3.69 (s, 6 H), 3.52 (M, 8 H), 3.12 (q, 2 H, J = 5.8 Hz), 3.01 (quintet, 2 H, J = 6.0 Hz), 2.75 (d, 6 H, J = 4.7 Hz), 2.61 (m, 4 H), 2.37 t, (t, J = 6.9 Hz), 2.03 (t, 6 H, J = 6.9 Hz), 2.01 (m, 8 H), 1.74 (m, 4 H), 1.63 (m, 2 H), 1.47 (m, 2 H); MALDI–TOF MS (monoisotopic) [M + H] 1666.7 (1666.7 calc'd for $C_{76}H_{98}Cl_2N_{27}O_{13}$)

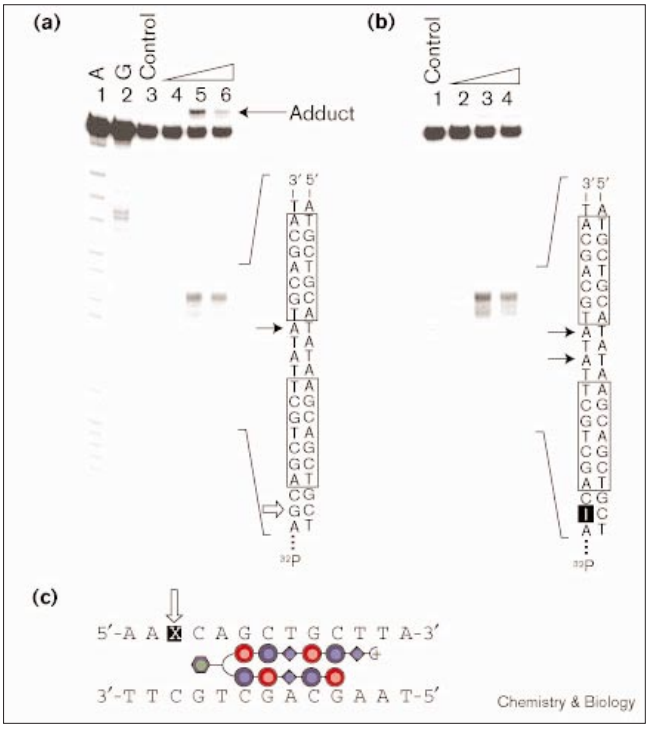
*Dp-CHL (**4**).* To a solution of chlorambucil (304 mg) in 500 µl DMF was added DCC (208 mg) and HOBt (143 mg). The solution was stirred for 1 h. The OBt ester solution was added to (dimethylamino)-propylamine (63 µl) in 500 µl DMF followed by 500 µl DIEA. The reaction mixture was stirred for 2 h. TFA (0.01% (wt/v), 6 ml) was added to the reaction mixture and the mixture was purified by reversed phase HPLC. Dp-CHL is recovered as a yellow oil upon lyophilization of the appropriate fraction (50 mg, 12.9% recovery). NMR; MALDI–TOF MS (monoisotopic) [M + H] 388.4 (388.4 calc'd for $C_{19}H_{31}Cl_2N_3O$).

Preparation of 3′- or 5′-^{32}P-labeled DNA restriction fragment
Plasmid pHIV–LTR [27] was digested with *Eco*RI, labeled at either the 3′ or 5′ end and digested with *Hind*III. The 241-base-pair restriction fragment was isolated by nondenaturing gel electrophoresis. Chemical sequencing reactions were performed as described [28,29]. Standard techniques were employed for DNA manipulation [30].

