Available online at www.sciencedirect.com
ScienceDirect
Bioorganic & Medicinal Chemistry
Bioorganic & Medicinal Chemistry 16 (2008) 5266–5276

Modifying the N-terminus of polyamides: PyImPyIm has improved sequence specificity over f-ImPyIm

Toni Brown,[a,b] Hilary Mackay,[a,b] Mark Turlington,[a] Arden Sutterfield,[a] Traci Smith,[b] Alan Sielaff,[b] Laura Westrate,[b] Chrystal Bruce,[c] Jerome Kluza,[d] Caroline O'Hare,[d] Binh Nguyen,[e] W. David Wilson,[e] John A. Hartley[d] and Moses Lee[a,b,*]

[a]*Department of Chemistry, Furman University, Greenville, SC 29613, USA*
[b]*Division of Natural Sciences and Department of Chemistry, Hope College, Holland, MI 49423, USA*
[c]*Department of Chemistry, Erskine College, Due West, SC 29639, 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 6 January 2008; accepted 3 March 2008
Available online 7 March 2008

Abstract—Seven N-terminus modified derivatives of a previously published minor-groove binding polyamide (f-ImPyIm, **1**) were synthesized and the biochemical and biophysical chemistry evaluated. These compounds were synthesized with the aim of attaining a higher level of sequence selectivity over f-ImPyIm (**1**), a previously published strong minor-groove binder. Two compounds possessing a furan or a benzofuran moiety at the N-terminus showed a footprint of 0.5 μM at the cognate ACGCGT site (determined by DNase I footprinting); however, the specificity of these compounds was not improved. In contrast, PyImPyIm (**4**) produced a footprint of 0.5 μM but showed a superior specificity using the same technique. When evaluated by thermal melting experiments and circular dichroism using ACGCGT and the non-cognate AAATTT sequence, all compounds were shown to bind in the minor-groove of DNA and stabilize the cognate sequence much better than the non-cognate (except for the non-amido-compound that did not bind either sequence, as expected). PyImPyIm (**4**) was interesting as the ΔT_m for this compound was only 4 °C but the footprint was very selective. No binding was observed for this compound with a third DNA (non-cognate, ACCGGT). ITC studies on compound **4** showed exothermic binding with ACGCGT and no heat change was observed for titrating the compound to the other two DNA sequences. The heat capacity (ΔC_p) of the PIPI/ACGCGT complex calculated from the hydrophobic interactions and SASA calculations was comparable to the experimental value obtained from ITC (-146 cal mol^{-1} K^{-1}). SPR results provided confirmation of the sequence specificity of PyImPyIm (**4**), with a K_{eq} value determined to be 7.1×10^6 M^{-1} for the cognate sequence and no observable binding to AAATTT and ACCGGT. Molecular dynamic simulations affirmed that PyImPyIm (**4**) binds as a dimer in an overlapped conformation, and it fits snugly in the minor-groove of the ACGCGT oligonucleotide. PyImPyIm (**4**) is an especially interesting molecule, because although the binding affinity is slightly reduced, the specificity with respect to f-ImPyIm (**1**) is significantly improved.

1. Introduction

Formamido(f)-Imidazole(Im)-Pyrrole(Py)-Im (**1**) is a polyamide (PA) molecule that binds to GC rich sequences in the minor-groove of DNA in a 2:1 fashion (PA:DNA) with a binding affinity of 1.9×10^8 M^{-1} (to ACGCGT, determined by SPR); however, the selectivity of this compound could be improved.[1] This molecule is analogous to distamycin (**2**), a naturally occurring pyrrole-containing molecule that interacts selectively with A/T rich sequences, also with high binding affinity,[2] nonetheless compound **1** is superior. The formamido-group at the N-terminus has been shown to influence the mode of binding of these compounds forcing them to bind in a 'staggered', rather than an 'overlapped' motif, and contributes to their high affinity.[3] In an effort to understand the role of the C-terminus on binding affinity, the authors' laboratory recently published the binding affinities of a series of f-ImPyIm analogues. Each of these compounds retained the f-group and the ImPyIm heterocycle order, but possessed various

Keywords: Binding; DNA; Minor-groove; N-terminus; Polyamide.
* Corresponding author. Tel.: +1 616 395 7190; fax: +1 616 395 7923; e-mail: lee@hope.edu

0968-0896/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2008.03.008

alkylamino-moieties at the C-terminus,[4] however, the binding affinity of f-ImPyIm (**1**) was still superior.

As part of an ongoing systematic study within the authors' laboratory to develop minor-groove binders with strong and selective DNA binding affinities (a lower binding affinity is acceptable providing selectivity is enhanced), seven compounds (**3**–**9**) (Fig. 1) with differing groups at the *N*-termini were synthesized. Compound ImPyIm (**3**) was chosen as a negative control as this compound has no formamido-group and should bind in an overlapped fashion (due to the absence of the f-group) and hence, should not target the same DNA sequence as the other compounds. PyImPyIm (**4**) was chosen as another overlapped molecule but one that should target the same sequence of DNA, and the 1-methylpyrrole-2-carboxylic acid starting material is commercially available.

