Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry 15 (2007) 474–483
Bioorganic & Medicinal Chemistry

Synthesis and biophysical evaluation of minor-groove binding C-terminus modified pyrrole and imidazole triamide analogs of distamycin

Toni Brown,[a,b] Zarmeen Taherbhai,[b] Jim Sexton,[b] Arden Sutterfield,[b] Mark Turlington,[b] Justin Jones,[b] Lindsay Stallings,[b] Michelle Stewart,[b] Karen Buchmueller,[b,c] Hilary Mackay,[a] Caroline O'Hare,[d] Jerome Kluza,[d] Binh Nguyen,[e] David Wilson,[e] Moses Lee[a,b,*] and John A. Hartley[d]

[a]*Division of Natural Sciences and Department of Chemistry, Hope College, Holland, MI 49423, USA*
[b]*Department of Chemistry, Furman University, Greenville, SC 29613, USA*
[c]*Department of Chemistry, Wake Forest University, Winston-Salem, NC 27109, USA*
[d]*Department of Oncology, Royal Free & University College Medical School, London, W1W 7BS, UK*
[e]*Department of Chemistry, Georgia State University, Atlanta, GA 30302, USA*

Received 18 July 2006; revised 13 September 2006; accepted 20 September 2006
Available online 10 October 2006

Abstract—Five polyamide derivatives with rationally modified C-terminus moieties were synthesized and their DNA binding specificity and affinity determined. A convergent approach was employed to synthesize polyamides containing an alkylaminopiperazine (**4** and **5**), a truncated piperazine (**6**), or an alkyldiamino-C-terminus moiety (**7** and **8**) with two specific objectives: to investigate the effects of number of potential cationic centers and steric bulk at the C-terminus. CD studies confirmed that compounds **4**, **5**, **7**, and **8** bind in the minor groove of DNA. The alkylpiperazine containing compounds (**4** and **5**) showed only moderate binding to DNA with ΔT_m values of 2.8 and 8.3 °C with their cognate sequence, respectively. The alkyldiamino compounds (**7** and **8**) were more impressive producing a ΔT_m of >17 and >22 °C, respectively. Compound **6** (truncated piperazine) did not stabilize its cognate DNA sequence. Footprints were observed for all compounds (except compound **6**) with their cognate DNA sequence using DNase I footprinting, with compound **7** producing a footprint of 0.1 μM at the expected 5′-ACGCGT-3′ site. SPR analysis of compound **7** binding to 5′-ACGCGT-3′, 5′-ACCGGT-3′, and 5′-AAATTT-3′ produced binding affinities of 2.2×10^6, 3.3×10^5, and 1×10^5 M^{-1}, respectively, indicating a preference for its cognate sequence of 5′-ACGCGT-3′. These results are in good agreement with the footprinting data. The results indicate that steric crowding at the C-terminus is important with respect to binding. However, the number of cationic centers within the molecule may also play a role. The alkyldiamino-containing compounds (**7** and **8**) warrant further investigation in the field of polyamide research.

1. Introduction

The field of polyamide DNA interactive ligands has been an area of interest since the discovery that distamycin (**1**, Fig. 1), a naturally occurring polyamide, interacts with A/T-rich regions of duplex DNA in the minor groove as an anti-parallel stacked dimer.[1] Since that time many research groups have strived to create more selective and specific polyamide molecules in the hope of recognizing G/C base pairs alongside A/T.[2–7]

Extensive structural and biophysical studies have identified three main components that affect the molecular recognition and binding affinity of triamide molecules to duplex DNA: the N-terminus moiety, heterocycle type and order, and the nature of the C-terminus. The formamido (f-) group at the N-terminus of distamycin (**1**) and related triamides has been shown by several groups, including the authors' laboratory, to play a major role in controlling the staggered binding motif of the 2:1 ligand/DNA complex.[8] In addition, binding affinity is increased 100-fold compared with a non-formamido

Keywords: Amine; Biophysical; C-terminus; Distamycin; DNA; Polyamides.
* Corresponding author. Tel.: +1 616 395 7190; fax: +1 616 395 7923; e-mail: lee@hope.edu

0968-0896/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2006.09.037

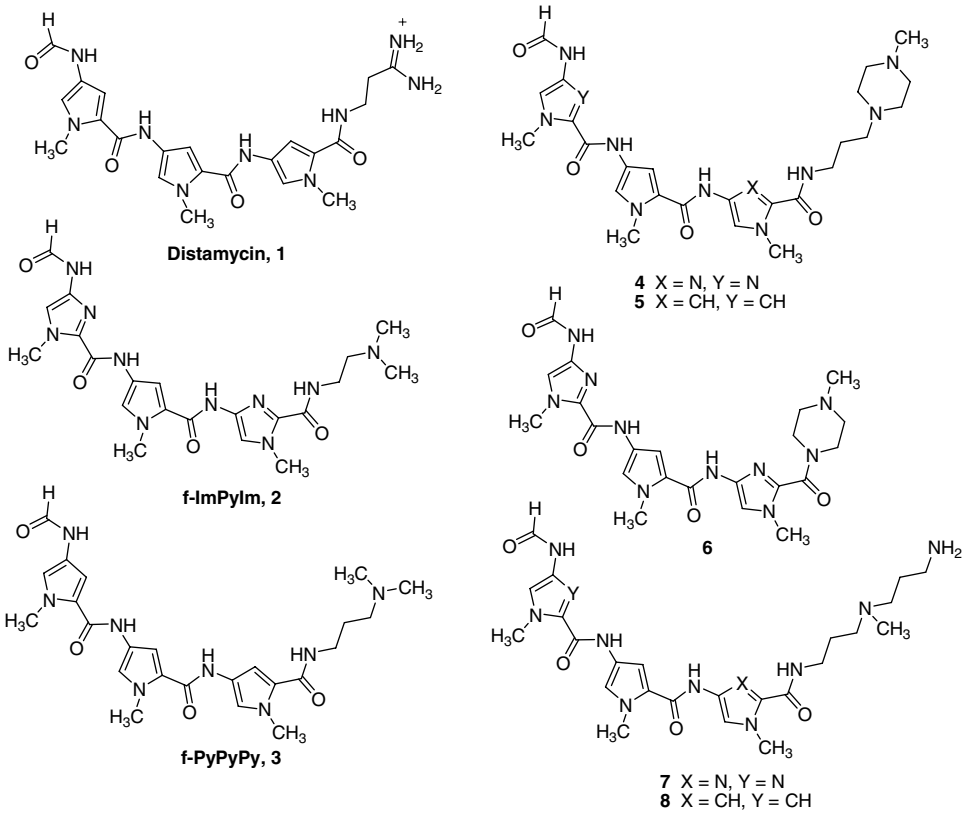


Figure 1. Structures of distamycin (**1**), f-ImPyIm (**2**), f-PyPyPy (**3**), and synthesized C-terminus modified compounds **4**–**8**.

equivalent.[8,9] This staggered dimeric binding to duplex DNA enabled a series of binding rules to be developed.[10–14] The order of heterocycles and the 'core' of the triamide unit are important with respect to binding affinity and selectivity. Work performed in the authors' laboratory[15] on all eight triamides containing the various combinations of imidazole and pyrrole showed that the following rank order of binding could be assigned if the two central units were considered: Imidazole(Im)Pyrrole(Py) > PyPy ≫ ImPy ∼ ImIm. As a result of these findings it was discovered that f-ImPyIm (**2**, Fig. 1) has the greatest binding affinity to its respective cognate sequence of 5′-T/A CGCG A/T-3′.[15] This molecule contains an *N*,*N*-dimethylaminopropyl C-terminus moiety which differs from the propanamidine found in the tripyrrole compound, distamycin (**1**). In comparison f-PyPyPy (**3**, Fig. 1), which also contains an *N*,*N*-dimethylaminopropyl C-terminus moiety, binds to 5′-A_3T_3-3′ with a lower binding affinity than distamycin (**1**), indicating that the amidine moiety offers excellent DNA interacting properties. Unfortunately, the amidine group is hygroscopic and is often difficult to handle. Accordingly, there is a need to develop polyamides with synthetically feasible C-terminus groups with comparable binding to alkylamidine.[16]

Compounds reported in this article were synthesized with the following heterocycle order: f-ImPyIm and f-PyPyPy, as compounds of this size have been shown to possess the strongest binding affinity with their respective cognates according to the 'core-binding-rules' study, mentioned above.[15] An alkylpiperazine moiety was chosen to extend the length of the polyamide at the C-terminus and increase the span of base-pair recognition. The methylpiperazine unit was incorporated into the target compounds in this study because it was demonstrated to be a pivotal factor in enhancing the DNA binding affinity in Hoechst 33258.[17–19] Moreover, the methylpiperazine moiety has the potential to become dicationic (pK_as 9.83 and 5.56[20]) at pH 6.4 used in these studies. To determine if the length of the C-terminus group affected the binding affinity, a truncated compound was synthesized in which the methylpiperazine group was attached directly to the polyamide via the carboxamide. To further investigate the effect of C-terminus length, compounds were synthesized by incorporating an *N*′*N*-*bis*-(3-aminopropyl)-*N*-methylamine group. These compounds were designed to resemble the original dimethylaminoethyl group found in f-ImPyIm (**2**), as they did not contain the bulky piperazine moiety but should possess two cationic centers felt necessary to retain water solubility.

Five compounds were synthesized employing the three different C-terminus groups and two combinations of heterocycles (f-ImPyIm and f-PyPyPy): two alkylpiperazine compounds (**4**, **5**, Fig. 1), a truncated piperazine compound (**6**, Fig. 1), and two alkyldiamino compounds (**7**, **8**, Fig. 1). Biophysical and biological data were obtained from circular dichroism (CD), thermal melting (T_m), surface plasmon resonance (SPR), and DNase I footprinting. All compounds were tested against two synthetic DNA hairpins, which contained either the 5′-ACGCGT-3′ or 5′-A_3T_3-3′ sequences. The former is

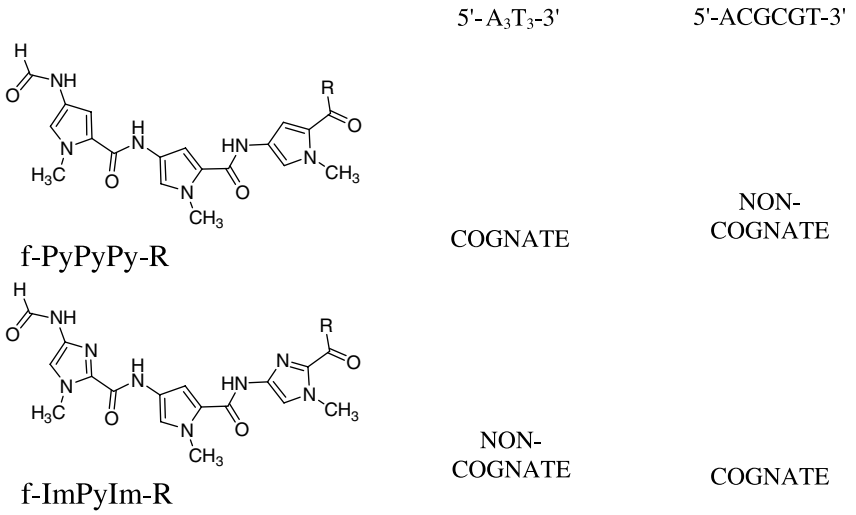


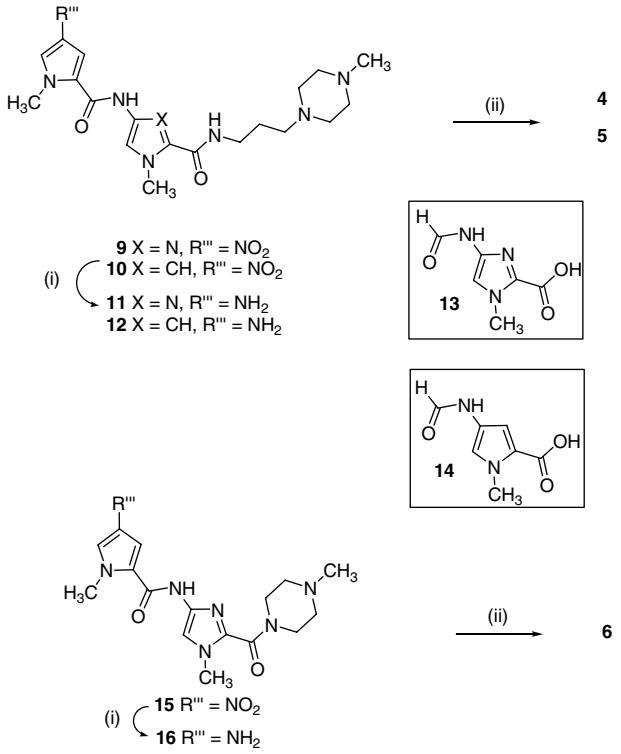
Figure 2. Diagram to show the cognate and non-cognate DNA for the compounds described in this study. (R = any C-terminus group.)

the cognate site for all compounds containing the heterocycle sequence of f-ImPyIm, and the latter sequence the non-cognate (Fig. 2). Cognate sequence identity is reversed in the case of the f-PyPyPy compounds (Fig. 2). All compounds were compared to f-ImPyIm (**2**), f-PyPyPy (**3**), and distamycin (**1**).[8,15]

2. Results and discussion

2.1. Synthesis

Scheme 1 represents the synthesis of the alkylpiperazine compounds (**4** and **5**) and the truncated piperazine compound (**6**). Compounds **9** and **10** were converted to the corresponding amino-derivatives, **11** and **12** (catalytic hydrogenation), and reacted with acid **13** or **14** using standard carbodiimide chemistry to produce alkylpiperazine compounds **4** and **5**, respectively. The truncated piperazine compound (**6**) was produced in a similar manner; compound **15** was reduced and subsequent reaction of **16** with acid **13** produced the final compound **6**.



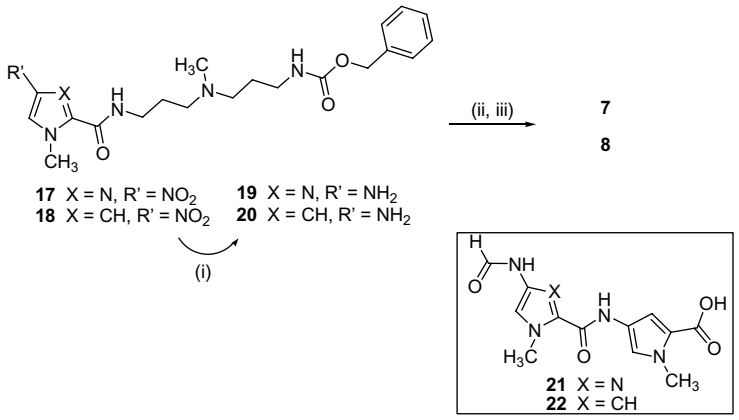
Scheme 1. Reagents and conditions: (i) 5% Pd/C, cold MeOH, H_2, ~18 h; (ii) **13** or **14** EDCI, DMAP, dry DMF, RT, ~96 h.

A slightly different strategy was employed for the formation of the alkyldiamino compounds (**7** and **8**) and is described in Scheme 2. The nitro-groups on compounds **17** and **18** were selectively reduced using PtO_2 catalyzed hydrogenation to produce amino-compounds **19** and **20**, respectively. These amines were then coupled with acids **21** and **22** using a DMAP catalyzed DCC coupling reaction to form the respective benzyl-protected triamide intermediates which were then de-benzylated via catalytic hydrogenation using 10% Pd/C to form the target compounds **7** and **8**.

2.2. DNase I footprinting

The sequence specificity of the C-terminus modified compounds was investigated by DNase I footprinting studies using a 130 bp 5′-[^{32}P]-radiolabeled fragment containing unique sequences 5′-ACGCGT-3′ (a), 5′-ACCGGT-3′ (b), 5′-A_3T_3-3′ (c), 5′-ACACCT-3′ (d), and 5′-AGCGCT-3′ (e) (Fig. 3). Sites (a) and (c) are the cognate sites for the f-ImPyIm and f-PyPyPy containing compounds, respectively. These studies were performed using f-ImPyIm (**2**) as control, which gave a distinct footprint at the 5′-ACGCGT-3′ site (a) at a concentration of 0.05 μM.

Protection of DNase I mediated cleavage (footprint) by the alkylpiperazine compounds was observed at 1 μM for compound **4** (f-ImPyIm-alkylpiperazine) with its cognate sequence (site (a), Fig. 4D); compound **5**



Scheme 2. Reagents and conditions: (i) PtO_2, cold MeOH, H_2, 50 psi, 1.5 h; (ii) **21** or **22**, DCC, DMAP, dry DMF, RT ~96 h; (iii) 10% Pd/C, cold MeOH, H_2, ~18 h.

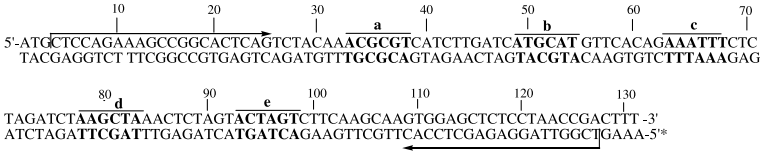


Figure 3. DNA sequence used to perform DNase1 footprinting showing the DNA sites of interest (a–e).

(f-PyPyPy-alkylpiperazine) showed binding at 5 μM with its cognate sequence 5′-A_3T_3-3′ (data not shown). It is interesting to note that despite having the preferred order of heterocycles, the truncated f-ImPyIm-piperazine compound (**6**) showed negligible binding to 5′-ACGCGT-3′ (site (a), Fig. 4C). This result is not surprising based on the bulky nature of the *N*-methylpiperazine group that inhibits the molecule from binding in the narrow minor groove, especially as a stacked dimer, when it is connected directly to the polyamide.[21] The results suggest that having an appropriate linker to provide flexibility to allow the positively charged C-terminus group to interact with DNA should be considered in the design process.

Alkyldiamine compound (**7**) shows a footprint of 0.1 μM for 5′-ACGCGT-3′ (Fig. 4B). Compound **8**, containing the tripyrrole sequence of heterocycles, produces a footprint of 1 μM with its cognate sequence (data not shown). These results show that the alkyldiamino compounds (**7** and **8**) bind with comparable affinity to their respective cognate sequences compared to distamycin (**1**) and f-PyPyPy (**3**) (0.5 and 1 μM with 5′-A_3T_3-3′[15]). However, as depicted in the gel shown in Figure 4A, f-ImPyIm (**2**) (0.05 μM with 5′-ACGCGT-3′) still binds more strongly to its cognate sequence of 5′-ACGCGT-3′. Nonetheless, the diaminoalkyl C-terminus moiety provides an attractive alternative to the DNA binding properties of a dimethylaminoalkyl group, especially with its improved solubility in water, which could be as a result of the ability of the diaminoalkyl compound to have two cationic centers in solution, rather than just one as with the dimethyl compounds. These binding data highlight that alongside heterocycle order, the length, flexibility, shape, and number of cationic centers can also play an influential role in binding to duplex DNA.

2.3. Circular dichroism

CD spectroscopy was employed (Fig. 5) to evaluate the binding motif of the polyamides with duplex DNA. DNA alone produces a CD signal (due to its inherent chirality), however the non-chiral polyamide compounds produce no response. When polyamide is added to a solution of DNA containing the appropriate recognition site, a ligand-induced band appears at approximately 300–330 nm. The observation of an isodichroic point is indicative of only one binding mode consistent with minor groove binding.[22] All compounds tested (with the exception of the truncated piperazine compound, **6**) showed binding to their cognate sequences, however differences in the magnitude and saturation ratios of the induced bands were observed.

The best compound in the series was the f-ImPyIm alkyldiamino compound **7** as this compound produced a strong induced band, and saturation was reached at a ratio of ~7:1 (polyamide/cognate DNA). As reported earlier,[15] f-ImPyIm (**2**) also produced a strong induced CD band at ~300 nm; however, the saturation point was reached at a 2:1 ratio (polyamide/cognate DNA). These results are consistent with the footprinting data, compound **7** demonstrated a weaker signal than that induced by f-ImPyIm (**2**) (Fig. 4B vs D). It is worth noting that the amine (*N*-benzyloxycarbonylaminopropyl-*N*-aminopropyl-*N*-methylamine) alone (**23**, structure given in Table 1)[23,24] did not produce any induced CD band

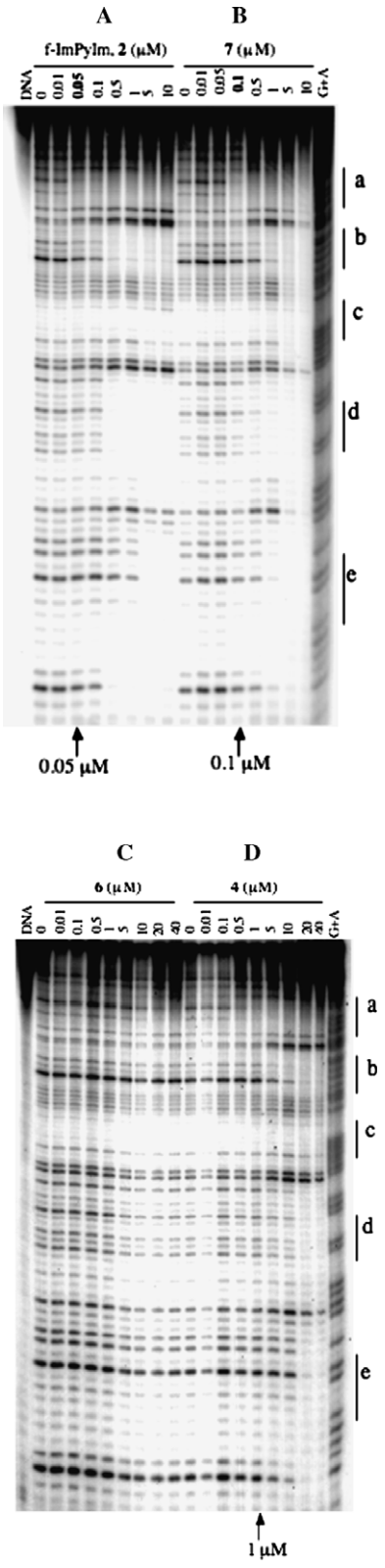


Figure 4. DNase1 footprinting data of compounds f-ImPyIm **2** (A); **7** (B); **6** (C) and **4** (D) using a 5′-[^{32}P]-radiolabeled probe corresponding to the DNA explained in Figure 2, with the sites of interest labeled a–e. DNA is a blank control, G + A is the sequencing lane, and the remaining lanes show increasing concentration of the compounds (in μM).

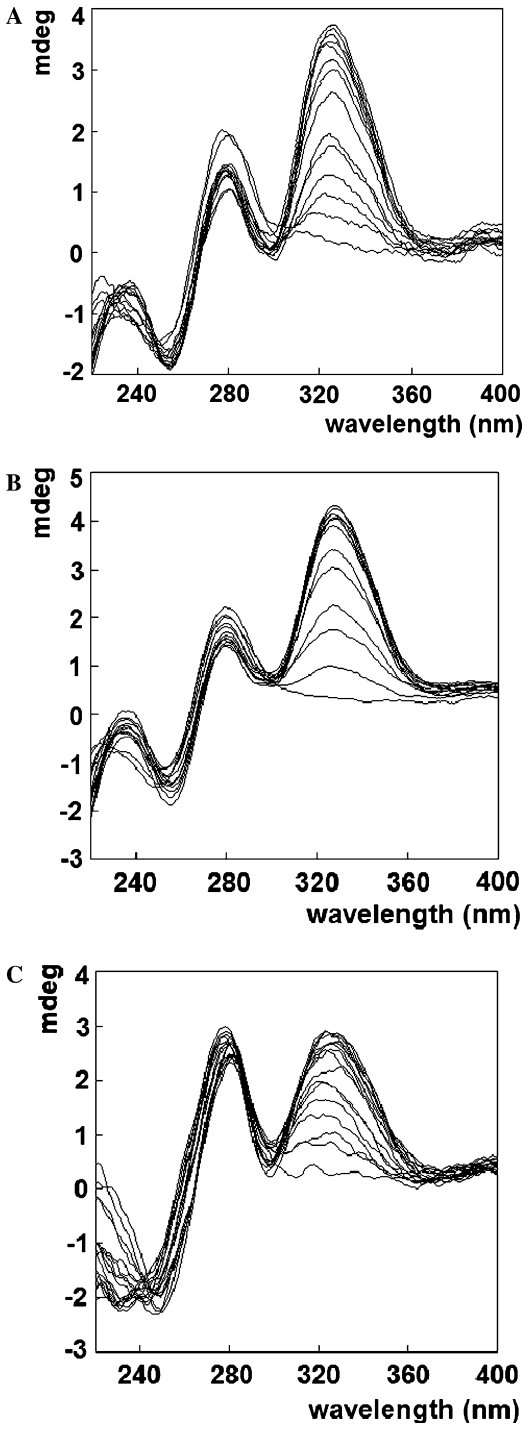
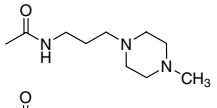
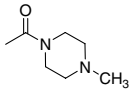
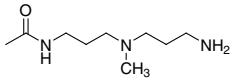
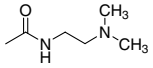
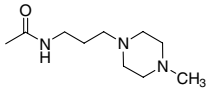
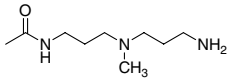
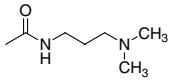
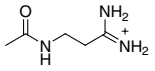
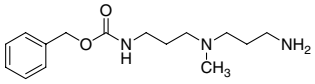


Figure 5. Circular dichroism (CD) spectra of alkylpiperazine **4** (A) and alkyldiamino **7** (B) binding to 5′-ACGCGT-3′, and alkylpiperazine **5** (C) binding to 5′-AAATTT-3′. Each compound was added to a 9 μM DNA solution in aliquots of 1 M equivalents until saturation was achieved.

up to 12:1 polyamide/DNA (5′-ACGCGT-3′). The inability of compound **23** to bind to the DNA minor groove indicates that the polyamide moiety within the molecule drives the binding. The compounds recognize the sequence they were designed to and the amine moiety is then able to 'fine-tune' the interaction; clearly

Table 1. Thermal melt data for C-terminus modified compounds

Compound		5′-ACGCGT-3′ ΔT_m°C	5′-A_3T_3-3′ ΔT_m°C
f-ImPyIm-R R=			
	**4**	2.8[a]	0[a]
	**6**	0	0
	**7**	>17[a]	5[a]
	**2**	7.8[b]	1.1[b]
f-PyPyPy-R′ R′=			
	**5**	n.d.	8.3[a]
	**8**	2[b]	>22[b]
	**3**	0.1[b]	8.8[b]
	**1**	0[b]	14[b]
	**23**	0[c]	0[c]

Ratio of polyamide to DNA [a](5:1), [b](3:1), [c](10:1). All T_m and CD data were obtained using $PO_4$0 buffer.

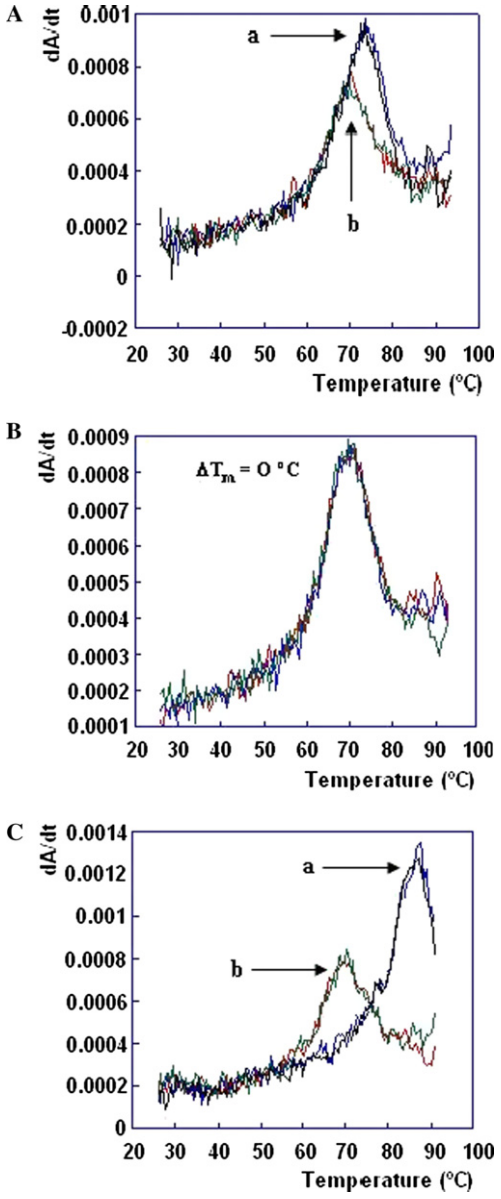


Figure 6. Thermal melt derivative plots of f-ImPyIm-containing compounds **4** (A), **6** (B) and **7** (C) with the cognate 5′-ACGCGT-3′, where a and b denote the T_m (°C) of the complex and free DNA, respectively, performed in duplicate.

when the amine moiety is too bulky, for example **6**, the compound is unable to bind.

2.4. Thermal melting (ΔT_m)

Thermal denaturation experiments were performed to assess the ability of the polyamides to stabilize the melting temperature of hairpin DNA at elevated temperature. The results of these experiments are shown in Table 1.

Compounds containing the Im-Py-Im order of heterocycles were considered, all compounds (with the exception of **6**) stabilized 5′-ACGCGT-3′ more strongly than they did their non-cognate 5′-A_3T_3-3′ DNA. The most impressive stabilization was found for the alkyldiamino compound **7** (ΔT_m = >17 °C, Fig. 6C). Compound **6** did not stabilize either DNA, in agreement with the footprinting data, again confirming the inability of the piperazine moiety to form stacked dimers in the DNA minor groove (Fig. 6B).

A similar trend was observed for the tripyrrole series. The alkylpiperazine compound (**5**) stabilized its cognate 5′-A_3T_3-3′ with a ΔT_m of 8.3 °C, however, the alkyldiamino compound (**8**) stabilized the same sequence with a ΔT_m of >22 °C even at a lower ratio of polyamide to DNA (Table 1). Compound f-PyPyPy (**3**) and distamycin, **1**, (ΔT_m = 8.8 and 14 °C, respectively [data not shown]) do not stabilize the same sequence of DNA as strongly as the alkyldiamine compounds.

The amine alone (**23**) produced no measurable stabilization at ratios up to 10:1 polyamide/DNA (Table 1). Again this indicates that the specificity of binding is driven by the heterocycle order, and that C-terminus functionality affects the binding affinity, unless the

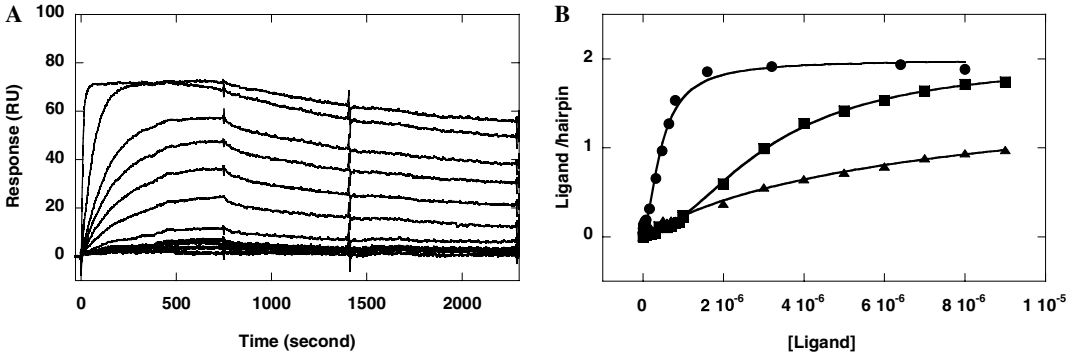


Figure 7. (A) Typical sensorgrams for the interaction of compound **7** with the 5′-ACGCGT-3′ hairpin. The concentrations of these sensorgrams are from 0.3 nM to 8 μM. The experiment was conducted in 200 mM Na^+ with a flow rate of 20 μL/min and a 15-min dissociation period. (B) Binding isotherms of compound **7** with three different DNA hairpins. Positive cooperative binding was observed with 5′-ACGCGT-3′ (circles) and 5′-ACCGGT-3′ (squares). Negative cooperative binding behavior was observed with 5′-AAATTT-3′ (triangles).

group is too bulky or rigid, in which case binding is prevented entirely, as with compound **6**.

2.5. Surface plasmon resonance (SPR)

The sensorgrams for interaction of compound **7** with its cognate sequence 5′-ACGCGT-3 exhibit a slow observed kinetic pattern (Fig. 7A) as previously seen with f-ImPyIm (**2**), which gave a binding affinity of $1.9 \times 10^8\ M^{-1}$.[15] In contrast, much faster observed kinetics was observed for 5′-AAATTT-3′ and 5′-ACGCGT-3′ sequences. The binding isotherms of this ligand with three DNA hairpins and their fits are shown in Figure 7B. The alkyldiamino-f-ImPyIm compound **7** retains its positive cooperative binding behavior with the C/G sites (5′-ACGCGT-3′ and 5′-ACCGGT-3′). The binding to the 5′-AAATTT-3′ site has a negative cooperative feature and is remarkably weak. For comparison purposes, the average of two binding constants $(K_1 * K_2)^{1/2}$ is reported for all three sequences. The binding affinities of compound **7** to the 5′-ACGCGT-3′, 5′-ACCGGT-3′, and 5′-AAATTT-3′ sequences are 2.2×10^6 (M^{-1}), 3.3×10^5 (M^{-1}), and 1×10^5 (M^{-1}), respectively. The reproducibility error in K is 20% for the G/C sequences and higher for weak binding to the A/T sequences. It is clear, however, that the binding constant for the 5′-ACGCGT-3′ site is much greater than that for the 5′-AAATTT-3′. This result is in good agreement with the footprinting data given in Figure 4 and the thermal melting studies in Table 1.

2.6. Conclusions

Five compounds were synthesized and the DNA binding ability of these compounds assessed by biochemical and biophysical techniques. Overall the alkylpiperazine compounds (**4** and **5**) did show binding to their cognate sequence but this was weaker than that shown by the alkyldiamino compounds (**7** and **8**). All techniques presented show the same relationship, with compound **7** producing the strongest binding with a binding constant of $2.2 \times 10^6\ M^{-1}$ determined by SPR with its cognate sequence. This paper has illustrated that steric bulk, flexibility, and the number of possible protonation sites of the C-terminus group are important with respect to binding, however the strong binding of compound **7** could also be due to anchorage of the extended diamine in the groove. This investigation indicates that work still needs to be done to surpass the binding affinities exhibited by distamycin (**1**) and f-ImPyIm (**2**).

3. Experimental

3.1. Synthesis

Solvents and organic reagents were purchased from Aldrich or Fisher and in most cases were used without further purification. DCM (P_2O_5) and DMF (BaO) were distilled prior to use. Melting points (mp) were performed using a Mel-temp instrument and are uncorrected. Infrared (IR) spectra were recorded using a Perkin-Elmer Paragon 500 FT-IR instrument as films on NaCl discs, unless otherwise stated. ^1H NMR spectra were obtained using a Varian Unity Inova 500 instrument unless otherwise stated. Chemical shifts (δ) are reported at 20 °C in parts per million (ppm) downfield from internal tetramethylsilane (Me_4Si). High-resolution mass spectra (HRMS) and low-resolution mass spectra (LRMS) were provided by the Mass Spectrometry Laboratory, University of South Carolina, Columbia. Reaction progress was assessed by thin-layer chromatography (TLC) using Merck silica gel (60 F_{254}) on aluminum plates unless otherwise stated. Visualization was achieved with UV light at 254 nm and/or 366 nm, I_2 vapor staining, and ninhydrin spray.

3.2. General method A: reduction of a nitro-group to an amino-group

The required nitro-compound (1 mmol equiv) was dissolved/suspended in cold MeOH and 5% Pd/C was added (50% by mass). The reaction mixture was sonicated briefly to ensure the reaction mixture was homogeneous. The reaction flask was degassed, purged with hydrogen (3×) with stirring, and then stirred for ~18 h at room

temperature. The catalyst was removed by filtration over Celite and washed thoroughly with MeOH. The solvent was removed, co-evaporated with DCM (3× 2 mL), and used directly.

3.3. General method B: carbodiimide coupling reaction

The required acid (1.2 mmol equiv), EDCI (3 mmol equiv), and DMAP (0.1 mmol equiv) were dried over P_2O_5, under vacuum, for ~2 h. The reaction flask was then flushed with Ar. The required amine (1 mmol equiv) was dissolved in dry DMF (1.5 mL per 100 mg) and then added to the dry reagents. Sonication was used to ensure a homogeneous mixture and the reaction mixture was stirred under Ar, for 4 days. The DMF was removed by using Kügelrohr apparatus (5 mmHg and 100 °C). The residue was purified by silica gel column chromatography using a gradient system of $CHCl_3$/MeOH.

3.4. Synthesis of 1-methyl-*N*-(3-[4-methylpiperazin-1-yl]propyl)-4-(4-[4-formamido-1-methylimidazole-2-carboxamido]-1-methylpyrrole-2-carboxamido)-1-methylimidazole-2-carboxamide (4)

General methods A and B were employed using 1-methyl-*N*-(3-[4-methylpiperazin-1-yl]propyl)-4-(4-1-methylpyrrole-2-carboxamido)-1-methylimidazole-2-carboxamide, **9**, and 4-formamido-1-methylimidazole-2-carboxamide carboxylic acid, **13**, to yield **4** as a brown oil (40 mg, 66%): ^{1}H NMR ($CDCl_3$) 1.80 (quin, 2H, J = 9 Hz); 2.17 (m, 4H); 2.38 (s, 3H); 2.55 (t, 2H, J = 6 Hz); 2.62 (m, 4H); 3.46 (quin, 2H, J = 6 Hz); 3.96 (s, 3H); 4.02 (s, 3H); 4.06 (s, 3H); 6.80 (d, 1H, J = 1.5 Hz); 7.30 (d, 1H, J = 3 Hz); 7.41 (s, 1H); 7.44 (s, 1H); 7.89 (s, 1H); 8.35 (s, 2H); 8.95 (s, 1H); IR (neat) ν 3393, 3082, 2958, 2916, 2854, 1664, 1643, 1550, 1534, 1467, 1446, 1405, 1259, 1099, 1016, 803 cm^{-1}; MS (ES+) m/z (rel intensity) 554 ([M+H], 100%); HRMS $[M+H]^+$ calcd for $C_{25}H_{35}N_{11}O_4$ m/z 554.2951; found m/z 554.2956.

3.5. Synthesis of 1-methyl-*N*-(3-[4-methylpiperazin-1-yl]propyl)-4-(4-[4-formamido-1-methylpyrrole-2-carboxamido]-1-methylpyrrole-2-carboxamido)-1-methylpyrrole-2-carboxamide (5)

General methods A and B were employed using 1-methyl-*N*-(3-[4-methylpiperazin-1-yl]propyl)-4-nitro-(4-1-methylpyrrole-2-carboxamido)-1-methylpyrrole-2-carboxamide, **10**, and 4-formamido-1-methylpyrrole-2-carboxylic acid, **14**, followed by typical purification to yield **5** as a brown oil (53 mg, 56%): ^{1}H NMR ($CDCl_3$) 1.69 (quin, 2H, J = 5 Hz); 2.56 (m, 4H); 2.51 (m, 4H); 3.14 (q, 2H, J = 6 Hz); 3.19 (t, 2H, J = 5 Hz); 3.39 (s, 3H); 3.88 (s, 3H); 3.91 (s, 3H); 6.63 (quin, 1H, J = 3.3 Hz); 6.65 (d, 1H, J = 5 Hz); 6.73 (s, 1H, J = 5 Hz); 6.77 (d, J = 1 Hz); 6.79 (d, 1H, J = 5 Hz); 7.10 (s, 1H); 7.12 (d, 1H, J = 5 Hz); 8.82 (d, 1H, J = 5 Hz); 8.47 (s, 1H); 8.28 (s, 1H); IR (neat) ν 3351.9, 2947.2, 2864.2, 2833.0, 2355.7, 1638.2, 1565.6, 1435.8, 1259.4, 1160.8, 1098.6, 1057.1, 673.1 cm^{-1}; MS (ES+) m/z (rel intensity) 552 ([M+H], 10%); HRMS $[M+H]^+$ calcd for $C_{27}H_{37}N_9O_4$ m/z 552.3047; found m/z 552.3041.

3.6. Synthesis of 1-methyl-*N*-(3-[4-methylpiperazin-1-yl])-4-(4-[4-formamido-1-methylpyrrole-2-carboxamido]-1-methylpyrrole-2-carboxamido)-1-methylpyrrole-2-carboxamide (6)

General methods A and B were employed using 1-methyl-*N*-(3-(4-methylpiperazin-1-yl))-4-(1-methyl-pyrrole-2-carboxamido)-1-methylimidazole-2-carboxamide, **15**, and 4-formamido-1-methylimidazole-2-carboxylic acid, **13**, to yield **6** as a light brown oil (20 mg, 37%): ^{1}H NMR ($CDCl_3$) 2.79 (s, 3H); 2.5 (m, 4H); 3.78 (m, 4H); 3.87 (s, 3H); 3.98 (s, 3H); 4.19 (s, 3H); 6.77 (d, 1H, J= 2.5 Hz); 7.2 (d, 1H, J = 1 Hz); 7.40 (s, 1H); 7.46 (s, 1H); 7.63 (s, 1H); 8.0 (s, 1H); 8.39 (s, 1H); 8.78 (s, 1H); IR (neat) ν 3403.8, 3341.5, 2957.5, 2345.3, 1669.3, 1560.3, 1560.4, 1461.8, 1259.4, 1119.3, 1020.8, 792.5, 662.7 cm^{-1}; MS (ES+) m/z (rel intensity) 497 ([M+H] 10%); HRMS $[M+H]^+$ calcd for $C_{22}H_{28}N_{10}O_4$ m/z, 497.2373; found m/z 497.2368.

3.7. Synthesis of *N*-(3-[*N*,*N*-{3-aminopropan-1-yl}methyl]propan-1-yl)-4-(4-[4-formamido-1-methylimidazole-2-carboxamido]-1-methyl-pyrrole-2-carboxamido)-1-methylimidazole-2-carboxamide (7)

General methods A and B were used with *N*-(3-[*N*,*N*-{3-aminopropan-1-yl}methyl]propan-1-yl)-4-nitro-1-methyl-1*H*-imidazole-2-carboxamide **17** and 4-(4-formamido-1-methylimidazole-2-carboxamido)-1-methylpyrrole carboxylic acid (**21**) except DCC (3 mmol equiv) and PtO_2 (25% by weight) for ~1.5 h were used, followed by typical purification to yield the benzyl-protected intermediate as a tan oil (147 mg, 60%); ^{1}H NMR ($CDCl_3$) 1.74 (m, 4H); 2.2 (br t, 2H, J = 2.5 Hz); 2.25 (br t, 2H, J = 2.5 Hz), 2.46 (br s, 3H); 3.28 (q, 2H, J= 3 Hz); 3.55 (q, 2H, J = 3 Hz); 3.98 (s, 3H); 4.01 (s, 3H); 4.08 (s, 3H); 5.08 (s, 2H); 6.65 (s, 1H); 6.90 (d, 1H, J= 2 Hz); 7.29 (s, 1H); 7.31 (s, 1H); 7.36 (d, 1H, J = 2 Hz); 7.61 (s, 1H); 7.84 (s, 1H); 8.35 (s, 1H), 10.41 (s, 1H); MS (ES+) m/z (rel intensity) 676 ([M+H] 100%).

The benzyl-protecting group was removed using general method C except 10% Pd/C was used. The residue was purified by preparative TLC (70:30 v/v, $CHCl_3$/MeOH) yielding compound **7** as a pale brown solid (67 mg, 57%), mp 280–285 °C: ^{1}H NMR (CD_3OD) 2.04 (quin, 2H, J = 5.0 Hz); 2.09 (quin, 2H, J = 5.0 Hz) 2.8 (br t, 3H); 3.07 (t, 4H, J = 7.5 Hz); 3.47 (t, 2H; J = 6.5 Hz); 3.95 (s, 3H); 4.58 (s, 3H); 7.31 (d, 1H, J = 2.0 Hz); 7.44 (d, 1H, J = 1.5 Hz); 7.43 (s, 1H); 7.44 (s, 1H); 8.25 (s, 1H); IR (neat) ν 3292, 2927, 1655, 1533, 1467, 1441, 1404, 1253, 1207, 1123, 1061, 895, 758, 626 cm^{-1}; MS (ES+) m/z (rel intensity) 542 ([M+H], 100%); HRMS $[M+H]^+$ calcd for $C_{24}H_{35}N_{11}O_4$ m/z, obsd 542.2951; found m/z 542.2944.

3.8. Synthesis of *N*-(3-[*N*,*N*-{3-aminopropan-1-yl}methyl]propan-1-yl)-4-(4-[4-formamido-1-methylpyrrole-2-carboxamido]-1-methylpyrrole-2-carboxamido)-1-methylpyrrole-2-carboxamide (8)

General methods A and B were employed with *N*-(3-[*N*,*N*-{3-aminopropan-1-yl}methyl]propan-1-yl)-4-nitro-1-methyl-pyrrole-2-carboxamide (**18**) and 4-(4-formamido-1-methylpyrrole-2-carboxamido)-1-methylpyrrole-2-carboxylic **22**, except DCC (3 mmol equiv) and PtO_2 (25% by weight) for ~1.5 h were used, followed by typical purification to yield the intermediate as a tan oil (126 mg, 53%): ^{1}H NMR ($CDCl_3$) 1.74 (m, 4H); 2.25 (s, 3H); 2.46 (br m, 4H); 3.28 (br t); 3.55 (br t); 3.98 (s, 3H); 4.00 (s, 3H); 4.08 (s, 3H); 5.32 (s, 2H); 6.89 (d, 1H, J = 1.5 Hz) 7.29 (s, 1H); 7.31 (s, 1H); 7.32 (m, 5H); 7.35 (s, 1H); 7.39 (s, 1H); 7.45 (s, 1H); 7.61 (br s, 1H); 7.84 (br s, 1H); 8.35 (s, 1H); IR (neat) ν 3283, 3138, 2940, 1639, 1530, 1434, 1259 cm^{-1}; MS (ES+) *m/z* (rel intensity) 674 ([M+H] 60%); HRMS $[M+H]^+$ calcd for $C_{34}H_{43}N_9O_6$ *m/z* 674.3417; found 674.3414.

The benzyl protecting group in the intermediate was removed using general method A except 10% Pd/C was used to yield **8** as a brown solid (80.4 mg, 84%), mp 270–275 °C: ^{1}H NMR (CD_3OD) 1.38 (m, 4H); 2.03 (t, 2H, J = 7 Hz); 2.13 (t, 2H, J = 7 Hz); 2.89 (br s, 3H); 3.07 (t, 2H, J = 7 Hz); 3.44 (t, 2H, J = 7 Hz); 3.90 (s, 3H) 3.91 (s, 3H); 3.92 (s, 3H); 6.90 (d, 1H, J= 1.5 Hz); 6.92 (d, 1H, J= 1.5 Hz); 6.97 (d, 1H, J = 1.5 Hz); 7.17 (d, 1H, J= 2.0 Hz); 7.18 (d, 1H, J = 1.5 Hz); 7.19 (d, 1H, J = 2 Hz); 8.13 (s, 1H); IR (neat) ν 3261, 2947, 1635, 1528, 1433, 1402, 1262 cm^{-1}; MS (ES+) *m/z* (rel intensity) 540 ([M+H] 100%); HRMS $[M+H]^+$ calcd for $C_{26}H_{37}N_9O_4$ *m/z* 540.3047; found 540.3041.

3.9. Biological evaluation

3.9.1. DNase1 footprinting. *Taq* polymerase and DNase I were purchased from Promega, T4 polynucleotide kinase from Invitrogen. The PCR forward primers: 5′-CTCCAGAAAGCCGGCACTCAG-3′, PCR reverse primers: 5′-GTCGGTTAGGAGAGCTCCACTTG-3′, the templates IM18: 5′-ATGCTCCAGAAAGCCGGCACTCAGTCTACAAACGCGTCATCTTGATCATGCATGTTCACAGAAATTTCTCTAGATCTAAGCTAAACTCTAGTACTAGTCTTCAAGCAAGTGGAGCTCTCCTAACCGACTTT-3′ and IM20: 5′-AAAGTCGGTTAGGAGAGCTCCACTTGCTTGAAGACTAGTACTAGAGTTTAGCTTAGATCTAGAGAAATTTCTGTGAACATGCATGATCAAGATGACGCGTTTGTAGACTGAGTGCCGGCTTTCTGGAGCAT-3′ were synthesized by MWG. All other chemicals are analytical grade reagents.

3.9.2. Preparation of the DNA substrate, radiolabeling, and purification. A 132 bp fragment was designed and amplified by PCR. Reverse primer (4 ng) was radioactively labeled by reaction with γ-[^{32}P]-ATP (3 μL) and T4 polynucleotide kinase (1 μL) in 10 μL of forward reaction buffer (70 mM Tris–HCl, pH 7.6, 10 mM $MgCl_2$, 100 mM KCl, and 1 mM 2-mercaptoethanol) for 1 h at 37 °C. PCR was performed in 50 μL thermophilic DNA poly buffer containing dNTPs (125 μM), $MgCl_2$ (1 mM), *Taq* polymerase (5 μ), [^{32}P]-labeled forward primer (4 ng), reverse primer (4 ng), and IM18/IM20 templates (20 ng). The PCR products were desalted using a P6 biogel spin column (BIO-RAD) and purified by 3% agarose gel electrophoresis. DNA was isolated using MERMAID kit (Q-Biogene) according to the manufacturer's instructions.

3.9.3. DNase I footprinting reaction. DNase I reactions were conducted in a total volume of 8 μL. The labeled DNA fragment (2 μL, 200 counts/s) was incubated for 30 min in 4 μL TN binding buffer (10 mM Tris, pH 7.0, 10 mM NaCl) containing the desired drug concentration. DNA cleavage was performed by addition of 2 μL DNase I solution (20 mM NaCl, 2 mM $MgCl_2$, and 2 mM $MnCl_2$, pH 8.0) adjusted to yield a final enzyme concentration of around 0.1 U mL^{-1} in the reaction mixture. The extent of digestion was limited to less than 30% of the starting material so as to minimize the incidence of multiple cuts in any strand. After 8 min, the digestion was stopped by transferring the tube to dry ice and the samples were lyophilized. DNA was resuspended in 4 μL of 80% formamide containing tracking dye (bromophenol and cyanol blue) and denatured by heating the samples to 90 °C and rapid cooling on ice prior loading in a conventional 6% denaturing polyacrylamide gel containing 7.5 M urea. Electrophoresis was performed for 120 min at 2000 V (around 70 W and 55 °C) in TBE buffer (89 mM Tris base, 89 mM boric acid, and 2.5 mM Na_2EDTA, pH 8.3). Finally the gel was transferred onto Whatman 3 mm paper and dried under vacuum at 80 °C for 2 h. The gel was exposed overnight to X-ray film at −80 °C.

3.9.4. Biophysical evaluation. All buffer reagents were obtained from Sigma–Aldrich or Fisher and used without further purification. The synthetic oligonucleotides were obtained from Operon (Huntsville, AL): **5′-ACGCGT-3′**: 5′-GA ACGCGT CG CTCT CGACGCGTTC-3′ and **5′-A$_3$T$_3$-3′**: 5′-CG AAATTT CC CTCT GG AAATTT CG-3′ and were used in the thermal denaturation and circular dichroism experiments without further purification. DNA used in the footprinting and SPR experiments is described below. T_m and CD data were obtained using a Varian-Cary 100Bio UV–vis spectrophotometer, equipped with a Peltier temperature controller and a precision cuvette-mounted temperature probe and a Jasco J-710 spectrophotometer, respectively. Phosphate buffer: $PO_4$0 (10 mM sodium phosphate, 1 mM EDTA, pH 6.4) was used for CD and T_m experiments. All data were analyzed using KaleidaGraph (Snergy Software, Reading PA), unless otherwise stated. Ligand stock solutions were prepared in double distilled water at a concentration of 0.5 mM, unless otherwise specified.

3.9.5. Circular dichroism (CD). CD studies were performed using published procedures,[8,9] and were conducted at ambient temperature in a 1-mm pathlength quartz cell using $PO_4$0 buffer. Buffer and stock DNA were added to the cuvette to give a final DNA concentration of 9 μM. Ligand was titrated in aliquots of 0.5 (compound **8**) or 1 molar equivalents (all other compounds) until saturation was achieved. The CD response at the λ_{max} of the induced peak was plotted against the mole ratio of ligand:DNA.

3.9.6. Thermal melts (T_m). Thermal denaturation studies were performed using published procedures.[8,9] Experiments for compounds **4**, **5**, and **6** were performed at a concentration of 5 μM ligand and 1 μM DNA, compounds **7** and **8** were performed at concentrations of 3 μM ligand to 1 μM ligand, and compound **23** was performed at a concentration of 10 μM ligand to 1 μM DNA. All experiments were run in $PO_4$0 buffer. Oligonucleotide samples were reannealed prior to denaturation studies by heating at 70 °C for 1 min then cooling to RT. Heating runs were typically performed between 25 and 95 °C, with a heating rate of 0.5 °C min^{-1}, while continuously monitoring the absorbance at 260 nm (digitally sampled at 200 ms intervals). All melts were performed in 10-mm pathlength quartz cells. T_m values were determined as the maximum of the first derivative.

3.9.7. Surface plasmon resonance (SPR). The biosensor experiments were conducted in degassed phosphate buffer (200 mM Na^+, 10 mM sodium phosphate, 1 mM EDTA, and 0.00005 v/v of 10% surfactant P20—BIACORE, pH 6.25) at 25 °C. The 5′-biotin-labeled DNA hairpins were purchased from Midland Certified Reagent Company (Midland, TX) with HPLC purification. The DNA sequences are 5′-biotin-GAACGCGTCCTCTGACG CGTTC-3′, 5′-biotin-GAACCGGTCCTCTGACCGG TTC-3′, and 5′-biotin-CGAAATTTCCTCTGAAATTT CG-3′ (denoted as ACGCGT, ACCGGT, and AAATTT, respectively). The experiments were conducted with a BIACORE 2000 or BIACORE 3000 instrument (Biacore AB). The DNA hairpins were immobilized on a streptavidin-derivatized gold chip (SA chip from BIAcore) by manual injection of 25 nM hairpin DNA solution with flow rate of 1 μL/min until the response units reach about 400–500 RUs. Flow cell 1 was left blank while flow cells 2, 3, and 4 were immobilized with three different DNA hairpins. Typically, a series of different concentrations of ligand were injected (injected volume = 250 μL) onto the chip with a flow rate of 10 or 20 μL/min followed by a dissociation period of 15 or 25 min. The chip surface was regenerated with a 1/2 min injection of a regenerating solution (0.4 M NaCl, 20 mM NaOH). The steady-state responses are converted to a stoichiometry-normalized binding isotherm and fitted with a two-site model as previously described.[8,15]

$$r = (K_1 \times C_{free} + 2 \times K_1 \times K_2 \times C_{free}^2)/(1 + K_1 \times C_{free} + K_1 \times K_2 \times C_{free}^2)$$

C_{free} : ligand concentration in the flow solution.

Acknowledgments

Support from The Camille and Henry Dreyfus Foundation, the National Science Foundation, and Cancer Research, UK are gratefully acknowledged.

References and notes

1. (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723; (b) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393.
2. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
3. Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376.
4. Lown, J. W.; Krowicki, K.; Bhat, U. G. *Biochemistry* **1986**, *25*, 7408.
5. Lee, M.; Preti, C. S.; Vinson, E.; Wyatt, M. D.; Hartley, J. A. *J. Med. Chem.* **1994**, *37*, 4073.
6. Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983.
7. Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104.
8. Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2153.
9. Buchmueller, K. L.; Staples, A. M.; Uthe, P. B.; Howard, C. M.; Pacheco, K. A. O.; Cox, K. K.; Henry, J. A.; Bailey, S. L.; Horick, S. M.; Nguyen, B.; Wilson, W. D.; Lee, M. *Nucleic Acids Res.* **2005**, *33*, 912.
10. Kopka, M. L.; Goodsell, D. S.; Hand, G. W.; Chiu, T. K.; Lown, J. W.; Dickerson, R. E. *Structure* **1997**, *5*, 1033.
11. Yang, X.-L.; Kaenzig, C.; Lee, M.; Wang, A. H.-J. *Eur. J. Biochem.* **1999**, *263*, 616.
12. Yang, X.-L.; Hubbard, R. B.; Lee, M.; Tao, Z.-F.; Sugiyama, H.; Wang, A. H.-J. *Nucleic Acids Res.* **1999**, *27*, 4183.
13. Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284, and references contained within.
14. Lacy, E. R.; Nguyen, B.; Le, M.; Cox, K. K.; O'Hare, C.; Hartley, J. A.; Lee, M.; Wilson, W. D. *Nucleic Acids Res.* **2004**, *32*, 2000.
15. Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Le, N. M.; Cox, K. K.; Nguyen, B.; Pacheco, K. A. O.; Wilson, W. D.; Lee, M. *J. Am. Chem. Soc.* **2005**, *127*, 742.
16. Brooks, N.; Hartley, J. A.; Simpson, J. E., Jr.; Wright, S. R.; Woo, S.; Centioni, S.; Fontaine, M. D.; McIntyre, T. E.; Lee, M. *Bioorg. Med. Chem* **1997**, *5*, 1497.
17. Pjura, P. E.; Grzeskowiak, K.; Dickerson, R. E. *J. Mol. Biol.* **1987**, *197*, 257.
18. Carrondo, M. A. A. F de C. T.; Coll, M.; Ayamami, J.; Wang, A. H.-J.; Marel van der, G. A.; Boom van, J. H.; Rich, A. *Biochemistry* **1989**, *28*, 7849.
19. Searle, M. S.; Embrey, K. J. *Nucleic Acids Res.* **1990**, *18*, 3753.
20. *CRC Handbook of Chemistry and Physics*, 70th ed.; CRC Press: Weast, 1989–1990; p D162.
21. Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *Biophys. J.* **2001**, *81*, 1588.
22. (a) Lyng, R.; Rodger, A.; Norden, B. *Biopolymers* **1992**, *32*, 1201; (b) Lyng, R.; Rodger, A.; Norden, B. *Biopolymers* **1991**, *31*, 1709.
23. Allainmat, M.; Haridon, P. L.; Toupet, L.; Pluquelled, D. *Synthesis* **1990**, 27.
24. Makino et al. U.S. Patent 6,368,807 B2, 2002.