Detection of heat-induced single-strand DNA cleavage
All reactions were executed in a total volume of 40 µl. A polyamide stock solution or H_2O was added to an assay buffer containing radiolabeled restriction fragment (10,000 cpm), affording final solution conditions of 10 mM Tris–HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, and either 1 pM–10 nM nitrogen mustard–polyamide conjugate or no conjugate (for reference lanes). The solutions were equilibrated at 37°C for the desired time period. The reactions were stopped by adding 60 µl of a solution containing 0.6 M NaOAc, 12.5 mM EDTA, 0.15 mM calf thymus DNA, 2.0 M NaCl, 0.8 mg/ml glycogen, and ethanol precipitated. Reactions were resuspended in 40 µl 10 nM sodium citrate buffer, pH 7.2, and heated for 30 min at 95°C. The reactions were ethanol precipitated and resuspended in TBE/80% formamide loading buffer, denatured by heating at 85°C for 10 min, and placed on ice. The reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross-link, 7 M urea) in TBE at 2000 V. Gels were dried and exposed to a storage phosphor screen (Molecular Dynamics) [31].

Quantitative DNase I footprint titration experiments
All reactions were executed in a total volume of 400 µl. A polyamide stock solution or H_2O (for reference lanes) was added to an assay

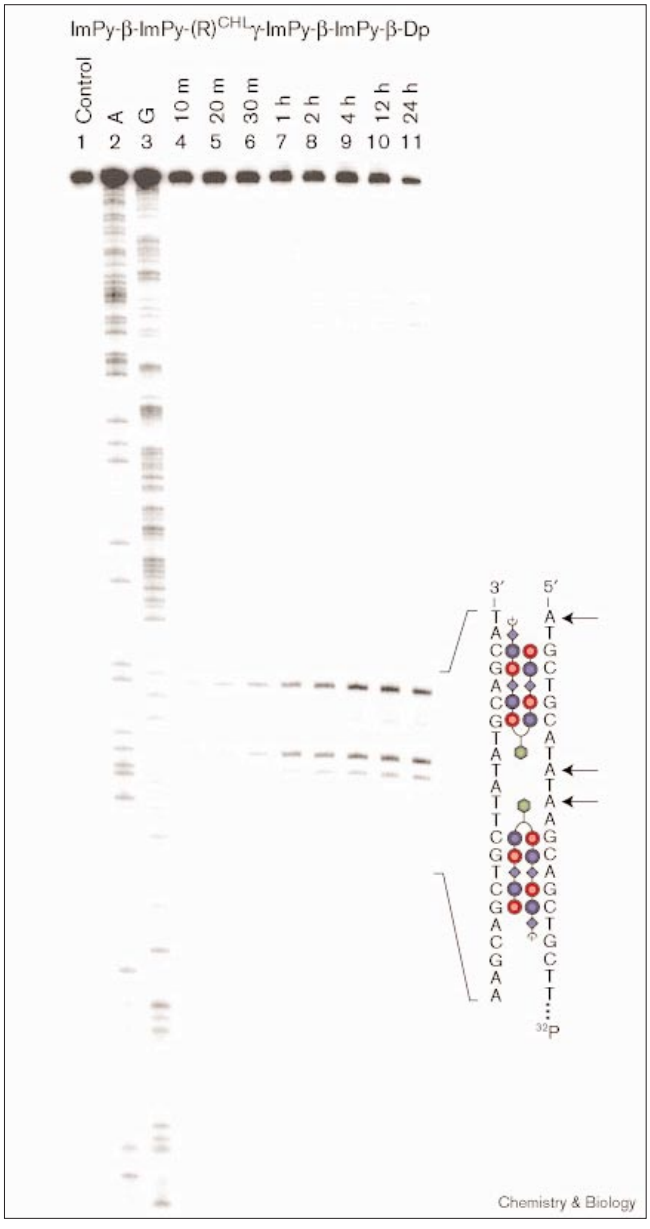
Figure 8



Thermal cleavage assay experiments with ImPy-β-ImPy-(*R*)^{CHL}γ-ImPy-β-ImPy-β-Dp (**2**) on 5′-^{32}P-end-labeled 120 base pair oligonucleotides with and without inosine substitutions. **(a,b)** Storage phosphor autoradiograms of 8% denaturing polyacrylamide gels used to separate the fragments generated by heat-induced DNA cleavage at alkylation sites. All reactions contained 10,000 cpm restriction fragment, 10 mM Tris–HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$ and were performed at 37°C. Following 24 h of equilibration, the DNA pellet was resuspended in sodium citrate buffer (pH 7.2) and heated to 90°C for 15 min to thermally cleave at sites of adenine- or guanine-N3 lesions. Cleavage sites of adenine-N3 lesions are shown. The solid arrows indicate cleavage bands from cleavage sites of adenine-N3 lesions on the restriction fragment. The hollow arrow indicates alkylation at the exocyclic amine of guanine. (a) The fragment contains the site 5′-AGC-AGCTGCT. Lane 1, A-specific reaction; lane 2, G-specific reaction; lanes 3–6, 100 pM, 1 nM, 10 nM, respectively. (b) The fragments contain the site 5′- A*I*CAGCTGCT-3′. Lanes 1–4, 100 pM, 1 nM, 10 nM, respectively. (c) Model bound to match site 5′-AGCTGCT-3′. X indicates site of inosine substitution. Models are colored as in Figure 2.

buffer containing radiolabeled restriction fragment (20,000 cpm) affording final solution conditions of 10 mM Tris–HCl, 10 mM KCl, 10 mM $MgCl_2$, 5 mM $CaCl_2$, pH 7.0, and either 1 pM–10 nM nitrogen mustard-polyamide conjugate or no conjugate (for reference lanes). The solutions were allowed to equilibrate at 22°C for 18 h. Footprinting reactions were initiated by the addition of 4 µl of a DNase I stock solution (at the appropriate concentration to give about 55% intact DNA) containing 1 mM DTT and allowed to proceed for 7 min at 22°C. The reactions were stopped by adding 50 µl of a solution containing 1.25 M NaCl, 100 mM EDTA, 0.2 mg/ml glycogen, and 28 µM (bp) calf thymus DNA, and ethanol precipitated. Reactions were resuspended in TBE/80% formamide loading buffer, denatured by heating at 85°C for 10 min, and then placed on ice. The reaction products were separated by electrophoresis on an 8% polyacrylamide gel (5% cross link, 7 M urea) in TBE at 2000 V for 1.5 h. Gels were dried and exposed to a storage phosphor screen (Molecular Dynamics).

Figure 9



Thermal cleavage assay experiments with ImPy-β-ImPy-(*R*)^{CHL}γ-ImPy-β-ImPy-β-Dp (**2**) on the 3′-^{32}P-end-labeled 241-base pair *Eco*RI/*Hind*III restriction fragment from plasmid pHIV-LTR. Storage phosphor autoradiogram of 8% denaturing polyacrylamide gels used to separate the fragments generated by heat-induced DNA cleavage at alkylation sites. All reactions contained 10,000 cpm restriction fragment, 10 mM Tris–HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$ and 5 mM $CaCl_2$ and were performed at 37°C. Following of equilibrations of 10 min–24 h, the DNA pellet was resuspended in sodium citrate buffer (pH = 7.2) and heated to 90°C for 15 min to thermally cleave at sites of adenine- or guanine-N3 lesions. Lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lanes 4–11, 500 pM ImPy-β-ImPy-(*R*)^{CHL}γ-ImPy-β-ImPy-β-Dp (**2**), equilibrations for 10 min, 20 min, 30 min, 1 h, 2 h, 4 h, 12 h and 24 h, respectively. Right, models bound to match sites on the DNA fragment. The polyamide is colored as in Figure 2. The solid arrows indicate cleavage bands from cleavage sites of adenine- or guanine-N3 lesions on the restriction fragment.

Quantitation and data analysis

Data from the footprint titration gels were obtained using a Molecular Dynamics 400S PhosphorImager followed by quantitation using ImageQuant software (Molecular Dynamics). Background-corrected volume integration of rectangles encompassing the footprint sites and a reference site at which DNase I reactivity was invariant across the titration generated values for the site intensities (I_{site}) and the reference intensity (I_{ref}). The apparent fractional occupancy (θ_{app}) of the sites was calculated using the equation:

$$\theta_{app} = 1 - \frac{I_{site}/I_{ref}}{I^{o}_{site}/I^{o}_{ref}} \quad (1)$$

where I^{o}_{site} and I^{o}_{ref} are the site and reference intensities, respectively, from a control lane to which no polyamide was added. The ($[L]_{tot}$, θ_{app}) data points were fitted to a Langmuir binding isotherm (eq 2, $n = 1$) by minimizing the difference between θ_{app} and θ_{fit}, using the modified Hill equation:

$$\theta_{fit} = \theta_{min} + (\theta_{max} - \theta_{min}) \frac{K_a{}^{n}[L]^{n}tot}{1 + K_a{}^{n}[L]^{n}tot} \quad (2)$$

where $[L]_{tot}$ is the total polyamide concentration, K_a is the equilibrium association constant, and θ_{min} and θ_{max} are the experimentally determined site saturation values when the site is unoccupied or saturated, respectively. The data were fitted using a nonlinear least-squares procedure with K_a, θ_{max} and θ_{min} as the adjustable parameters. All acceptable fits had a correlation coefficient of $R > 0.97$. Five sets of data were used in determining each association constant. All lanes from each gel were used unless visual inspection revealed a data point to be obviously flawed relative to neighboring points.

Acknowledgements

We are grateful to the National Institutes of Health General Medical Sciences for research support, and to the Ralph M Parsons Foundation for a predoctoral fellowship to N.R.W.

References

1. Trauger, J.W., Baird, E.E. & Dervan, P.B. (1996). Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **382**, 559-561.
2. Swalley, S.E., Baird, E.E. & Dervan, P.B. (1997). Discrimination of 5′-GGGG-3′,5′-GCGC-3′, and 5′-GGCC-3′ sequences in the minor groove of DNA by eight-ring hairpin polyamides. *J. Am. Chem. Soc.* **119**, 6953-6961.
3. Turner, J.M., Baird, E.E. & Dervan, P.B. (1997). Recognition of seven base pair sequences in the minor groove of DNA by ten-ring pyrrole-imidazole polyamide hairpins. *J. Am. Chem. Soc.* **119**, 7636-7644.
4. Trauger, J.W., Baird, E.E. & Dervan, P.B. (1998). Cooperative hairpin dimers for recognition of DNA by pyrrole-imidazole polyamides. *Angew. Chem. Int. Ed.* **37**, 1421-1423.
5. Dervan, P.B. & Burli, R.W. (1999). Sequence-specific DNA recognition by polyamides. *Curr. Opin. Chem. Biol.* **3**, 688-693.
6. Wade, W.S., Mrksich, M. & Dervan, P.B. (1992). Design of peptides that bind in the minor groove of DNA at 5′-(A,T)G(A,T)C(A,T)-3′ sequences by a dimeric side-by-side motif. *J. Am. Chem. Soc.* **114**, 8783-8794.
7. Mrksich, M., Wade, W.S., Dwyer, T.J., Geierstanger, B.H., Wemmer, D.E. & Dervan, P.B. (1992). Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin. *Proc. Natl Acad. Sci. USA* **89**, 7586-7590.
8. White, S., Baird, E.E. & 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.
9. Kielkopf, C.L., Baird, E.E., Dervan, P.D. & Rees, D.C. (1998). Structural basis for G•C recognition in the DNA minor groove. *Nat. Struct. Biol.* **5**, 104-109.
10. Kielkopf, C.L., *et al.,* & Rees, D.C. (1998). A structural basis for recognition of A•T and T•A base pairs in the minor groove of B-DNA. *Science* **282**, 111-115.
11. White, S., Szewczyk, J.W., Turner, J.M., Baird, E.E. & Dervan, P.B. (1998). Recognition of the four Watson–Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **391**, 468-471.
12. Trauger, J.W., Baird, E.E., Mrksich, M. & Dervan, P.B. (1996). Extension of sequence-specific recognition in the minor groove of DNA by pyrrole-imidazole polyamides to 9–13 base pairs. *J. Am. Chem. Soc.* **118**, 6160-6166.
13. Swalley, S.E., Baird, E.E. & Dervan, P.B. (1997). A pyrrole-imidazole polyamide motif for recognition of eleven base pair sequences in the minor groove of DNA. *Chem. Eur. J.* **3**, 1600-1607.
14. Trauger, J.W., Baird, E.E. & Dervan, P.B. (1998). Recognition of 16 base pairs in the minor groove of DNA by a pyrrole-imidazole polyamide dimer. *J. Am. Chem. Soc.* **120**, 3534-3535.
15. Gottesfeld, J.M., Neely, L., Trauger, J.W., Baird, E.E. & Dervan, P.B. (1997). Regulation of gene expression by small molecules. *Nature* **387**, 202-205.
16. Dickinson, L.A., *et al.,* & 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.
17. Mattes, W.B., Hartley, J.A. & Kohn, K.W. (1986). DNA-sequence selectivity of guanine-N7 alkylation by nitrogen mustards. *Nucleic Acids Res.* **14**, 2971-2987.
18. Arcamone, F.M. *et al.*, & Mongelli, M. (1989). Synthesis, DNA-binding properties and antitumor activity of novel distamycin derivatives. *J. Med. Chem.* **32**, 774-788.
19. Baraldi, P.G., Cacciari, B., Guiotto, A., Romagnoli, R., Zaid, A.N. & Spalluto, G. (1998). Heterocyclic analogs of DNA minor groove alkylating agents. *Curr. Pharm. Des.* **4**, 249-276.
20. Herman, D.M., Baird, E.E. & Dervan, P.B. (1998). Stereochemical control of the DNA binding affinity, sequence specificity, and orientation preference of chiral hairpin polyamides in the minor groove. *J. Am. Chem. Soc.* **120**, 1382-1391.
21. Baird, E.E. & Dervan, P.B. (1996). Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146.
22. Reynolds, V.L., Molineux, I.J., Kaplan, D.J., Swenson, D.H. & Hurley, L.H. (1985). Reaction of the antitumor antibiotic CC-1065 with DNA – location of the site of thermally induced strand breakage and analysis of DNA-sequence specificity. *Biochemistry* **24**, 6228-6237.
23. Wyatt, M.D., Lee, M., Garbiras, B.J., Souhami, R.L. & Hartley, J.A. (1995). Sequence specificity of alkylation for a series of nitrogen mustard-containing analogs of distamycin of Increasing binding-site size – evidence for increased cytotoxicity with enhanced sequence specificity. *Biochemistry* **34**, 13034-13041.
24. Brooks, N., *et al.,* & Lee, M. (1997). Structure-activity relationship of a series of C-terminus modified aminoalkyl, diaminoalkyl- and anilino-containing analogues of the benzoic acid mustard distamycin derivative tallimustine: synthesis, DNA binding and cytotoxicity studies. *Bioorg. Med. Chem.* **5**, 1497-1507.
25. Kundu, G.C., Schullek, J.R. & Wilson, I.B. (1994). The alkylating properties of chlorambucil. *Pharm. Biochem. Behav.* **49**, 621-624.
26. Zhi-Fu, T., Fujiwara, T., Saito, I. & Sugiyama, H. (1999). Rational design of sequence-specific DNA alkylating agents based on duocarmycin A and pyrrole-imidazole hairpin polyamides. *J. Am. Chem. Soc.* **121**, 4961-4967.
27. Jones, K.A. & Peterlin, B.M. (1994). Control of RNA initiation and elongation at the HIV-1 promoter. *Annu. Rev. Biochem.* **63**, 717-743.
28. Maxam, A.M. & Gilbert, W. (1986). Sequencing DNA by labeling the end and breaking at bases – DNA segments, end labels, cleavage reactions, polyacrylamide gels and strategies. *Mol. Biol.* **20**, 461-509.
29. Iverson, B.L. & Dervan, P.B. (1987). Adenine-specific DNA chemical sequencing reaction. *Nucleic Acids Res.* **15**, 7823-7830.
30. Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989). *Molecular Cloning.* Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
31. Johnston, R.F., Pickett, S.C. & Barker, D.L. (1990). Autoradiography using storage phosphor technology. *Electrophoresis* **11**, 355-360.