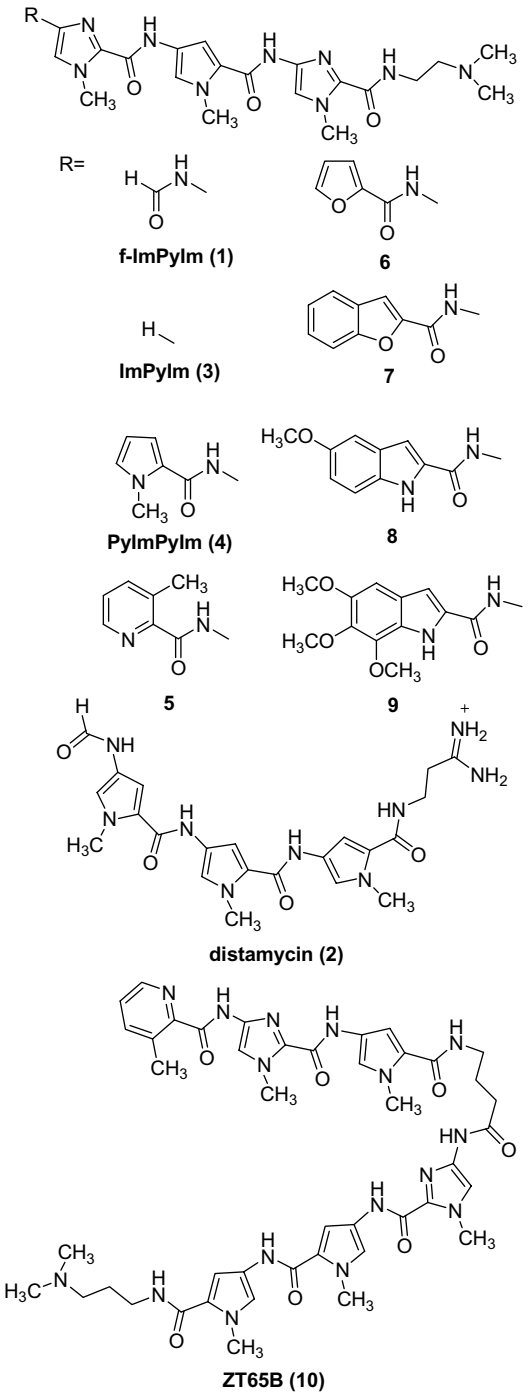


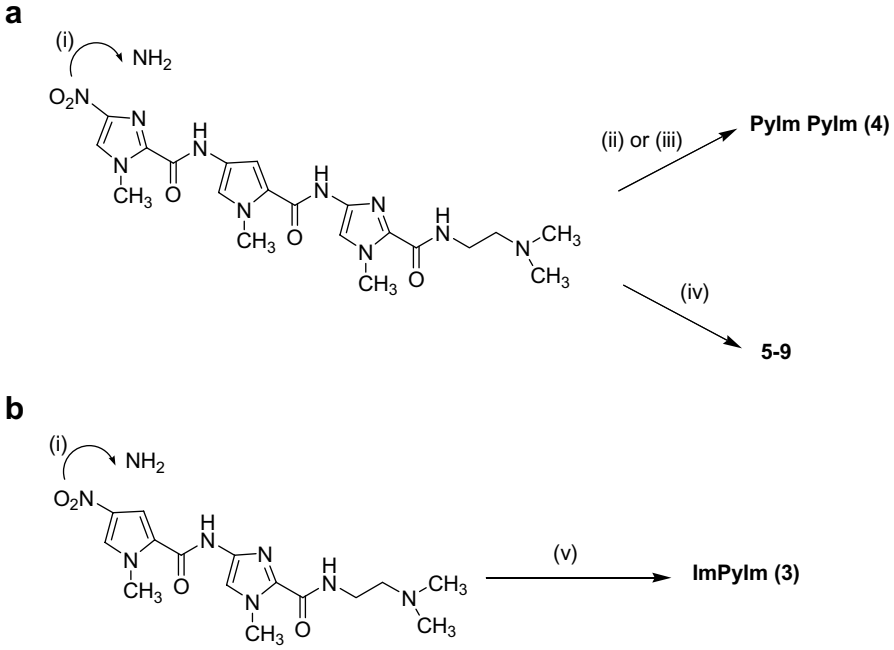
Figure 1. Structures of f-ImPyIm (**1**), distamycin (**2**) the synthesized compounds **3**–**9**, and ZT65B, a hairpin polyamide (**10**).

Compound **5** was synthesized as this compound contains the methylpicolinate group found in ZT-65B (**10**, Fig. 1), a hairpin molecule that also contains this moiety, and has been shown to selectively bind to ICB sites in the promoter region of the human multi-drug resistance gene (mdr1).[5] It also binds selectively to ICB sites 1 and 5 in the topoisomerase IIα promoter region.[5] Compound **6** was designed with a furan at the N-terminus based on previous research that showed that comparable binding affinity was observed when a furan was incorporated into a hairpin molecule.[6,7] Compound **6** was also investigated to determine if any differences in binding affinity arise between compounds that contain an oxygen instead of a nitrogen atom or a methine group. Compound **7** also contains an oxygen moiety but has increased hydrophobicity due to the benzofuran system, and compounds **8** and **9** were developed based on the non-covalent binding subunit of duocarmycins; a potent class of anti-tumor antibiotics that alkylate A/T rich sequences of DNA.[8]

Each of these compounds were evaluated by DNase I footprinting to determine the sequence(s) of DNA that these compounds preferentially recognized. The ligands were also investigated by two biophysical techniques using two different DNA's: the predicted cognate ACGCGT and a non-cognate AAATTT. The initial techniques employed included: thermal denaturation (T_m) and circular dichroism (CD). PyImPyIm (**4**) was also subjected to more in-depth analysis using isothermal titration microcalorimetry (ITC), surface plasmon resonance (SPR), molecular dynamics calculations, and SASA calculations using the two DNA's described above and a similar non-cognate DNA: ACCGGT. In all the compounds, the C-terminus dimethylamino moiety and heterocyclic core (-ImPyIm-) was kept constant.

2. Results and discussion

The synthetic approach is outlined in Scheme 1a and b. The N-terminus of f-ImPyIm (**1**) was varied by the reduction of NO_2-ImPyIm (**11**) using standard palladium catalyzed hydrogenation to the amino-derivative (**12**) as described previously.[1] Compounds **5**–**9** were produced using standard coupling conditions of the relevant carboxylic acid using a parallel synthetic approach with benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) and *N*,*N*-diispropyl-*N*-ethylamine (DiPEA) in dry dichloromethane (DCM) (Scheme 1a). PyImPyIm (**4**) was synthesized by two different routes: one employed 1-methylpyrrole-2-carboxylic acid, *N*,*N*-dicyclohexylcarbodiimide (DCC), and 4-(*N*,*N*-dimethylamino)pyridine (DMAP), also in dry DMF;



Scheme 1. Reagents and conditions: (a) i—5% Pd/C, H_2, cold MeOH; ii—EDCI, DMAP, dry DMF, 0–50 °C, pyrrole-2-carboxylic acid; iii—pyrrole-2-carobonyl chloride, dry TEA, dry DCM, 0 °C to rt; iv—PyBOP, DiPEA, dry DCM, relevant acid; (b) v—imidazole-carboxylic acid (lithium salt), EDCI, DMAP, dry DMF, 0–50 °C.

and the other was achieved by forming the acid chloride of the commercially available carboxylic acid, and reacting this with the amine in dry triethylamine (TEA) and DCM (Scheme 1a). Compound **3** was produced from NO_2-PyIm (**13**) that was reduced as described previously,[1] and then coupled to the lithium salt of imidazole-2-carboxylic acid[9] using EDCI and DMAP in dry DMF (Scheme 1b).

DNase I footprinting was performed to analyze the sequence specificity and affinity of the N-terminally modified derivatives of compound **1** (f-ImPyIm) using a 131 base pairs 5′-[^{32}P]-radiolabeled fragment containing the unique sequences *5′-ACGCGT-3′, *5′-ACCGGT-3′, *5′-ACGTGT-3′ and *5′-AGCGCT-3′ (the asterisk is used here to denote the sequence as read along the antisense strand). The sequence 5′-ACGCGT-3′ is the cognate site for all the compounds except **3**. This compound was synthesized as a control and was not expected to bind to this site. The concentration at which the footprint begins to appear for each compound can be seen in Table 1. Typical footprint titrations for f-ImPyIm and its structural derivatives compounds **4** (PyImPyIm), **6** (furan-ImPyIm), and **3** (ImPyIm) are shown in Figure 2. The titration result for compound **1** (f-ImPyIm) has been reported previously[1] but is included here for comparison with the titrations obtained for its N-terminal derivatives (Fig. 2A). Binding was observed at the cognate site (*5′-ACGCGT-3′) for all but one of the f-ImPyIm derivatives. Compounds **4** (Fig. 2B), **6** (Fig. 2C), and **7** (data not shown) produced the strongest binding at *5′-ACGCGT-3′, with the appearance of footprints from 0.5 μM. This was a 10-fold decrease in affinity relative to that observed with f-ImPyIm (**1**) at that site, which produced binding at the cognate from 0.05 μM. As expected compound **3** did not show binding at any of the sites examined (Fig. 2D). A large secondary binding site containing the sequences *5′-ACGTGT-3′ and *5′-AGCGCT-3′ was observed at concentrations of 5 μM and above with compounds **4**, **6**, **7**, and **8**. This was also observed with compound **1**, however, only f-ImPyIm (**1**) exhibited a third binding site at *5-ACCGGT-3′, with binding occurring from 5 μM. Thus the DNase I footprint titrations indicated that whilst compounds **4** (PyImPyIm), **6** (furan-ImPyIm), and **7** (benzofuran-ImPyIm) demonstrated a reduced affinity for their cognate site relative to f-ImPyIm (**1**), they showed an increased selectivity for this sequence particularly for PyImPyIm (**4**).

Thermal melting studies to determine whether the compounds stabilize duplex DNA at elevated temperatures were attempted with the previously described compounds using the cognate ACGCGT DNA and a non-cognate DNA sequence AAATTT. Further studies with compound **7** were discontinued due to solubility issues not encountered during the footprinting assay. The results (Table 1) showed that all tested compounds (excluding compound **3**) stabilized ACGCGT as predicted. An interesting result was obtained for PyImPyIm (**4**) with the cognate ACGCGT sequence. The ΔT_m value was found to be 4 °C. In contrast, furan compound **6** exhibited a ΔT_m of 10 °C. Both of these compounds produced a similar footprint, although compound **6** appeared less selective since at 0.5 μM additional footprints began to appear. This could be due to rotation about the furan and the carboxamido groups because in contract to *N*-methylpyrrole, furan lacks a methyl group. However, neither compounds **4** nor **6** stabilized AAATTT by measuring the T_m value.

Table 1. Footprinting and T_m data for N-terminus modified compounds **4**–**10** and f-ImPyIm (**1**)

Compound	Footprinting[a] (μM)	ΔT_m[b] (°C)		
		ACGCGT	AAATTT	ACCGGT
5	5	1	0	—
6	0.5	10	0	—
7	0.5	Insoluble		
8	5	7	5	—
9	50	1	6	—
PyImPyIm (**4**)	0.5	4	0	0
10	0	0	0	—
f-ImPyIm (**2**)[1]	0.05	9	1	1

[a] All experiments were performed using the appropriate sequence of DNA.

[b] All experiments were performed in PO_40 buffer which contains 13 mM salt.

Compound **3** (ImPyIm) did not show any DNA stabilization with either DNA evaluated as predicted.[3a]

CD studies (the ability of chiral molecules to absorb circularly polarized light: the DNA/PA complex is chiral due to the DNA's inherent chirality) were performed on each of the described compounds (except compound **7**) (data not shown) and confirmed that each compound bound in the minor-groove of DNA due to the appearance of an induced band at 330 nm.[10] Again, PyImPyIm (**4**) produced the most impressive results (Fig. 3) among the tetraamides studied, especially for binding to its cognate ACGCGT. The significantly weaker induced CD bands when compound **4** was titrated into the solutions of ACCGGT and AAATTT provided additional indication of sequence selectivity.

Based on the data obtained from DNase I footprinting, and the initial biophysical evaluation, PyImPyIm (**4**) was selected for further investigation. Compound **7** was not examined further due to the solubility issues, and compound **6** was not continued due to the non-selective binding observed with the footprinting studies.

A further T_m experiment was performed with compound **4** using a similar non-cognate DNA (ACCGGT, Table 1) and confirmed that this compound is selective for ACGCGT. The CD experiment using this DNA also confirmed minimal binding to both non-cognate DNA's (Fig. 3). These findings correlate well with the ITC studies of PyImPyIm (**4**) with ACGCGT, ACCGGT, and AAATTT. Figure 4A shows the thermograms and integrated data for all three DNA's tested. Compound **4** does bind to the cognate DNA sequence (ACGCGT) as indicated by the heat of reaction. No heat was produced with the two non-cognates mentioned above (Fig. 4). Table 2 shows the data relating to ITC. ΔC_p was calculated from an ITC experiment performed at three different temperatures (15, 25, and 35 °C) with ACGCGT and was found to be -146 cal mol^{-1} K^{-1}. A negative ΔC_p of this magnitude is consistent with minor-groove binders, including f-ImPyIm (**1**), which has a heat capacity of -142 cal mol^{-1} K^{-1}.[1] However, it is worth noting that the thermodynamic contributions are quite different between these two polyamides. The binding of f-ImPyIm to ACGCGT is strongly driven by enthalpy, whereas the entropic component for PyImPyIm is almost two-times the enthalpic contribution. This distinct difference is interesting because it again demonstrates the importance of the formamido group, which greatly adds to the binding constant.[3a] Even though the exact mechanism by which the formamido group contributes to binding affinity is unknown, the results suggest that it is perhaps mediated through enthalpic effects, for example, through dipole–dipole interactions.[11a]

Molecular modeling studies show that PyImPyIm (**4**) binds in the minor-groove as a dimer in an overlapped fashion to the sequence ACGCGT depicted in Figure 5. The PyImPyIm (**4**) dimer was modeled by modifying the f-ImPyIm (**1**) dimer from Buchmueller et al.,[1] and docked into $(CCACGCGTGG)_2$ using SYBYL 7.0. The complex was solvated and subjected to molecular dynamics simulation using AMBER 9.0 (*xleap* subprogram). A 45 ps constant pressure simulation was conducted and then the 5 ns production run was initiated. The structure and conformation of the complex given in Figure 5 are in good agreement with the complex of f-ImPyIm bound to the same oligonucleotide, in which the anti-parallel stacked dimeric tetraamides sit snugly in the minor-groove of the DNA. In this conformation, the ligands are able to make specific van der Waals contacts and hydrogen bonds with groups on the floor of the minor-groove. The complex given in Figure 5 is also remarkably similar to previously reported complexes of tetraamides bound the their cognate oligonucleotides, in terms of how well and deep the molecules fit in the minor-groove, limited distortion on the B-conformation, the extent to which the polyamide span over six base pairs, and the way the C-terminus dimethylamino moiety bind to the DNA. These previously disclosed complexes were determined by NMR and modeling studies (ImImPyPy-β-Dp)$_2$/5′-CCAGGCCTGG-3′ (β = β-alanine, Dp = dimethylaminopropylamide)[11b] or by a single crystal X-ray crystallography (ImImPyPy-β-Dp)$_2$/5′-CCAGGCCTGG-3′[12] as well as the complex of $(ImHpPyPy)_2$ with 5′-CCAGTACTGG-3′, in which Hp is *N*-methyl-3-hydroxypyrrole.[13]

The heat capacity (ΔC_p) arising from changes in hydrophobicity was estimated using solvent accessible surface area (SASA) calculations and the results are given in Table 3, and employing Eqs. 2–4. The results obtained from these calculations per mole of ligand were found to be: -224 ± 41, -302 ± 23, -278 ± 38 cal mol^{-1} K^{-1}. The calculated values are close to the experimental value providing credibility on the structure and conformation of the complex of PIPI and ACGCGT determined from modeling studies.

Accurate binding constants of compound **4** for its cognate and non-cognate were determined using the surface plasmon resonance (SPR) biosensor method, using a similar protocol to that previously reported.[1,3–5] The results given in Figure 6 provided conclusive evidence for the exquisite specificity of PyImPyIm (**4**) for ACGCGT.

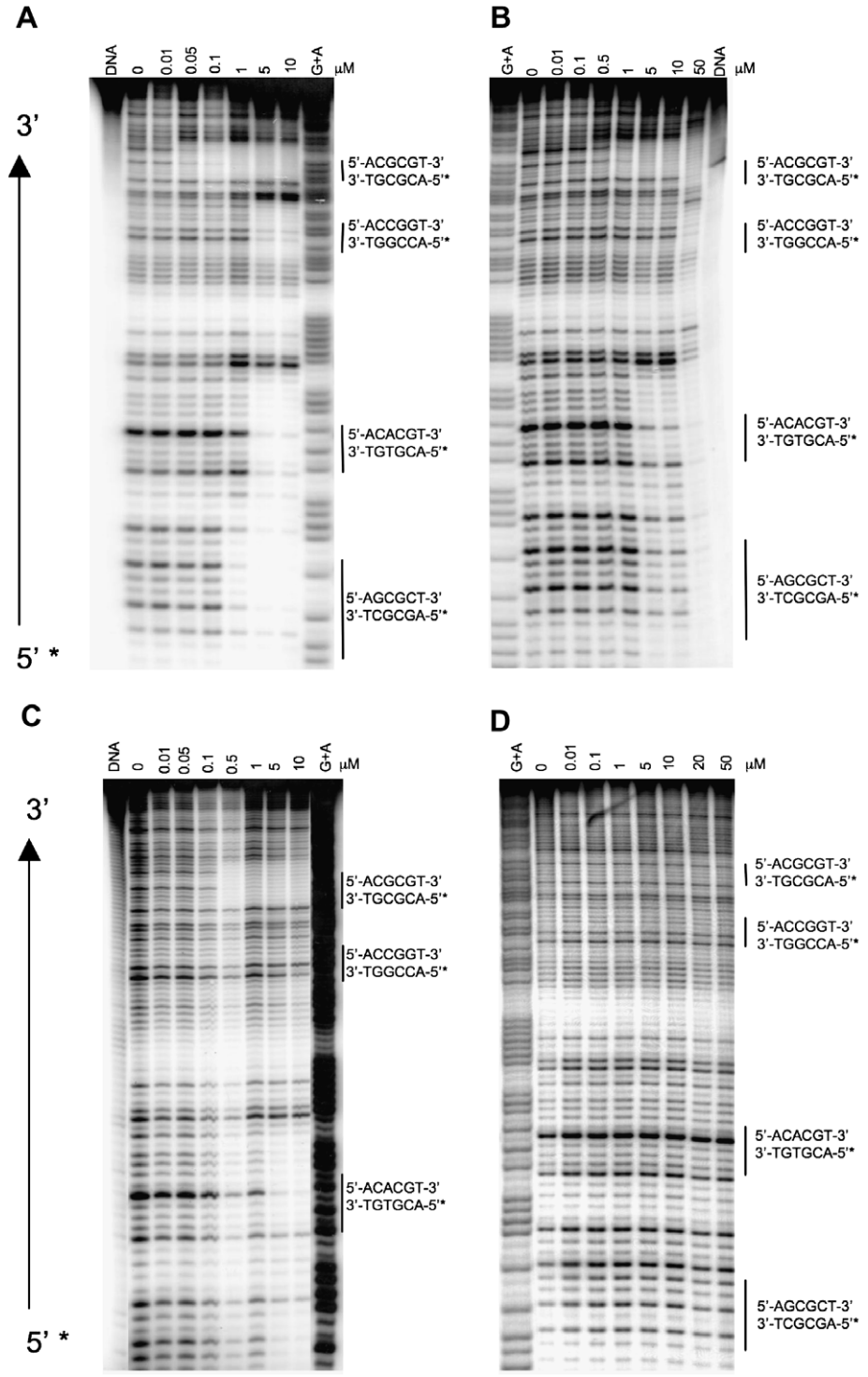


Figure 2. DNase I footprinting of (A) compound **1** (f-ImPyIm), (B) compound **4** (PyImPyIm), (C) compound **6** (furan-ImPyIm) and (D) compound **3** (ImPyIm) on the antisense strand of the 5′^{32}P-labeled 131 bp fragment. All reactions contain 1000 cps DNA fragment, 10 mM Tris, pH 7, 1 mM EDTA, 50 mM KCL, 1 mM MgCl2, 0.5 mM DTT, 20 mM Hepes. DNA denotes undigested DNA and G+A the purine sequencing lane. The sites *5-ACGCGT-3′, *5′-ACCGGT-3′, *5′-ACGTGT-3′ and *5′-AGCGCT-3′ are indicated by solid bars.

The sensorgram shows a fast on- but a relatively slow off-rate (Fig. 6A). The binding constant (K_{eq}) of compound **4** to ACGCGT was calculated using a 2:1 fit parameters to be $7.1 \times 10^6\ M^{-1}$. However, no signal was observed for the flow of PyImPyIm to sensor chips containing the ACCGGT and AAATTT hairpins, at the ligand concentrations employed in the studies. Hence, the binding constants could not be estimated. The binding constant with ACGCGT is in the same range as that obtained by the footprinting experiment. The binding constant of PyImPyIm (**4**) to ACGCGT is, however, reasonable and is consistent with the magnitude of several tetraamides reported previously. For instance, ImPyImPy-Dp (Dp is dimethylaminopropylamide),

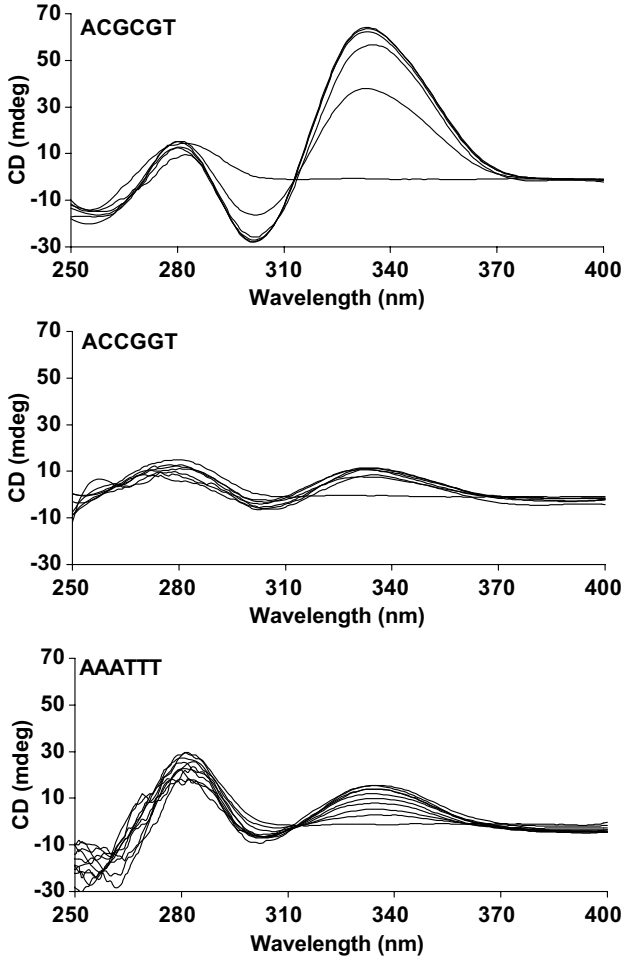


Figure 3. CD data for PyImPyIm (**4**) with ACGCGT (top), ACCGGT (middle) and AAATTT (bottom). Data were obtained using PO_45 buffer.

essentially an isomer of compound **4**, gave a binding constant of $3.6–3.8 \times 10^5\ M^{-1}$ for TGCGCA and AGCGCT according to quantitative MPE-footprinting studies.[14] Moreover, ImPyImPy-Dp was not highly sequence selective; it was found to recognize the sequence AACGCT that contained a single base pair mismatch, albeit at a threefold lower affinity. Additionally, ImPyImPy-β-Dp and ImImPyPy-β-Dp interacted with their cognates: TGCGCA and TGGCCA, at affinities of $<2 \times 10^5\ M^{-1}$ and $2.4 \times 10^7\ M^{-1}$, respectively.[15] It is interesting to note that, consistent with the 'core binding rules' reported by the authors' laboratories, polyamides bearing a –ImPy– core exhibit strong affinity for the corresponding –GC– site.[1,3b,4] Accordingly, ImImPyPy-β-Dp, which binds to a –GC–core site, demonstrated significantly higher affinity than ImPyImPy-β-Dp whose core site is –CG–. Moreover, ImImPyPy-β-Dp displayed a >100-fold greater affinity for its cognate than its non-cognates, 5′-TGCGCA-3′ and 5′-TGGGGA-3′.

3. Conclusion

In comparing the DNA binding properties of PyImPyIm (**4**) to f-ImPyIm (**1**), the SPR and footprinting results are worth emphasizing. Although PyImPyIm has a lower binding affinity than f-ImPyIm (**1**) ($1.9 \times 10^8\ M^{-1}$)[1] to the same ACGCGT sequence, it exhibits enhanced sequence specificity. Even though, f-ImPyIm (**1**) gave a binding affinity of 2.2×10^5 and $5.3 \times 10^4\ M^{-1}$ for ACCGGT and AAATTT, respectively,[1] as confirmed by DNase I footprinting studies, PyImPyIm (**4**) did not produce binding to these two sequences by footprinting or SPR studies. The results strongly indicate that tetraamide-based structures are excellent templates for the design of novel DNA sequence specific agents for biological activity, yet having acceptable binding affinity. As we have recently demonstrated, f-PyImPy, a simple triamide with a relatively low binding affinity of $2 \times 10^5\ M^{-1}$ for its cognate ATCGAT, was able to enter cells, move into the nucleus, and up-regulate the expression of the topoisomerase IIα gene in NIH 3T3 cells grown in confluence.[16] This strategy of designing gene control agents offers the advantage of developing small molecules, which unlike the larger linked molecules,[17] have favorable cellular permeability properties. Studies to develop tetraamides for the regulation of biologically relevant genes in anti-cancer drug discovery are in progress and the results will be reported in due course.

4. Experimental

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 disks. ^1H NMR spectra were obtained using a Varian Unity Inova 400 or 500 mHz 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.

4.1. Synthesis

4.1.1. PyImPyIm (4)

4.1.1.1. Method 1. Compound NO_2-ImPyIm (**11**, 60 mg, 0.124 mmol) was dissolved in cold MeOH and reduced to the corresponding amine via atmospheric catalytic hydrogenation using 5% Pd/C (30 mg). The catalyst was removed by filtration over celite and washed thoroughly with MeOH. The solvent was removed by evaporation and residual MeOH removed by co-evaporation with dry CH_2Cl_2 (3× 2 mL). The resulting amine was stored under vacuum, protected from light until required. DCC (70.9 mg, 0.37 mmol), DMAP (1.5 mg, 0.012 mmol) and 1-methy-2-pyrrole-

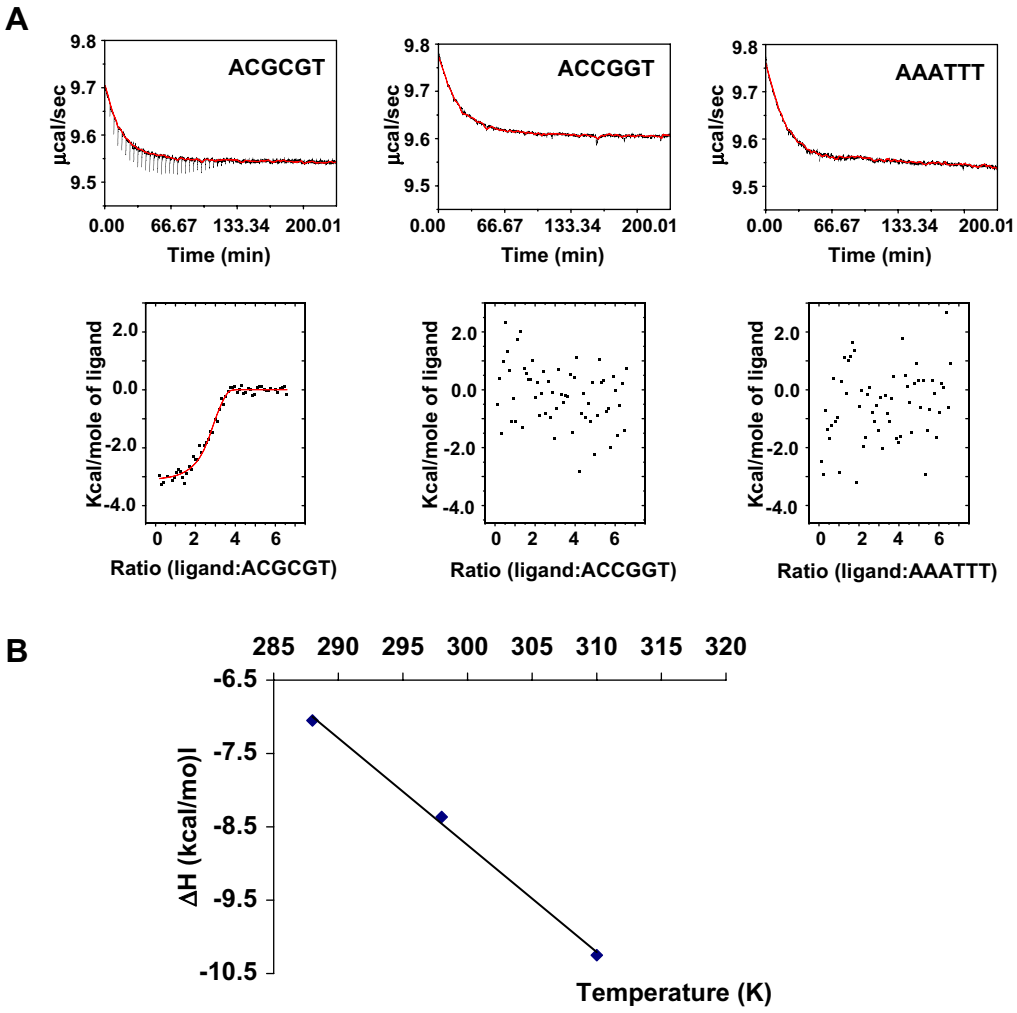


Figure 4. (A) ITC example thermograms (top) and integrated data (bottom) of PyImPyIm (**4**) with ACGCGT (left), ACCGGT (middle) and AAATTT (right); (B) plot of ΔH versus temperature in K for PyImPyIm (**4**) with ACGCGT at different temperatures (15, 25, 37, and 45 °C) used to calculate ΔC_p.

Table 2. ITC and SPR data for PyImPyIm (**4**) using ACGCGT

Compound	K_{eq}[a] (M^{-1})	ΔH[b] (kcal mol^{-1})	ΔG[c] (kcal mol^{-1})	$T\Delta S$[d] (kcal mol^{-1})
PyImPyIm (**4**)	7.1×10^6	−3.2	−9.3	6.1
f-ImPyIm (**1**)[1]	1.9×10^8	−7.6	−11.2	3.6

[a] Determined by SPR.
[b] Determined by ITC.
[c] Calculated from $\Delta G = -RT\ln K_{eq}$.
[d] Calculated from $\Delta G = \Delta H - T\Delta S$.

carboxylic acid (38.6 mg, 0.309 mmol) were dried under vacuum for ~2 hr and then placed under Ar. The amine was dissolved in dry DMF (1.5 mL), added to the dry reagents and the reaction mixture stirred under argon, protected from light, for 4 days at rt followed by heating at 50 °C (oil) for 2 days. The solvent was removed using Kuglröhr distillation and the resulting residue dissolved in $CHCl_3$ and washed with water (adjusted to pH ~ 12 with aq 1 M NaOH). The aqueous layer was back-extracted with $CHCl_3$ (3×) and EtOAc (2×). The crude product was purified by column chromatography (silica, gradient: $CHCl_3$/MeOH) to yield a yellow/brown crystalline solid (21.5 mg, 33%), mp decomp. 150 °C: IR (neat) υ 3328, 2960, 2891, 1769, 1723, 1658, 1541, 1459, 1442, 1411, 1368, 1323, 1289, 1248, 1190, 1121, 1067, 1038 cm^{-1}; ^1H NMR($CDCl_3$): δ 2.33 (s, 6H), 2.55 (t, J = 6.00 Hz, 2H), 3.50 (q, J = 6.00 Hz, 2H), 3.98 (s, 3H), 4.01 (s, 3H), 4.04 (s, 3H), 4.10 (s, 3H), 6.16–6.18 (m, 1H), 6.78–6.82 (m, 2H), 7.29 (d, J = 2.0 Hz, 1H), 7.42 (s, 1H), 7.50 (s, 1H), 7.60 (br t, J = 5.2 Hz, 1H), 8.10 (br s, 1H), 8.14 (br s, 1H), 8.84 (br s, 1H); MS (ES^+) m/z (rel intensity) 564 ([M+H], 10%), 210 (20%), 281 (100%); HRMS $[M+H]^+$ calcd for $C_{26}H_{34}N_{11}O_4$ m/z 564.2795; found m/z 564.2791.

4.1.1.2. Method 2. The amine was obtained as stated in Section 4.1.1.1 (0.41 mmol). *N*-Methylpyrrole-2-carboxylic acid (106.7 mg, 0.85 mmol) was dissolved in

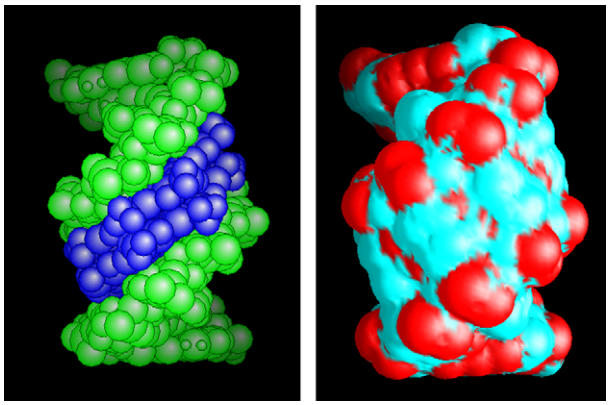


Figure 5. Molecular modeling of PyImPyIm (**4**) with ACGCGT. SASA calculated data for ΔC_p of PyImPyIm (**4**) with ACGCGT. (Red is polar, cyan is non-polar).

Table 3. The solvent accessible surface area (SASA) of PIPI/ACGCGT complex

	Polar	Non-polar	Total
Complex	2212	1914	4125
Ligand 1	180	933	1113
Ligand 2	192	920	1112
Free DNA[a]	2271	1650	3921
Change	−431	−1589	−2021

The area (Å^2) were computed using GRASP with Cornell et al. radii using a probe of 1.7683 Å.[21]

[a] The values were taken from previous publication[1] calculated with a NMR structure.[26]

dry THF (2 mL) and cooled to 0 °C (ice/H_2O) and dry TEA (0.1 mL ≡ 83 mg, 0.82 mmol) was added. The reaction mixture was stirred for ~15 min at rt and then cooled to 0 °C (ice/H_2O) and $SOCl_2$ (1 mL) was added and the reaction mixture stirred at 0 °C for 1 h, followed by 45 min at rt. The solvents were removed by aspiration followed by co-aspiration with dry DCM (3× 2 mL). The amine was dissolved in dry DCM (7 mL) and TEA (0.1 mL ≡ 83 mg, 0.82 mmol) and cooled to 0 °C (ice/H_2O) and the acid chloride [dissolved in dry DCM (6 mL)] was added dropwise to the amine mixture and allowed to warm to rt overnight with stirring. Work-up and purification of this compound was achieved in the same manner as above to yield PyIm-PyIm (**4**) as a yellow/brown crystalline solid (23.3 mg, 40%).

4.2. General procedure for the synthesis of compounds 5–9

NO_2-ImPyIm (**11**) (72 mg, 0.15 mmol) was reduced following the same procedure as outlined in Section 4.1.1.1 except cold MeOH (~40 mL) and 5% Pd/C (100 mg) were used. The resultant amine was dissolved in dry CH_2Cl_2 (6 mL). PyBOP (28.7 mg, 0.055 mmol) and the required carboxylic acid (1.2 eq.) were dried under vacuum for ~2 h. The vial was purged with argon and the amine solution (1 mL), followed by diisopropylamine (40 μL ≡ 28.8 mg, 0.28 mmol) were added to the dry reagents and stirred under an argon atmosphere for ~48 h at rt protected from light. The solution was diluted with CH_2Cl_2 and washed with H_2O and the aqueous layer adjusted (pH ~ 11, aq NaOH) and then

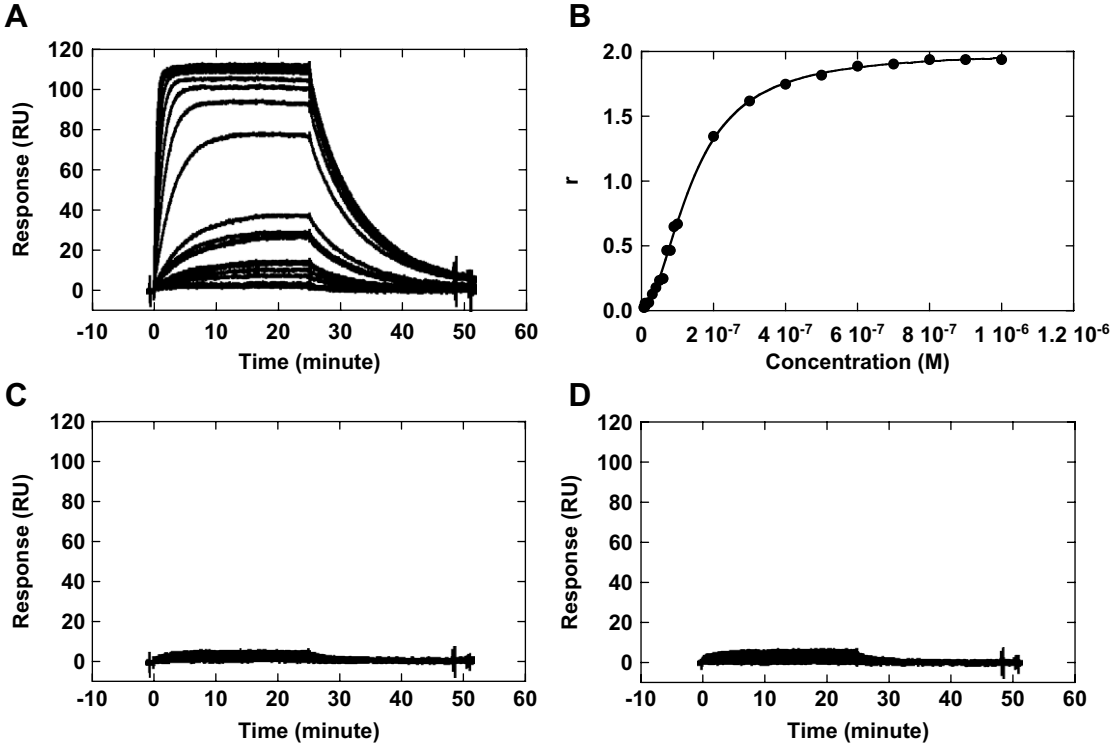


Figure 6. SPR of PyImPyIm (**4**) with ACGCGT (A) and its corresponding fit (B). Cooperative binding of two PyImPyIm molecules per ACGCGT hairpin is observable. On the same scale, the sensorgrams for the interaction with the ACCGGT (C) and AAATTT (D) are shown for comparison.

back-extracted with CH_2Cl_2 (3×) and EtOAc (2×). The organic layers were combined and dried (Na_2SO_4) and the solvent was removed. The crude product was purified by flash-column chromatography (silica, gradient CH_2Cl_2/MeOH).

4.2.1. Compound 5. 3-Methylpicolinic acid (41 mg, 0.3 mmol), appearance (67 mg, 78%), mp >300 °C: IR (neat) υ 371, 2925, 1658, 1528, 1466, 1440, 1253, 1124, 1098 cm^{-1}; ^1H NMR ($CDCl_3$) δ 2.36 (s, 6H), 2.58 (t, J = 6.50 Hz, 2 H), 2.82 (s, 3H), 3.98 (s, 3H), 4.04 (s, 3H), 4.12 (s, 3H), 6.81 (d, J = 1.5 Hz, 1H), 7.37–7.39 (m, 2H), 7.42 (s, 1H), 7.57 (br t, J = 5.2 Hz), 7.62 (s, 1H), 7.62–7.65 (m, 1H), 8.16 (br s, 1H), 8.48 (d, J = 4.5 Hz, 1H), 8.94 (s, 1H), 10.5 (br s, 1H); MS (ES^+) *m/z* (rel intensity) 562 ([M+H], 70%), 282 (100%), 136 (45%); HRMS $[M+H]^+$ calcd for $C_{27}H_{34}N_{11}O_4$ *m/z* 576.2795; found *m/z* 576.2797.

4.2.2. Compound 6. 2-Furoic acid (33 mg, 0.3 mmol), appearance, (65 mg, 78%) mp > 300 °C: IR (Nujol) υ 3393, 3165, 1706, 1649, 1540, 1244, 1125 cm^{-1}; ^1H NMR ($CDCl_3$) δ 2.32 (s, 6H), 2.53 (t, J = 6.0 Hz, 2H), 3.48 (q, (under H_2O peak), J = 6.0 Hz, 2H), 3.98 (s, 3H), 4.04 (s, 3H), 4.12 (s, 3H), 6.58–6.59 (m, 1H), 6.78 (d, J = 1.5 Hz, 1H), 7.25 (s, 1H), 7.42 (s, 1H), 7.54 (s, 2H), 8.02 (br s, 1H), 8.47 (br s, 1H), 8.82 (br s, 1H); MS (ES^+) *m/z* (rel intensity) 551 ([M+H], 60%), 276 (100%); HRMS $[M+H]^+$ calcd for $C_{25}H_{31}N_{10}O_5$ *m/z* 551.2479; found *m/z* 551.2468.

4.2.3. Compound 7. Benzofuran-2-carboxylic acid (49 mg, 0.3 mmol), appearance (73 mg, 81%), mp > 300 °C: IR (Nujol) υ 3402, 3163, 1663, 1544, 1378, 1253, 1212, 1124 cm^{-1}; ^1H NMR (DMSO-*d6*) δ 2.77 (br s, 2H), 3.33 (s, 6H), 3.54 (br s, 2H), 3.87 (s, 3H), 3.96 (s, 3H), 4.02 (s, 3H), 7.16 (d, J = 2.0 Hz,1H), 7.37 (t, J = 7.6 Hz, 1H), 7.40 (d, J = 2.0 Hz, 1H) 7.50 (td, J = 7.6 Hz, 1.6 Hz, 1H), 7.55 (s, 1H), 7.65 (s, 1H), 7.71 (d, J = 8 Hz, 1H), 7.60 (d, J = 7.6 Hz, 1H), 7.87 (br s, 1H), 9.22 (br s, 1H), 10.03 (br s, 1H), 10.22 (br s. 1H), 10.96 (br s, 1H); MS (ES^+) *m/z* (rel intensity) 601 ([M+H], 90%), 301 (100%); HRMS $[M+H]^+$ calcd for $C_{29}H_{33}N_{10}O_5$*m/z* 601.2635; found *m/z* 601.2634.

4.2.4. Compound 8. 5-Methoxyindole-2-carboxylic acid (58 mg, 0.3 mmol), appearance, (46 mg, 49%), mp > 300 °C: IR (Nujol) υ 3392, 1710, 1653, 1544, 1305, 1253, 1217, 1155 cm^{-1}; ^1H NMR (DMSO-d_6) δ 2.19 (s, 6H), 2.40 (m. 2H), 3.43 (m, 2H), 3.77 (s, 3H), 3.87 (s, 3H), 3.95 (s, 3H), 4.02 (s, 3H), 6.87 (dd, J = 9.00 Hz, 2.50, 1H), 7.10 (d, J = 2.50 Hz, 1H), 7.17 (d, J = 2.00 Hz, 1H), 7.34 (s, 1H), 7.37 (s, 1H), 7.40 (d, J = 2.00 Hz, 1H), 7.52 (s, 1H), 7.62 (s, 1H), 7.77 (m, 1H), 9.94 (s, 1H), 10.39 (s, 1H), 10.81 (s, 1H), 11.55 (s, 1H); MS (ES^+) *m/z* (rel intensity) 630 ([M+H], 60%), 316 (100%).

4.2.5. Compound 9. 3,4,5-Trimethoxyindole-2-carboxylic acid (41 mg, 0.3 mmol), (34 mg, 33%), mp > 300 °C: IR (Nujol) υ 3402, 1653, 1539, 1305, 1233, 1118, 1041 cm^{-1}; ^1H NMR (DMSO-d_6) 2.08 (s, 6H), 2.40 (m, 2H), 3.39 (m, 2H), 3.79 (s, 9H), 3.87 (s, 1H), 3.95 (s, 3H), 4.01 (s, 3H), 6.92 (d, J = 2.00 Hz, 1H), 7.16 (d, J = 2.00 Hz, 1H) 7.26 (d, J = 2.50 Hz, 1H), 7.40 (d, J = 2.00 Hz, 1H), 7.53 (s, 1H), 7.61 (s, 1H), 7.98 (br, 1H), 9.95 (s, 1H), 10.31 (br, 1H), 10.73 (s, 1H), 11.49 (s, 1H); MS (ES^+) *m/z* (rel intensity) 689 ([M+H], 40%), 346. ($[M+2H]^+$, 100%).

4.2.6. Compound 3. NO_2-PyIm (**13**) (60.1 mg, 0.165 mmol) was reduced in the presence of 5% Pd/C (31 mg), cold MeOH (~20 mL) as described in Section 4.1.1.1. *N*-Methyl-imidazole-carboxylic acid (31.2 mg, 0.380 mmol), EDCI (95.1 mg, 0.496 mmol), DMAP (2.05 mg, 0.017 mmol), and dry DMF (2 mL) were stirred under argon at rt for 4 days and then for ~24 h at 50 °C. The same work-up and purification were performed as stated in Section 4.1.1.1 and yielded **3** as a white solid (20.4 mg, 28%), mp 205–208 °C: IR (neat) υ 3393, 2947, 1659, 1555, 1534, 1472, 1259, 1125, 1093, 798 cm^{-1}; ^1H NMR ($CDCl_3$) 2.31 (s, 6H), 2.53 (t., J = 6.3 Hz, 2H), 3.48 (q, J = 6.3 Hz, 2H), 3.97 (s, 3H), 4.04 (s, 3H), 4.11 (s, 3H), 6.80 (d, J = 1.5 Hz, 1H), 7.01 (s, 1H), 7.06 (s, 1H), 7.27 (s, overlapping with $CHCl_3$, 1H), 7.41 (s, 1H), 7.54 (br t, 1H), 8.09 (br s, 1H), 9.13 (br s, 1H); MS (ES^+) *m/z* (rel intensity) 442 ([M+H], 70%), 222 (100%), 199 (60%); HRMS $[M+H]^+$ calcd for $C_{20}H_{28}N_9O_3$*m/z* 442.2315; found *m/z* 442.2304.

4.3. Biophysical

4.3.1. Buffers. Ten millimolar of phosphate buffers were prepared at the following salt (NaCl) concentrations: 12.5 mM ($PO_4$0), 50 mM ($PO_4$5), 200; mM ($PO_4$20) with the addition of 1 mM EDTA, pH 6.2.

4.3.2. Thermal denaturation (T_m). Thermal denaturation data were obtained using the Cary 100 BioMelt (Varian) spectrophotometer with DNA (1 μM) in $PO_4$0 and the required compound (3 μM), using the procedure previously reported.[3a] Thermal melts were obtained for each compound using ACGCGT and AAATTT and ACCGGT for compound **4**.

4.3.3. Circular dichroism (CD). CD studies were performed using the Olis DSM20 instrument. Each run was performed over 400–220 nm wavelength range (180 increments) and an integration time of 1 s and the average of two scans were used for analysis. The required compound (500 μM in double distilled H_2O) was titrated in 1 molar equivalents into the required DNA (160 μL of 9 μM DNA), in $PO_4$5, until saturation was observed. Data analysis was performed as previously described.[3a] CD experiments were performed for all the compounds with ACGCGT and AAATTT and ACCGGT with compound **4**.

4.3.4. