Pergamon

Bioorganic & Medicinal Chemistry 9 (2001) 2093–2103

BIOORGANIC & MEDICINAL CHEMISTRY

Inhibition of Major Groove DNA Binding bZIP Proteins by Positive Patch Polyamides

Ryan E. Bremer, Nicholas R. Wurtz, Jason W. Szewczyk and Peter B. Dervan*

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

Received 30 January 2001; accepted 4 April 2001

Abstract—Cell permeable synthetic ligands that bind to predetermined DNA sequences offer a chemical approach to gene regulation, provided inhibition of a broad range of DNA transcription factors can be achieved. DNA minor groove binding polyamides containing aminoalkyl substituents at the N-1 of a single pyrrole residue display inhibitory effects for a bZIP protein which binds exclusively in the DNA major groove. For major groove protein inhibition, specific protein-DNA contacts along the phosphate backbone were targeted with the positively charged dimethylamino substituent on the backbone of a minor groove binding polyamide hairpin. Remarkably, these polyamides bind DNA with enhanced affinity and uncompromised specificity when compared to polyamides with the aminoalkyl moiety at the C-terminus. By adding bZIP transcription factors to the class of protein–DNA complexes that can be disrupted by minor groove binding ligands, these results may increase the functional utility of polyamides as regulators of gene expression.

Introduction

Hairpin polyamides containing pyrrole (Py), imidazole (Im) and hydroxypyrrole (Hp) amino acids are synthetic ligands that bind predetermined DNA sequences in the minor groove with affinities and specificities comparable to many DNA binding proteins.[1,2] Rules have been developed that allow for sequence specific recognition of the DNA minor groove by relating each Watson–Crick base pair with a particular pairing of the aromatic Py, Im and Hp rings.[1–4] The crescent shaped polyamides bind in the minor groove of DNA with pairs of aromatic rings stacked against each other and the walls of the groove, allowing the backbone amide hydrogens and the substituents at the 3-position of the Py, Im and Hp residues to make specific contacts with the edges of the intact base pairs. A γ-aminobutyric acid residue (γ) connects the polyamide subunits in a 'hairpin' motif which enhances affinity and unambiguously locks the desired ring pairings in register.

A requirement for the general application of minor groove binding polyamides as regulators of gene transcription is motifs that can effectively inhibit all classes of DNA binding proteins, especially transcription factors which bind regulatory elements in gene promoter regions. Polyamides have been found to interfere with protein-DNA recognition in cases where contacts in the *minor* groove are important for protein-DNA binding affinity.[2] Polyamides targeted to minor groove contacts of transcription factors Ets-1, LEF-1, and TBP inhibited DNA-binding of the each protein in mobility gel shift assays.[5] In contrast, polyamides have been shown to bind simultaneously with a some proteins that exclusively occupy the DNA *major* groove,[6–8] such as the bZIP protein, GCN4. The ubiquity of major groove contacts in protein-DNA recognition provides the impetus to develop approaches for the inhibition of major groove proteins by DNA minor groove binding polyamides.

Hairpin polyamides with the tripeptide Arg-Pro-Arg at the C-terminus have been shown to inhibit GCN4, potentially by competing with the protein side chains for electrostatic contacts to the DNA phosphate backbone.[8] However, Arg-Pro-Arg modified polyamides may be limited by the use of α-amino acids with regard to biostability. This lead us to ask the question of whether a *N*-aminoalkylpyrrole residue could serve as a simple nonpeptide substitute for the Arg-Pro-Arg tripeptide positive patch that can be placed at any position along the polyamide backbone which would allow for discrete targeting of phosphodiester protein contacts.

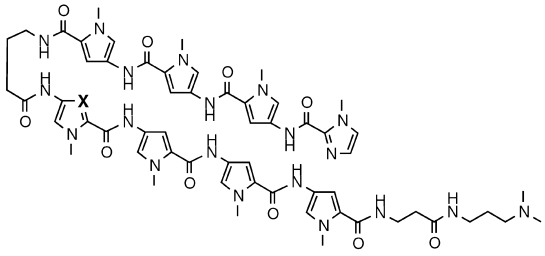
Analysis of an X-ray crystal structure of a polyamide bound to its cognate DNA site reveals that the N-1

*Corresponding author. Tel.: +1-626-356-6002; fax: +1-626-683-8753; e-mail: dervan@caltech.edu

0968-0896/01/$ - see front matter
PII: S0968-0896(01)00122-5

substituent of each pyrrole ring makes its closest phosphate contact two phosphates to the 3′ side of the base the ring pair recognizes, consistent with the minor groove binding nature of the polyamide.[9] Polyamides displaying a *N*-aminoalkyl positive patch could interfere with protein binding to this specific phosphate and consequently inhibit specific contacts with DNA. The increasing availability of protein-DNA crystal structures provide the necessary information to design a polyamide that targets important protein-DNA contacts. Previous attempts to inhibit major groove binding transcription factors with minor groove binding ligands have used polyamine substituents with 10 charges attached to a tripyrrole, likely resulting in a ligand with significantly reduced sequence specificity.[8,10]

We report here the synthesis and DNA binding properties of a series of hairpin polyamides containing *N*-aminoalkyl substitutions at a single pyrrole residue and their ability to inhibit DNA binding by a major groove binding protein, GCN4 (222-281), a member of the bZIP family of transcriptional activators. The X-ray crystal structure of the GCN4–DNA complex reveals several electrostatic contacts between positively charged side chains and the DNA phosphate backbone.[11] Of these, ethylation interference experiments suggest that the Lys246-DNA phosphate interaction is particularly important for DNA binding by GCN4.[12] Polyamides were designed to place a competing positive charge at the Lys246-DNA phosphate position to disrupt this potentially crucial interaction. A series of polyamides were synthesized that contain multiple amines and selected ether substitutions as controls at the N-1 position of the pyrrole. The DNA binding affinity and specificity of each of these compounds for their target sites was evaluated using quantitative DNase I footprinting and compared to analogues containing the traditional *N*-methyl substituents. Gel mobility shift assays were employed to investigate the ability of these polyamides to inhibit DNA binding by a major groove binding protein.



1) X=CH, ImPyPyPy-γ-PyPyPyPy-β-Dp
2) X=N, ImPyPyPy-γ-ImPyPyPy-β-Dp

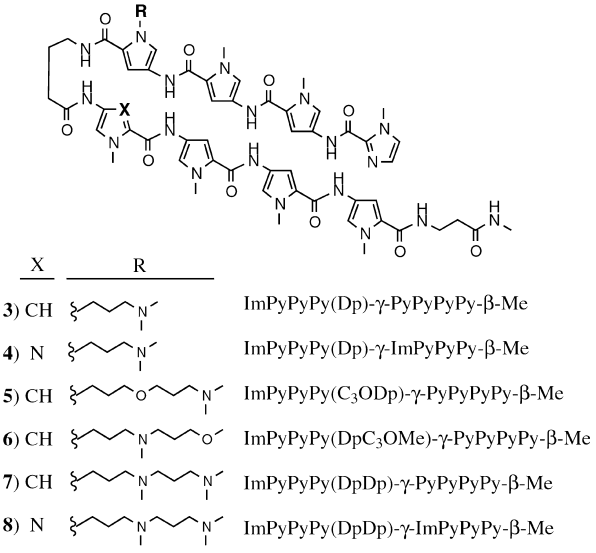
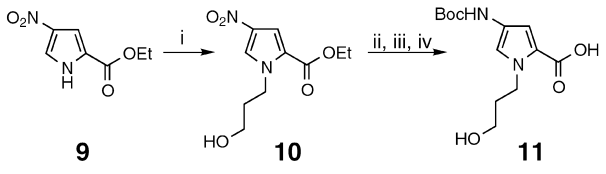


Figure 1. Structures of eight-ring hairpin polyamides containing *N*-methylpyrrole residues and a C-terminal Dp tail (top) and polyamides with *N*-aminoalkylpyrrole residues and a C-terminal *N*-methylamide. ImPyPyPy-γ-PyPyPyPy-β-Dp (**1**), and ImPyPyPy-γ-ImPyPyPy-β-Dp (**2**) (top). ImPyPyPy(Dp)-γ-PyPyPyPy-β-Me (**3**), ImPyPyPy(Dp)-γ-ImPyPyPy-β-Me (**4**), ImPyPyPy(C_3ODp)-γ-PyPyPyPy-β-Me (**5**), ImPyPyPy(DpC_3OMe)-γ-PyPyPyPy-β-Me (**6**), ImPyPyPy(DpDp)-γ-PyPyPyPy-β-Me (**7**) and ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Me (**8**).

Results

Polyamide synthesis

The synthesis of *N*-aminoalkylpyrrole-containing polyamides **3**–**8** (Fig. 1) required the preparation of a new pyrrole monomer suitable for solid phase synthesis and subsequent modification at the N-1 position. (3-Hydroxypropyl)-4-[(*tert*-butoxycarbonyl)amino]-pyrrole-2-carboxylic acid, (**11**), introduces a 3-hydroxypropyl moiety at the N-1 of pyrrole which can be modified by sulfonylation followed by nucleophilic displacement to afford a variety of derivatives from a single polyamide precursor. By cleaving the polyamide from resin with methylamine, polyamides (**3** and **4**) can be obtained that are structural isomers of those previously reported (**1** and **2**) and limit the overall charge of the *N*-diaminoalkylpyrrole-containing polyamides (**7** and **8**).[13] Ethyl 4-nitropyrrole-2-carboxylate (**9**) is available in 500 g quantities as described.[14] Alkylation of the pyrrole N-1 with 3-iodopropanol provided **10**. Reduction and Boc protection of the amine followed by hydrolysis of the ethyl ester afforded **11** in gram quantities in 53% yield (Scheme 1). The *N*-(3-hydroxypropyl)pyrrole monomer **11** was used without further protection in the Boc-chemistry machine assisted solid phase synthesis of polyamides **12** and **13** from Boc-β-Ala-PAM resin (Fig. 2).[15] In order to provide polyamides that differ from those previously reported only by the placement of the positively charged *N,N*-dimethylaminopropyl moiety and the *N*-methyl group, it was necessary to cleave the polyamide from resin using methylamine. This was accomplished by condensing methylamine in a Parr apparatus with resin-



Scheme 1. Synthesis of the Boc-*N*-(3-hydroxypropyl)pyrrole-acid monomer (**11**) for solid phase synthesis. (i) Iodopropanol, acetone, K_2CO_3. (ii) H_2, 10% Pd/C, DMF. (iii) Boc-anhyride, DIEA. (iv) 1 M KOH, EtOH, 70 °C.

bound polyamide, followed by heating overnight (55 °C) to provide crude polyamides ImPyPyPy(C_3OH)-γ-PyPyPyPy-β-Me, (**12**), and ImPyPyPy(C_3OH)-γ-ImPyPyPy-β-Me, (**13**), containing the *N*-(3-hydroxypropyl)pyrrole residue suitable for further modification. The free hydroxyl of the crude polyamide was activated with *p*-toluenesulfonyl chloride or methanesulfonyl chloride in dry pyridine. Nucleophilic displacement of the tosylate or mesylate with the appropriate amine (**3, 4**, **6**–**8**) or sodium salt (**5**) followed by preparatory reverse phase HPLC provided the desired *N*-aminoalkyl substituted polyamides in recoveries similar to those observed for *N*-methyl derivatives cleaved with dimethylaminopropylamine.

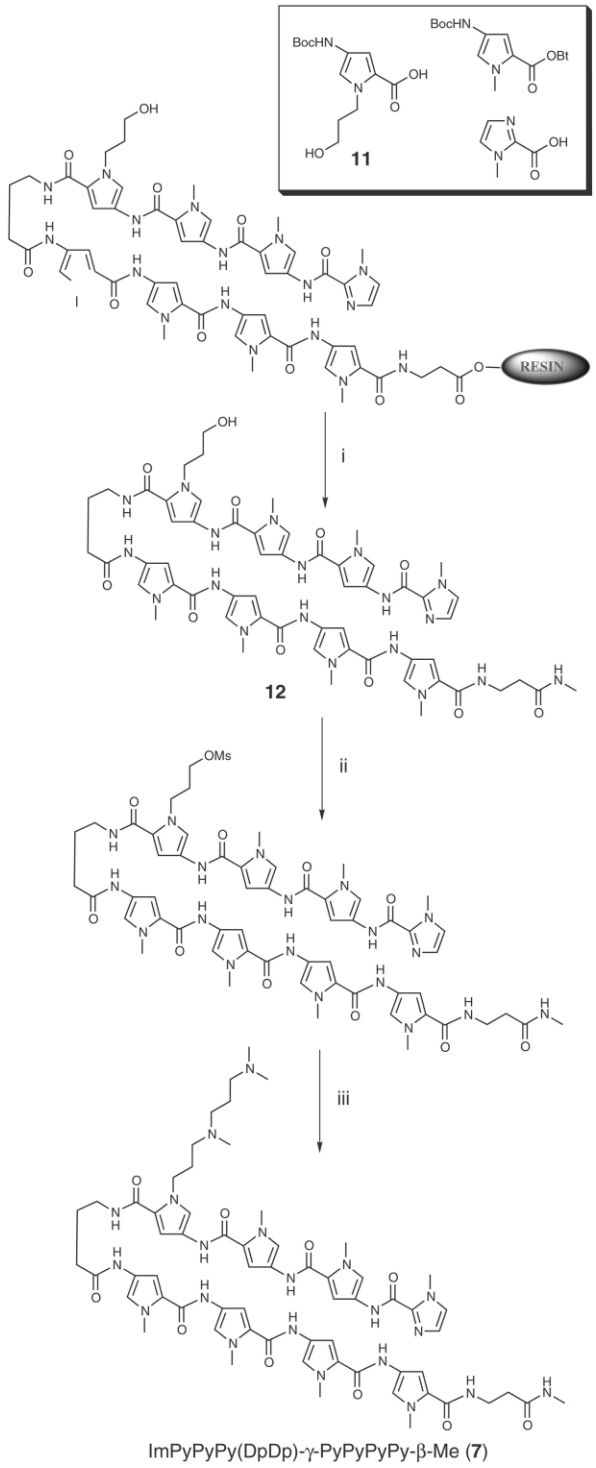


Figure 2. Synthetic scheme for solid phase preparation of *N*-aminoalkylpyrrole containing polyamides.[15] Cycling protocols consisted of trifluoroacetic acid (TFA) deprotection followed by coupling with HOBt-activated Boc-pyrrole, Boc-imidazole, Boc-*N*-(3-hydroxypropyl)pyrrole (**11**) or aliphatic amino acid ester. (i) The PAM resin is then cleaved by treatment with methylamine. (ii) Activation with methanesulfonyl chloride or toluenesulfonyl chloride followed by reaction with the appropriate amine or alkoxide (iii) and reverse phase HPLC provided the desired products. The synthesis of ImPyPyPy(DpDp)-γ-PyPyPyPy-β-Me (**7**) is shown as a representative example.

Quantitative DNase I footprint titrations

The DNA binding affinity and specificity for the six base pair target sites were measured for the *N*-aminoalkylpyrrole-containing polyamides by quantitative DNase I footprinting on the 229-bp *Afl*II/*Fsp*I restriction fragment of pJT8.[13] Polyamides of the basic composition ImPyPyPy-γ-ImPyPyPy (**2**, **4** and **8**) were designed to bind the site 5′-AGTACT-3′ according to the pairing rules. Conversely, the single atomic substitution that converts an imidazole to a pyrrole residue provides polyamides based on ImPyPyPy-γ-PyPyPyPy (**1**, **3** and **5**–**7**), targeted to 5′-AGTATT-3′. The sites 5′-AGTACT-3′ and 5′-AGTATT-3′ are for ImPyPyPy-γ-ImPyPyPy polyamides 'match' and 'single base pair mismatch' sites, respectively, and for ImPyPyPy-γ-PyPyPyPy polyamides 'single base pair mismatch' and 'match' sites, respectively.

DNase I footprinting revealed that ImPyPyPy(Dp)-γ-ImPyPyPy-β-Me (**4**) bound the 5′-AGTACT-3′ with an equilibrium association constant $K_a = 1.9\ (\pm 0.4)\times 10^{11}$ M^{-1}, and a 56-fold preference over the single base pair mismatch site [$K_a = 3.3\ (\pm 0.4)\times 10^{9}$ M^{-1}] (Fig. 3A, Table 1). This represents ≈10-fold increase in affinity relative to the parent polyamide differing in the placement of the dimethylaminopropyl and *N*-methyl moieties, ImPyPyPy-γ-ImPyPyPy-β-Dp (**2**) [$K_a = 1.4\ (\pm 0.5)\times 10^{10}$ M^{-1} for 5′-AGTACT-3′], and uncompromised specificity [61-fold, $K_a = 2.3\ (\pm 0.5)\times 10^{8}$ M^{-1} for 5′-AGTATT-3′]. The increase in affinity upon placement of the alkylamine on the pyrrole ring was also observed for ImPyPyPy(Dp)-γ-PyPyPyPy-β-Me (**3**) relative to ImPyPyPy-γ-PyPyPyPy-β-Dp (**1**). Polyamide **3** bound the 5′-AGTATT-3′ match site with an affinity $K_a = 1.3\ (\pm 0.6)\times 10^{10}$ M^{-1} and a 4-fold preference over the single base pair mismatch site [$K_a = 2.9\ (\pm 0.6)\times 10^{9}$ M^{-1}]. In comparison, the parent polyamide, **1**, bound both the match and mismatch sites with reduced affinity [$K_a = 1.5\ (\pm 0.3)\times 10^{9}$ M^{-1} for 5′-AGTATT-3′ and $K_a = 2.4\ (\pm 1.2)\times 10^{8}$ M^{-1} for 5′-AGTACT-3′].

In order to explore the potential applications of aminoalkyl substitutions of the pyrrole ring, the DNA binding affinities and specificities of analogues containing multiple amines or selected ether linkages were determined. The presence of an additional aminopropyl moiety to provide a polyamide with potentially two positive charges as in ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Me (**8**), resulted in an additional slight increase in affinity for the match site relative to **4** [$K_a = 3.0$

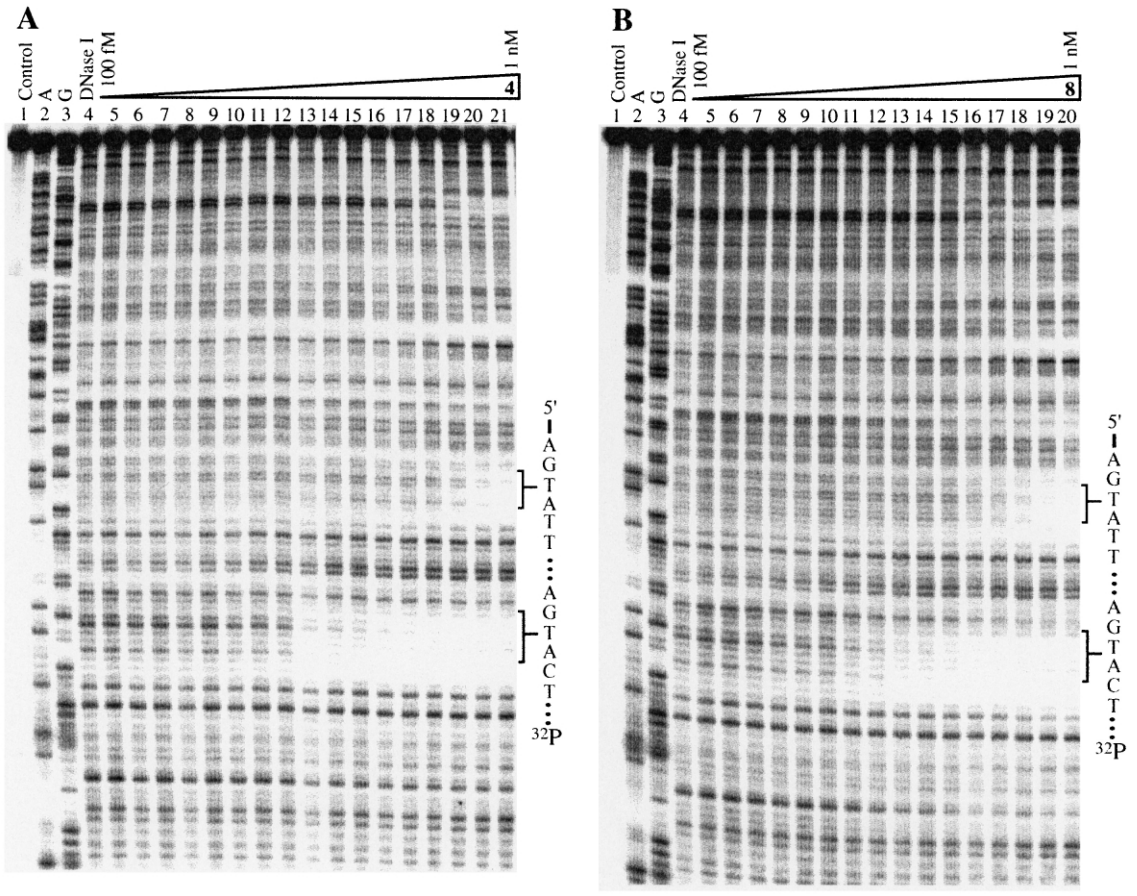


Figure 3. Storage phosphor autoradiograms of quantitative DNase I footprint titration experiments with (A) ImPyPyPy(Dp)-γ-ImPyPyPy-β-Dp (**4**) and (B) ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Dp (**8**) on the 3′-^{32}P-end-labeled *Aff*II/*Fsp*I restriction fragment from pJT8.[13] All reactions contained 8 kcpm 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 5–21, 100 fM, 200 fM, 500 fM, 1 pM, 1.5 pM, 2.5 pM, 4.0 pM, 6.5 pM, 10 pM, 15 pM, 25 pM, 40 pM, 65 pM, 100 pM, 200 pM, 500 pM, 1 nM, respectively of **3** or **8**. The 65 pM lane has been omitted for **8**. The positions of the match (5′-AGTACT-3′) and single base pair mismatch (5′-AGTATT-3′) sites are indicated.

$(\pm 0.4)\times 10^{11}$ M^{-1}], and a 36-fold preference over the mismatch site [$K_a = 8.2\ (\pm 3.1)\times 10^9$ M^{-1}] (Fig. 3B, Table 1). This represents only a modest decrease in specificity for **8** relative to **4** (36-fold and 56-fold, respectively). A similar decrease in specificity was observed for ImPyPyPy(DpDp)-γ-PyPyPyPy-β-Me (**7**) relative to polyamide **3** [$K_a = 6.9\ (\pm 1.2)\times 10^9$ M^{-1} for 5′-AGTATT-3′ and $K_a = 3.1\ (\pm 1.0)\times 10^9$ M^{-1} for 5′-AGTACT-3′]. Substitution of either tertiary amine in **7** to provide the singly charged polyamides ImPyPyPy(C_3ODp)-γ-PyPyPyPy-β-Me (**5**) and ImPyPyPy(Dp-C_3OMe)-γ-PyPyPyPy-β-Me (**6**) restored the optimal affinity and specificity observed with **3**, while retaining the overall size and length of **7**. Polyamide **5**, containing an internal ether linkage and a terminal dimethylamine, bound the 5′-AGTATT-3′ match site with a $K_a = 1.2\ (\pm 0.1)\times 10^{10}$ M^{-1} with a 7-fold preference over the 5′-AGTACT-3′ mismatch [$1.8\ (\pm 0.9)\times 10^9$ M^{-1}]. Similarly, the introduction of an internal amine and a terminal methyl ether in **6** resulted in $K_a = 1.5\ (\pm 0.3)\times 10^{10}$ M^{-1} and $2.2\ (\pm 0.7)\times 10^9$ M^{-1} for the match and mismatch sites, respectively.

Table 1. Equilibrium association constants (M^{-1})[a]

Polyamide	Match	Mismatch
ImPyPyPy-γ-PyPyPyPy-β-Dp (**1**)	$1.5\ (\pm 0.3)\times 10^9$	$2.4\ (\pm 1.2)\times 10^8$
ImPyPyPy(Dp)-γ-PyPyPyPy-β-Me (**3**)	$1.3\ (\pm 0.6)\times 10^{10}$	$2.9\ (\pm 0.6)\times 10^9$
ImPyPyPy-γ-ImPyPyPy-β-Dp (**2**)	$1.4\ (\pm 0.5)\times 10^{10}$	$2.3\ (\pm 0.5)\times 10^8$
ImPyPyPy(Dp)-γ-ImPyPyPy-β-Me (**4**)	$1.9\ (\pm 0.4)\times 10^{11}$	$3.3\ (\pm 0.4)\times 10^9$
ImPyPyPy(C_3ODp)-γ-PyPyPyPy-β-Me (**5**)	$1.2\ (\pm 0.1)\times 10^{10}$	$1.8\ (\pm 0.9)\times 10^9$
ImPyPyPy(DpC_3OMe)-γ-PyPyPyPy-β-Me (**6**)	$1.5\ (\pm 0.3)\times 10^{10}$	$2.2\ (\pm 0.7)\times 10^9$
ImPyPyPy(DpDp)-γ-PyPyPyPy-β-Me (**7**)	$6.9\ (\pm 1.2)\times 10^9$	$3.1\ (\pm 1.0)\times 10^9$
ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Me (**8**)	$3.0\ (\pm 0.4)\times 10^{11}$	$8.2\ (\pm 0.4)\times 10^9$

[a]Values reported are the mean values obtained from at least three DNase I footprint titration experiments. The assays were carried out at 22 °C, 10 mM Tris·HCl (pH 7.0), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. The match and mismatch sites for **1**, **3** and **5**–**7** are 5′-ttAGTATTtg-3′ and 5′-ttAGTACTtg-3′, respectively. The match and mismatch sites for **2, 4** and **8** are 5′-ttAGTACTtg-3′ and 5′-ttAGTATTtg-3′, respectively.

Protein binding inhibition by polyamides determined by gel shift

The ability of *N*-aminoalkylpyrrole-containing polyamides to inhibit DNA binding by GCN4 (222–281), was investigated using gel mobility shift assays. The carboxy-terminal 60 amino acids of GCN4 (222–281) contain the 'leucine zipper' dimerization domain and the basic region DNA binding domain and have been shown to be sufficient for sequence specific DNA binding.[16,17] GCN4 (222–281) binds its target site as a dimer of α-helices that make specific hydrogen bonds, van der

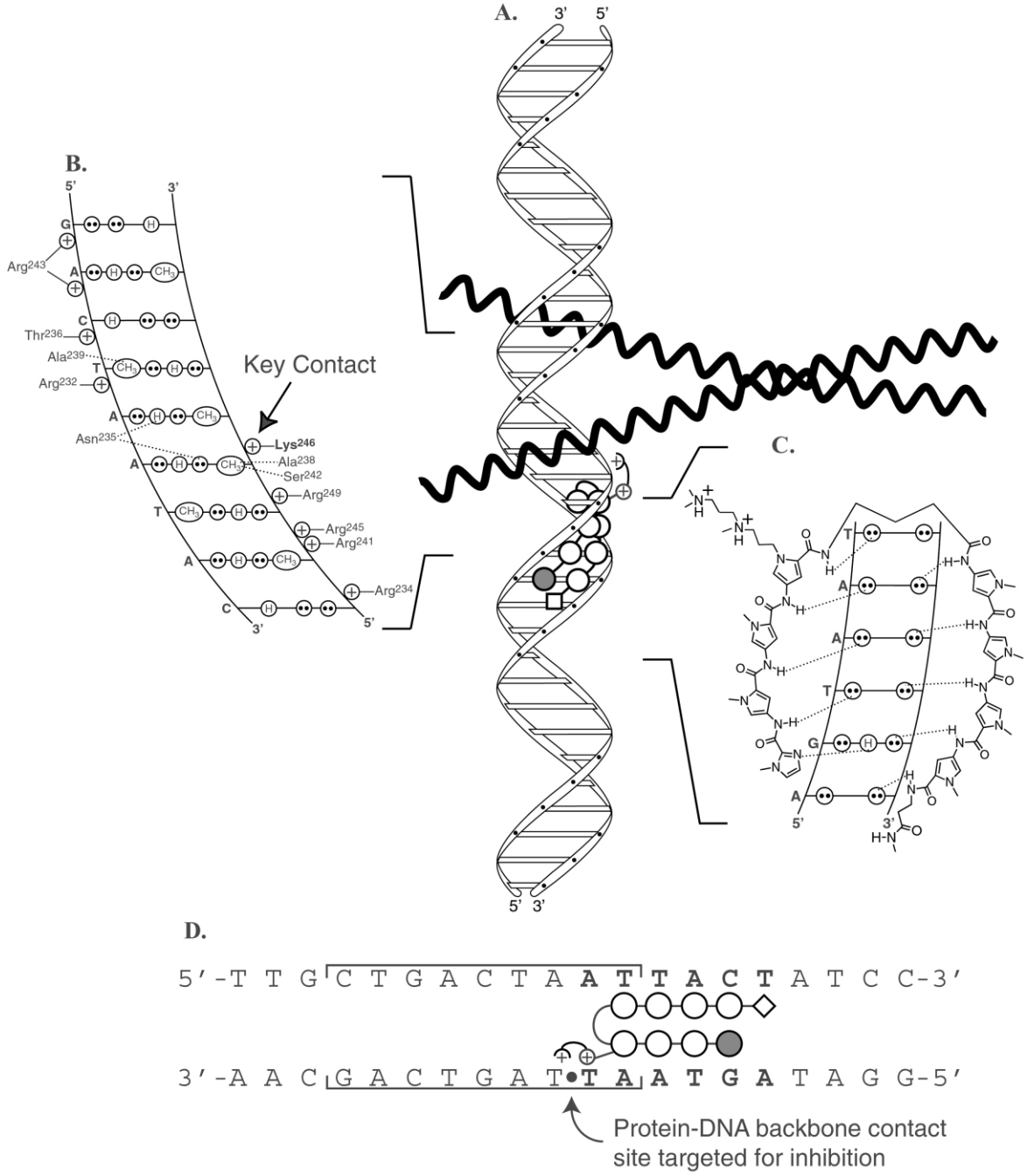


Figure 4. A schematic model of *N*-(aminoalkyl)pyrrole polyamides targeted to the major groove transcription factor, GCN4. (A) The α-helical GCN4 dimer is shown binding to adjacent major grooves.[11] The *N*-(aminoalkyl)pyrrole polyamide is shown as black and white balls which represent imidazole and pyrrole amino acids, respectively. The diamond represents β-alanine; γ-aminobutyric acid is designated as a curved line and the aminoalkyl chain is represented by a plus sign in a circle. (B) The contacts between one GCN4 monomer and the major groove of one half-site of 5′-CTGACTAAT-3′ are depicted (adapted from ref 11). Circles with two dots represent the lone pairs of the N7 of purines, the O4 of thymine, and the O6 of guanine. Circles containing an H represent the N6 and N4 hydrogens of the exocyclic amines of adenine and cytosine, respectively. The C5 methyl group of thymine is depicted as a circle with CH_3 inside. Plus signs represent protein residues which electrostatically contact the phosphate backbone. (C) The hydrogen bonding model of the eight-ring hairpin polyamide ImPyPyPy(DpDp)-γ-PyPyPyPy-β-Me bound to the minor groove of 5′-AGTAAT-3′. Circles with two dots represent the lone pairs of N3 of purines and O2 of pyrimidines. Circles containing an H represent the N2 hydrogens of guanines. Putative hydrogen bonds are illustrated by dotted lines. (D) The model of the polyamide binding its target site (bold) adjacent to the GCN4 binding site (brackets). Polyamide residues are as in A.

Waal contacts and phosphate interactions in the DNA major groove (Fig. 4A,B).[11,18] Gel mobility shift assays were performed as previously described.[6,8,17]

The polyamides of the sequence ImPyPyPy-γ-PyPyPyPy were designed to bind the 5′-AGTAAT-3′ site adjacent to the GCN4 binding site (5′-CTGACTAAT-3′) on the synthetic DNA duplex ARE-4. Aminoalkyl substitutions at the N-1 of the Py residue immediately to the N-terminal side of the γ-turn were designed to target the phosphate of the 3′-TpT-5′ step (Fig. 4). Incubation of the radiolabeled DNA fragment with the parent *N*-methyl containing polyamide ImPyPyPy-γ-PyPyPyPy-β-Dp (**1**) resulted in only a slight inhibition of protein binding, consistent with our earlier observation that standard Py/Im polyamides bind simultaneously with major groove binding bZIP proteins (Figure 5A). The placement of *N*′, *N*′-dimethylaminopropyl at the N-1 position of the specified pyrrole in **3** did not offer a significant improvement in GCN4 (222–281) inhibition over **1** (Figure 5B). However, incorporation of a second positive charge and a longer tether, as in **7**, provided 54% inhibition of GCN4 binding relative to the amount of DNA probe bound in the absence of polyamide (Figs

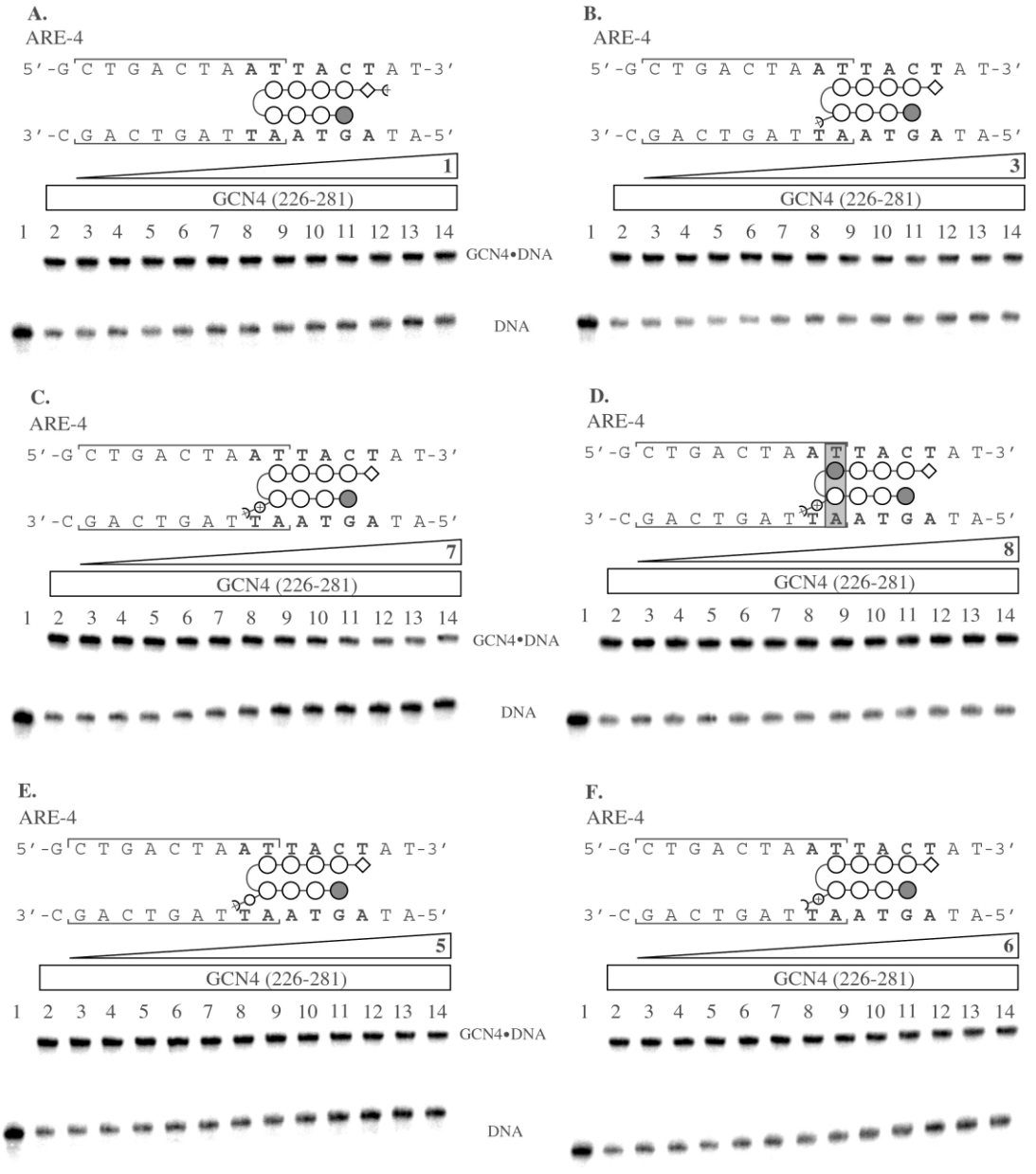


Figure 5. GCN4 (222–281) gel mobility shift experiments in the presence of *N*-methylpyrrole and *N*-aminoalkylpyrrole polyamides. (A–F) Top, model of polyamide binding the DNA fragment ARE-4. Mismatches are indicated by a grey box. Bottom, storage phosphor autoradiogram of non-denaturing polyacrylamide gel showing GCN4 (222–281) binding to the radiolabeled ARE-4 fragment in the presence of increasing concentrations of polyamide (10 mM bisTris pH 7.0, 100 mM NaCl, 1 mM DTT, 1 mM EDTA, 50 μg/mL poly(dI-dC)•poly(dI-dC), 22 °C). The upper band is the GCN4 (222–281)-DNA complex and the lower band is free DNA. Lane 1 is DNA only. Lane 2 contains DNA incubated with 200 nM GCN4 (222–281). Lanes 3–14 are 200 nM GCN4 (222–281) and 100 pM, 200 pM, 500 pM, 1 nM, 2 nM, 5 nM, 10 nM, 20 nM, 50 nM, 100 nM, 200 nM, and 500 nM polyamide: (A) ImPyPyPy-γ-PyPyPyPy-β-Dp (**1**), (B) ImPyPyPy(Dp)-γ-PyPyPyPy-β-Me (**3**), (C) ImPyPyPy(DpDp)-γ-PyPyPyPy-β-Me (**7**), (D) ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Me (**8**), (E) ImPyPyPy(C_3ODp)-γ-ImPyPyPy-β-Me (**5**), (F) ImPyPyPy(DpC_3OMe)-γ-ImPyPyPy-β-Me (**6**).

5C and 6). The specificity of GCN4 inhibition was also tested using ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Me (**8**), which is a single base pair mismatch for the polyamide site on ARE-4 and binds the DNA with significantly reduced affinity relative to **7**. Indeed, controls reveal that single mismatch polyamide **8** did not inhibit GCN4 binding (Fig. 5D).

Optimal GCN4 inhibition requires *N*-diaminoalkylpyrrole

It was important to determine if the basis for GCN4 inhibition by **7** was dependent upon the presence of both potentially charged amines or if perhaps the added length of the second aminopropyl was sufficient for inhibition. Polyamide **5** contains an ether linkage in place of the internal amine in **7**, resulting in a single charge placed distal to the polyamide. This change was sufficient to abolish GCN4 inhibition (Fig. 5E). Furthermore, **6**, containing a single charge internal and a terminal methyl ether, exhibited similar behavior with only slight inhibition on GCN4 binding (Fig. 5F). These results combined suggest that both of the positive charges present in **7** are necessary for inhibition of GCN4.

Directing the positive patch to adjacent phosphates

To ask whether GCN4 binding inhibition by polyamide **7** was specific for a particular phosophate residue, DNA probes were designed to shift the polyamide binding site by one or two base pairs relative to that in ARE-4. It was necessary to not alter the DNA sequence of the protein binding site in this new design. To accomplish this, the A/T-tract at the 3′ side of the GCN4 binding site and the degeneracy of Py/Py pairs for A•T and T•A base pairs was exploited. By simply shifting the G•C bp bound by the polyamide to the 3′ or 5′ side, probes were obtained that would deliver the polyamide positive patch to a different phosphate on the DNA backbone. ARE-3 and ARE-6 shift the polyamide to one or two base pairs deeper into the GCN4 binding site, respectively, while ARE-5 moves the polyamide one bp distal to the GCN4 cognate sequence, relative to ARE-4. Significant inhibition of GCN4 to any of these probes was not obtained at concentrations of **7** up to 20 times higher than that required for the maximum inhibition of GCN4 on ARE-4 (Fig. 7). On ARE-3, **7** showed a slight amount of GCN4 inhibition (Fig. 7C), while an apparent enhancement in GCN4 binding was observed on ARE-5 with as little as 100 nM **7** (Fig. 7A). This study reveals that specific targeting of a protein-DNA contact is essential for successful inhibition of bZIP protein binding to DNA.

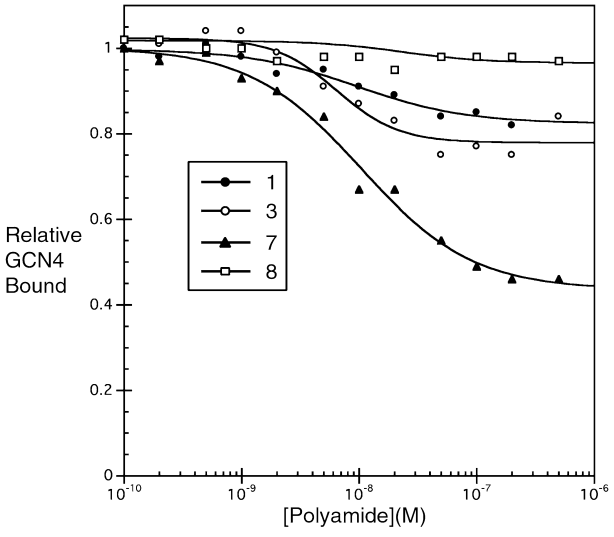


Figure 6. The level of GCN4 inhibition achieved by selected polyamides is compared. For each concentration of polyamide, the fraction of bound probe was determined by phosphorimager analysis followed by dividing the signal of the bound band by the sum of the bound and free bands. This value was normalized across experiments by dividing the value for each concentration by that for a control lane on the gel performed only with GCN4 (222–281). Typically, the fraction of bound probe in the absence of polyamide was ≈0.70. The optimal GCN4 inhibition observed for **7** is depicted with triangles. The slight inhibition provided by **1** and **3** is shown as unfilled and filled circles, respectively. The absence of inhibition with **8** is depicted with squares.

Discussion

A Py(Dp) residue increases polyamide affinity without compromising specificity

Generally, Py/Im polyamides are prepared using *N*-methyl aromatic amino acids and a dimethyalaminopropylamide tail on the C-terminus as in **1** and **2**. Exchanging the placement of the Dp and N-Me to provide a polyamide with a *N*-(*N*′, *N*′-dimethylaminopropyl)pyrrole residue and a C-terminal *N*-methyl amide (**3** and **4**) affords an isomer of identical molecular weight and composition; however, at least in this case, affinity for the DNA target site is increased 10-fold without compromising specificity. The affinity enhancement observed may result from a specific interaction between the positively charged alkyl chain and the phosphate oxygen or relief of unfavorable steric interactions of the dimethylaminopropyl tail at the C-terminus clashing with the floor of the minor groove, or a combination of these factors. Similar affinity enhancements were observed for each of the six aminoalkylpyrrole containing polyamides reported here. As the generality of the pairing rules are explored for the sequence specific recognition of DNA, the incorporation of *N*-aminoalkylpyrrole residues to increase affinity may be useful for DNA sequences that otherwise cannot be targeted with subnanomolar affinity.[19]

N-diaminoalkylpyrrole polyamides inhibit DNA binding by GCN4

The incorporation of a doubly positively charged alkyl chain on the N-1 of a pyrrole residue has afforded a new class of hairpin polyamides that can effectively inhibit DNA binding by an exclusively major groove binding bZIP protein. The presence of the positive patch is designed to compete with protein side chains that make electrostatic contacts to the phosphate backbone, potentially by partially neutralizing the negative charge of the phosphate. The maximal GCN4 inhibition

achieved with **7** is equivalent to that of the Arg-Pro-Arg analogue; however, it is observed at 10-fold lower polyamide concentration for **7** than for the Arg-Pro-Arg polyamide.[8] Furthermore, the *N*-diaminoalkyl moiety could be delivered from any ring in the polyamide, providing a versatility in design that allows for a polyamide target site selected for optimal recognition to be combined with the positive patch, potentially offering a greater number of targetable major groove proteins.

The optimal location of the positive patch for GCN4 inhibition is consistent with X-ray crystal structure data and phosphate ethylation interference experiments.[11,12] As described above, the positive patch of the polyamide on ARE-4 is likely directed to the phosphate of the 3′-TpT-5′ step. X-ray crystallography of GCN4 bound to a different DNA sequence reveals the residue Lys246 makes an electrostatic contact to the corresponding phosphate. Ethylation at the analogous phosphate on another GCN4 binding site resulted in a severe loss in affinity of GCN4 for the DNA target.[12] While structural studies show other electrostatic contacts between protein side chains and the phosphates on either side of the 3′-TpT-5′ step, ethylation of these phosphates has less of an effect on GCN4 binding affinity.[11,12] This latter observation is consistent with a lack of significant inhibition for **7** on ARE-3 and ARE-5. The slight inhibition observed for ARE-3 may be due to the flexible linker of **7** interacting with the 3′-TpT-5′ step shown to inhibit GCN4 binding. Ethylation of the phosphate corresponding to the 3′-GpT-5′ (ARE-6) reduced GCN4 affinity; however, no protein contacts to this phosphate have been identified by crystallography.[11,12] Taken together, these results suggest that the inhibition of a given protein: phosphate contact may not be universally equivalent in its effect on protein binding. Each contact plays a unique role in DNA recognition and inhibition of protein binding by positive patch polyamides may need to be optimized for each application.

The inability of polyamides containing ether linkages in place of the amines in **7** (**5** and **6**) to effectively inhibit GCN4 binding demonstrates at a minimum that an alkyl chain of this length is insufficient to obtain protein

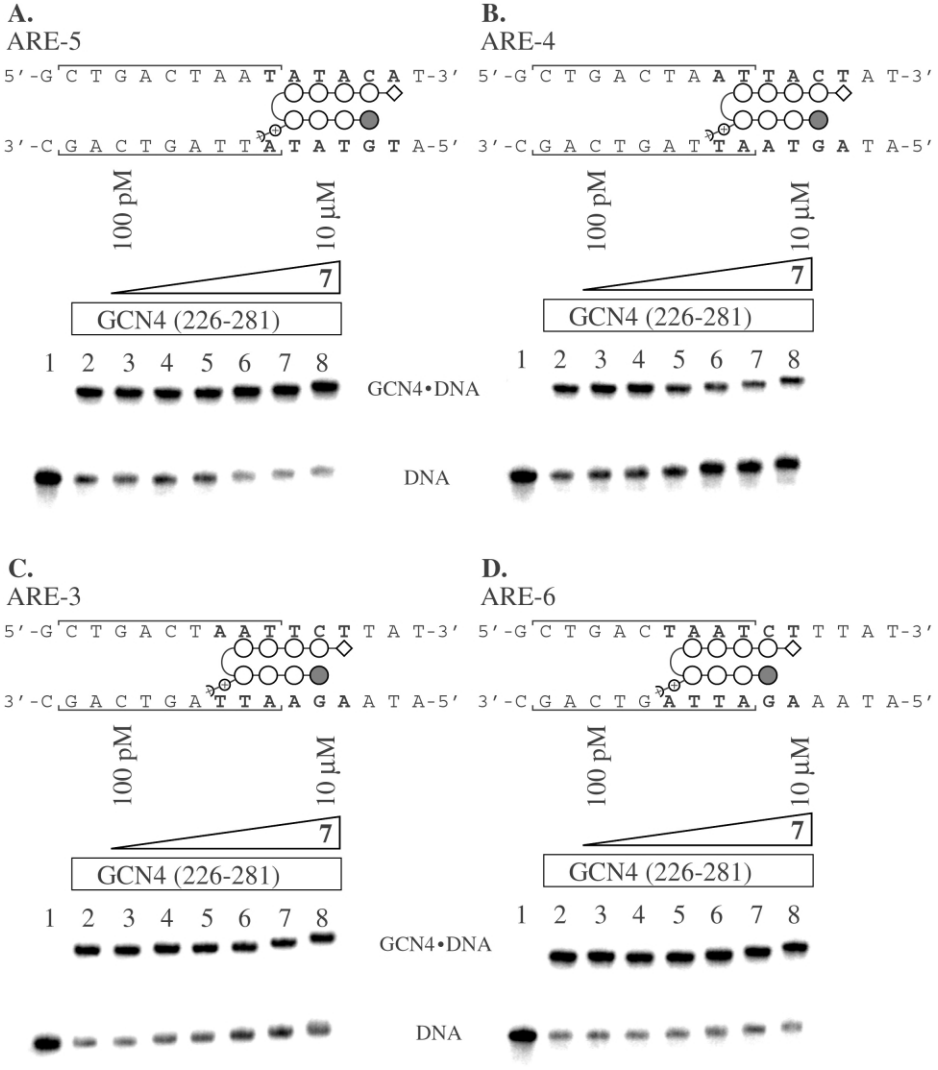


Figure 7. The effect of moving the polyamide binding site on GCN4 inhibition is investigated. Results are presented as in Figure 6. All panels include Lane 1: DNA only, Lane 2: 200 nM GCN4(222–281), Lane 3–8: 100 pM, 1 nM, 10 nM, 100 nM, 1 μM, and 10 μM **7**, respectively. (A) ARE-5 probe, (B) ARE-4 probe, (C) ARE-3 probe, (D) ARE-6 probe.

inhibition and that the amine nitrogens play a specific role in inhibiting GCN4 binding. The amines may be partially neutralizing one or more phosphates to interfere with a protein: DNA contact or potentially distort the DNA by phosphate neutralization.[20] If the protein were inhibited simply by steric bulk extending into the major groove one would have expected **5** and/or **6** to effectively inhibit the protein. Furthermore, the lack of inhibition provided by the mismatched polyamide **8** demonstrates that the mechanism of inhibition requires a polyamide specifically bound in the minor groove.

Conclusions

A new class of hairpin polyamides containing a positively charged alkyl amine at the N-1 of a pyrrole ring have been shown to sequence specifically recognize DNA with increased affinity and uncompromised specificity relative to analogues with the amine at the C-terminus. The positive patch provided by a *N*-diaminoalkylpyrrole residue proved reasonably effective in the inhibition of DNA binding by the major groove binding protein GCN4 (222–281), potentially by competing with the protein side chains for electrostatic contacts to the DNA phosphate backbone. However, there remains room for improvement and we consider this series to be a first generation design. Nevertheless, these positive patch polyamides may expand the functional repertoire of polyamides for use in chemical genetics.

Experimental

All synthetic reagents were as previously described or obtained from Aldrich.[15] Analytical HPLC was performed on a Beckman *Gold Nouveau* system with a model 126 pump and model 168 diode array detector. A Rainen C_{18}, Microsorb MV, 5 μm, 300×4.6 mm reverse phase column was employed with 0.1% (w/v) TFA:H_2O and 1.5% acetonitrile/min. Preparatory HPLC was performed on a Waters DeltaPak 25×100 mm 100 μm C_{18} column in 0.1% (w/v) TFA, gradient elution 0.25%/min. CH_3CN. Resin substitution of synthesized polyamides was calculated as $L_{new}(\text{mmol/g}) = L_{old}/(1 + L_{old}(W_{new} - W_{old}) \times 10^{-3})$ where L is the loading (mmol of amine per gram of resin) and W is the weight (g mol^{-1}) of the growing peptide attached to the resin.[15] DNA restriction fragment labeling, DNase I footprinting, and determination of equilibrium association constants were accomplished using previously described protocols.[21,22] Chemical sequencing reactions were performed according to published methods.[23,24]

Monomer synthesis

Ethyl (3-hydroxypropyl)-4-nitropyrrole-2-carboxylate (10). Potassium carbonate (22.8 g, 163 mmol) was added to ethyl 4-nitropyrrole-2-carboxylate (**9**) (10 g, 54.3 mmol) in acetone (740 mL) and the slurry stirred for 2 h, followed by the addition of 3-iodopropanol (23.5 g, 109 mmol) and refluxing for 2 h. The reaction was cooled to room temperature, filtered and concentrated in vacuo. Water (200 mL) was added to the resulting oil, the pH of the aqueous layer was reduced to 3 with 10% H_2SO_4, and the mixture extracted with ethyl acetate. The combined ethyl acetate extracts were dried ($MgSO_4$) and concentrated in vacuo. The solid was chromatographed with silica gel (ethyl acetate:hexanes) to provide a clear oil (12.8 g, 52.6 mmol, 99% yield): ^{1}H NMR (DMSO-d_6): δ 8.21 (s, 1H), 7.32 (s, 1H), 4.62 (t, 1H, J = 4.9 Hz), 4.47 (t, 2H, J = 6.8 Hz), 4.21 (q, 2H, J = 7.0 Hz), 3.31 (q, 2H, J = 6.1 Hz), 1.80 (m, 2H), 1.30 (t, 3H, J = 7.0 Hz).

(3-Hydroxypropyl)-4-[(*tert*-butoxycarbonyl)amino]-pyrrole-2-carboxylic acid (11). Ethyl (3-hydroxypropyl)-4-nitropyrrole-2-carboxylate (6.08 g, 25.1 mmol) was dissolved in DMF (20 mL). 10% Pd/C (1.0 g) was added and the mixture was stirred under hydrogen (300 psi) for 3 h. Pd/C was removed by filtering through Celite. Di-*tert*-butyl dicarbonate (5.41 g, 24.8 mmol) and DIEA (8.6 mL, 50 mmol) was added and stirred for 2 h. The reaction was cooled to room temperature, filtered, washed with methanol and concentrated in vacuo. Chromatography with silica gel (ethyl acetate:hexanes) provided a clear oil (5.10 g, 16.3 mmol, 65% yield): ^{1}H NMR (DMSO-d_6): δ 9.08 (s, 1H), 7.06 (s, 1H), 6.60 (s, 1H), 4.49 (t, 1H, J = 5.1 Hz), 4.20 (t, 2H, J = 6.8 Hz), 4.12 (q, 2H, J = 7.5 Hz), 3.30 (q, 2H, J = 5.5 Hz), 1.72 (m, 2H), 1.45 (s, 9H), 1.23 (t, 3H, J = 7.5 Hz). Ethyl (3-hydroxypropyl)-4-[(*tert*-butoycarbonyl)amino]-pyrrole-2-carboxylate (2.0 g, 6.4 mmol) was dissolved in ethanol (10 mL), followed by the addition of 1M KOH (40 mL) and heating (70 °C) for 2 h. The reaction was cooled to room temperature and washed with ethyl ether. The pH of the aqueous layer was reduced to 3 with 10% H_2SO_4, and the mixture extracted with diethyl ether. The combined ethyl acetate extracts were dried ($MgSO_4$) and concentrated in vacuo to provide a white, crystalline solid (1.5 g, 5.3 mmol, 83% yield): ^{1}H NMR (DMSO-d_6): δ 12.04 (s, 1H), 9.05 (s, 1H), 7.04 (s, 1H), 6.54 (s, 1H), 4.20 (t, 2H, J = 6.8 Hz), 3.27 (q, 2H, J = 6.1 Hz), 1.73 (m, 2H), 1.40 (s, 9H).

Polyamide synthesis

ImPyPyPy(C_3OH)-γ-PyPyPyPy-β-Me (12). ImPyPyPy(C_3OH)-γ-PyPyPyPy-β-PAM-resin was synthesized in a stepwise fashion by machine-assisted solid phase methods from Boc-β-PAM-resin (600 mg, 0.75 mmol/g). (3-hydroxylpropyl)-4-[(*tert*-butoxycarbonyl)amino]-pyrrole-2-carboxylic acid (212 mg, 0.75 mmol), HBTU (284 mg, 0.75 mmol) and HOBt (304 mg, 2.25 mmol) were added dry to a synthesis cartridge. Upon delivery of DMF and DIEA activation occurs as described. A sample of polyamide resin (380 mg, 0.43 mmol/g) was cleaved with methylamine (55 °C, 28 h) in a Parr apparatus. The methylamine was allowed to evaporate and the crude polyamide was redissolved in 1:1 1M NH_4OH:CH_3CN (20 mL), filtered to remove the resin and lyophilized dry (146 μmol).

ImPyPyPy(C_3OH)-γ-ImPyPyPy-β-Me (13). ImPyPyPy(C_3OH)-γ-ImPyPyPy-β-PAM-resin was synthesized in a stepwise fashion by machine-assisted solid phase

methods from Boc-β-PAM-resin (450 mg, 0.75 mmol/g) as described above. A sample of polyamide resin (178 mg, 0.43 mmol/g) was cleaved with methylamine (55 °C, 18 h) in a Parr apparatus. The methylamine was allowed to evaporate and the crude polyamide was redissolved in 1:1 1M NH_4OH/CH_3CN (20 mL), filtered to remove the resin and lyophilized dry (56 μmol).

ImPyPyPy(Dp)-γ-ImPyPyPy-β-Me (3). Crude ImPyPyPy(C_3OH)-γ-ImPyPyPy-β-Me (25 μmol) was dissolved in pyridine (600 μL) followed by the addition of methanesulfonyl chloride (100 μL). After 30 min, the reaction was precipitated with ethyl ether (1 mL), decanted, and redissolved in DMF (1 mL). 1:1 Dimethylamine (2.0 M in THF)/DMF (3 mL) was added and the reaction allowed to proceed for 45 min at 90 °C. The reaction was diluted to 8 mL with 0.1% (w/v) TFA and purified by preparatory reverse phase HPLC to afford ImPyPyPy(Dp)-γ-ImPyPyPy-β-Me (**3**) upon lyophilization of the appropriate fractions (1.3 mg, 1.0 μmol, 4.0% recovery); UV (H_2O) λ_{max} 312 (66,000); ^{1}H NMR (DMSO-d_6): δ 10.45 (s, 1H), 10.27 (s, 1H), 9.95 (s, 3H), 9.93 (s, 1H), 9.88 (s, 1H), 9.27 (m, 1H), 8.19 (m, 1H), 8.00 (m, 1H), 7.79 (m, 1H), 7.42 (s, 1H), 7.36 (s, 1H), 7.30 (s, 1H,), 7.25 (s, 1H), 7.22 (s, 1H), 7.19 (s, 1H), 7.15 (s, 1H), 7.13 (s, 1H), 7.10 (s, 1H), 7.05 (s, 1H), 7.00 (s, 2H), 6.87 (s, 1H), 6.78 (s, 1H), 4.26 (m, 2H), 3.94 (s, 3H), 3.91 (s, 3H), 3.80 (s, 12H), 3.75 (s, 3H), 3.35 (m, 2H), 3.20 (m, 2H), 2.97 (m, 2H), 2.74 (d, 6H, J = 4.8 Hz), 2.52 (d, 3H, J = 4.5 Hz), 2.29 (m, 4H), 2.02 (m, 2H), 1.75 (m, 2H); MS (MALDI-TOF): $[M+H]^+$ 1222.6, $[M+Na]^+$ 1244.5, $[M+K]^+$ 1260.5. ($C_{58}H_{72}N_{21}O_{10}^+$, $[M+H]^+$, calcd. 1222.6; $C_{58}H_{71}N_{21}NaO_{10}^+$, $[M+Na]^+$, calcd 1244.6; $C_{58}H_{71}KN_{21}O_{10}^+$, $[M+K]^+$, calcd 1260.5).

ImPyPyPy(Dp)-γ-PyPyPyPy-β-Me (4). Crude ImPyPyPy(C_3OH)-γ-PyPyPyPy-β-Me (49 μmol) was dissolved in pyridine followed by the addition of toluenesulfonyl chloride (2.2 M in pyridine, 350 μL) and kept at 0 °C for 10 h. After the addition of dimethylamine (2.0 M in THF, 2.0 mL), the reaction was heated at increasing temperatures until it was determined complete by analytical HPLC after 5 h at 80 °C. The reaction mixture was diluted to 8 mL with 0.1% (w/v) TFA and purified by preparatory reverse phase HPLC to provide ImPyPyPy(Dp)-γ-PyPyPyPy-β-Me (**4**) upon lyophilization of the appropriate fractions (5.8 mg, 4.8 μmol, 9.8% recovery); UV (H_2O) λ_{max} 312 (66,000); ^{1}H NMR (DMSO-d_6): δ 10.46 (s, 1H), 9.96 (s, 1H), 9.95 (s, 1H), 9.92 (s, 1H), 9.89 (s, 1H), 9.88 (s, 1H), 9.83 (s, 1H), 9.2 (br s, 1H), 8.18 (m, 1H), 7.99 (m, 1H), 7.80 (m, 1H), 7.38 (s, 1H), 7.31 (d, 1H, J = 1.8 Hz), 7.26 (d, 1H, J = 1.8 Hz), 7.20 (d, 1H, J = 1.5 Hz), 7.16 (m, 3H), 7.06 (d, 1H, J = 1.8 Hz), 7.03 (s, 1H), 7.02 (s, 1H), 7.01 (s, 1H), 6.89 (d, 1H, J = 1.5 Hz), 6.84 (d, 1H, J = 1.5 Hz), 6.79 (d, 1H, J = 1.5 Hz), 4.24 (m, 2H), 3.96 (s, 3H), 3.82 (s, 9H), 3.81 (s, 3H), 3.80 (s, 3H), 3.76 (s, 3H), 3.35 (m, 2H), 3.20 (m, 2H), 3.00 (m, 2H), 2.74 (d, 6H, J = 5.1 Hz), 2.53 (d, 3H, J = 4.5 Hz), 2.27 (m, 2H), 2.03 (m, 2H), 1.97 (m, 2H), 1.78 (m, 2H); MS (MALDI-TOF): $[M+H]^+$ 1221.6. ($C_{59}H_{73}N_{20}O_{10}^+$, $[M+H]^+$, calcd 1221.6).

ImPyPyPy(C_3ODp)-γ-PyPyPyPy-β-Me (5). Crude ImPyPyPy(C_3OH)-γ-PyPyPyPy-β-Me (24 μmol) in pyridine (0.4 mL) at 0 °C was activated with MsCl (75 μL) at 0 °C for 3 h, followed by precipitation with ether (1 mL). The supernatant was redissolved in DMF (0.5 mL) and cooled to 0 °C. *N*,*N*-dimethyl-3-amino-propanol (14 μL, 120 μmol) was added to NaH (60% dispersion in mineral oil, 4.7 mg, 114 μmol) in DMF (150 μL) and added to the reaction. Repeating the addition of NaH and *N*,*N*-dimethyl-3-amino-propanol and warming to room temperature did not provide sufficient reaction after 27 h as determined by analytical HPLC. More NaH (60% dispersion in mineral oil, 60 mg, 1.5 mmol) and *N*,*N*-dimethyl-3-amino-propanol (180 μL, 1.5 mmol) was added, followed by reaction at room temperature for 22 h. The reaction mixture was diluted to 8 mL with 0.1% (w/v) TFA and purified by preparatory reverse phase HPLC to afford ImPyPyPy(C_3ODp)-γ-PyPyPyPy-β-Me (**5**) upon lyophilization of the appropriate fractions (1.4 mg, 1.1 μmol, 4.6% recovery); UV (H_2O) λ_{max} 312 (66,000); ^{1}H NMR (DMSO-d_6): δ 10.43 (s, 1H), 9.94 (s, 1H), 9.93 (s, 1H), 9.90 (s, 1H), 9.87 (s, 2H), 9.81 (s, 1H), 9.25 (m, 1H), 8.16 (m, 1H), 7.97 (m, 1H), 7.77 (m, 1H), 7.35 (s, 1H), 7.32 (s, 1H), 7.30 (s, 1H,), 7.24 (s, 1H), 7.18 (s, 3H), 7.14 (s, 1H), 7.13 (s, 2H), 7.05 (s, 1H), 7.00 (s, 2H), 6.88 (s, 1H), 6.82 (s, 1H), 6.77 (s, 1H), 4.25 (m, 2H), 3.94 (s, 3H), 3.80 (s, 15H), 3.75 (s, 3H), 3.34 (m, 4H), 3.20 (m, 2H), 3.00 (m, 2H), 2.72 (d, 6H, J = 4.8 Hz), 2.52 (d, 3H, J = 4.2 Hz), 2.24 (m, 6H), 2.02 (m, 2H), 1.76 (m, 4H); MS (MALDI-TOF): $[M+H]^+$ 1279.7. ($C_{62}H_{79}N_{20}O_{11}^+$, $[M+H]^+$, calcd 1279.6).

ImPyPyPy(DpC$_3$OMe)-γ-PyPyPyPy-β-Me (6). MsCl activation of crude ImPyPyPy(C_3OH)-γ-PyPyPyPy-β-Me (24 μmol) as for **4** followed by redissolving of the precipitate in DMF (3 mL), reaction with 3-methoxypropylamine (1 mL, 90 °C, 30 min.) and preparatory reverse phase HPLC afforded ImPyPyPy(DpC$_3$OMe)-γ-PyPyPyPy-β-Me (**6**) upon lyophilization of the appropriate fractions (1.5 mg, 1.2 μmol, 5.0% recovery); UV (H_2O) λ_{max} 312 (66,000); ^{1}H NMR (DMSO-d_6): δ 10.44 (s, 1H), 9.96 (s, 1H), 9.95 (s, 1H), 9.91 (s, 1H), 9.88 (s, 1H), 9.88 (s, 1H), 9.83 (s, 1H), 8.24 (m, 2H), 8.20 (m, 1H), 7.99 (m, 1H), 7.80 (m, 1H), 7.37 (d, 1H, J = 0.6 Hz), 7.32 (d, 1H, J = 1.8 Hz), 7.26 (d, 1H, J = 1.8 Hz), 7.20 (m, 3H), 7.16 (m, 3H), 7.06 (d, 1H, J = 1.8 Hz), 7.01 (m, 3H), 6.89 (d, 1H, J = 1.5 Hz), 6.84 (d, 1H, J = 1.8 Hz), 6.79 (d, 1H, J = 2.1 Hz), 4.30 (m, 2H), 3.96 (s, 3H), 3.82 (s, 9H), 3.81 (s, 3H), 3.81 (s, 3H), 3.76 (s, 3H), 3.19 (s, 3H), 2.90 (m, 4H), 2.52 (m, 5H), 2.42 (m, 2H), 2.27 (m, 6H), 2.00 (m, 2H), 1.78 (m, 4H); MS (MALDI-TOF): $[M+H]^+$ 1265.8, $[M+Na]^+$ 1287.8, $[M+K]^+$ 1303.8. ($C_{61}H_{76}N_{20}O_{11}^+$, $[M+H]^+$, calc. 1265.6; $C_{61}H_{76}N_{20}NaO_{11}^+$, $[M+Na]^+$, calcd 1287.6; $C_{61}H_{76}KN_{20}O_{11}^+$, $[M+K]^+$, calcd 1303.6).

ImPyPyPy(DpDp)-γ-PyPyPyPy-β-Me (7). MsCl activation of crude ImPyPyPy(C_3OH)-γ-PyPyPyPy-β-Me (23 μmol) as for **4** followed by redissolving of the precipitate in pyridine (0.75 mL), reaction with *N*,*N*,*N*′-trimethyl-1,3-propanediamine (0.75 mL, 90 °C, 40 min) and preparatory reverse phase HPLC afforded ImPy-

PyPy(DpDp)-γ-PyPyPyPy-β-Me (**7**) upon lyophilization of the appropriate fractions (6.7 mg, 5.2 μmol, 23% recovery); UV (H_2O) λ_{max} 312 (66,000); ^1H NMR (DMSO-d_6): δ 10.45 (s, 1H), 9.97 (s, 1H), 9.95 (s, 1H), 9.92 (s, 1H), 9.89 (s, 1H), 9.88 (s, 1H), 9.84 (s, 1H), 9.55 (m, 1H), 8.44 (m, 1H), 8.19 (m, 1H), 7.98 (m, 1H), 7.80 (m, 1H), 7.37 (m, 1H), 7.34 (d, 1H, $J=1.8$ Hz), 7.26 (d, 1H, $J=1.8$ Hz), 7.20 (s, 3H), 7.15 (m, 2H), 7.07 (d, 1H, $J=1.5$ Hz), 7.01 (m, 3H), 6.88 (d, 1H, $J=1.8$ Hz), 6.84 (d, 1H, $J=1.8$ Hz), 6.79 (d, 1H, $J=1.8$ Hz), 4.30 (m, 2H), 3.96 (s, 3H), 3.82 (s, 9H), 3.81 (s, 6H), 3.80 (s, 3H), 3.35 (m, 4H), 3.20 (m, 2H), 3.00 (m, 2H), 2.75 (m, 9H), 2.54 (d, 3H, $J=4.5$ Hz), 2.26 (m, 6H), 2.04 (m, 2H), 1.96 (m, 2H), 1.77 (m, 2H); MS (MALDI-TOF): $[M+H]^+$ 1292.7 ($C_{63}H_{82}N_{21}O_{10}^+$, $[M+H]^+$, calcd 1292.7).

ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Me (8). Crude ImPyPyPy(C_3OH)-γ-ImPyPyPy-β-Me (23 μmol) was dissolved in pyridine (150 μL) and activated with TsCl (4.8 M in pyridine, 200 μL) for 1 h at 0 °C. The addition of *N*,*N*,*N*′-trimethyl-1,3-propanediamine (0.30 mL, 55 °C, 30 min.) followed by preparatory reverse phase HPLC afforded ImPyPyPy(DpDp)-γ-ImPyPyPy-β-Me (**8**) upon lyophilization of the appropriate fractions (0.4 mg, 0.3 μmol, 1.3% recovery); UV (H_2O) λ_{max} 312 (66,000); ^1H NMR (DMSO-d_6): δ 10.43 (s, 1H), 10.26 (s, 1H), 9.96 (s, 1H), 9.93 (s, 1H), 9.91 (s, 2H), 9.86 (s, 1H), 9.45 (m, 1H), 8.17 (m, 1H), 7.97 (m, 1H), 7.78 (m, 1H), 7.41 (s, 1H), 7.36 (s, 2H), 7.33 (s, 1H,), 7.25 (s, 1H), 7.22 (s, 1H), 7.18 (s, 2H), 7.14 (s, 1H), 7.14 (s, 1H), 7.11 (s, 1H), 7.00 (s, 2H), 6.86 (s, 1H), 6.77 (s, 1H), 4.24 (m, 2H), 3.96 (s, 3H), 3.91 (s, 3H), 3.80 (s, 12H), 3.75 (s, 3H), 3.35 (m, 4H), 3.20 (m, 2H), 3.00 (m, 2H), 2.74 (d, 9H, $J=3.6$ Hz), 2.52 (d, 3H, $J=4.8$ Hz), 2.27 (m, 6H), 2.05 (m, 2H), 1.94 (m, 2H), 1.78 (m, 2H); MS (MALDI-TOF): $[M+H]^+$ 1293.9 ($C_{62}H_{81}N_{22}O_{10}^+$, $[M+H]^+$, calcd 1293.7).

Acknowledgements

The authors are grateful to the National Institutes of Health for research support, the National Science Foundation for a predoctoral fellowship to R.E.B., Bristol-Myers Squibb and the Ralph M. Parson Foundation for predoctoral fellowships to R.E.B. and N.R.W., and the National Institutes of Health for a research service award to J.W.S. We thank G.M. Hathaway and the Caltech Protein/Peptide Microanalytical Laboratory for MALDI-TOF mass spectrometry.

References and Notes

1. Dervan, P. B.; Bürli, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688.
2. Gottesfeld, J. M.; Turner, J. M.; Dervan, P. B. *Gene Expression* **2000**, *9*, 77.
3. White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569.
4. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468.
5. Dickinson, L. A.; Gulizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12890.
6. Oakley, M. G.; Mrksich, M.; Dervan, P. B. *Biochemistry* **1992**, *31*, 10969.
7. Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B.; Gottesfeld, J. M. *J. Mol. Biol.* **1997**, *274*, 439.
8. Bremer, R. E.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1998**, *5*, 119.
9. Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104.
10. Bruice, T. C.; Sengupta, D.; Blasko, A.; Chiang, S. Y.; Beerman, T. A. *Bioorg. Med. Chem.* **1997**, *5*, 685.
11. Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* **1992**, *71*, 1223.
12. Gartenberg, M. R.; Ampe, C.; Steitz, T. A.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 6034.
13. Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559.
14. Bremer, R. E.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B. *Bioorg. Med. Chem.* **2000**, *8*, 1947.
15. Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.
16. Hope, I. A.; Struhl, K. *Cell* **1986**, *46*, 885.
17. Oakley, M. G.; Dervan, P. B. *Science* **1990**, *248*, 847.
18. Konig, P.; Richmond, T. J. *J. Mol. Biol.* **1993**, *233*, 139.
19. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 6953.
20. Maher III, L. J. *Curr. Opin. Chem. Biol.* **1998**, *2*, 688.
21. Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 7636.
22. Swalley, S. E.; Baird, E. E.; Dervan, P. B. *Chem.-Eur. J.* **1997**, *3*, 1600.
23. Maxam, A. M.; Gilbert, W. S. *Methods Enzymol.* **1980**, *65*, 499.
24. Iverson, B. L.; Dervan, P. B. *Nucleic Acids Res.* **1987**, *15*, 7823.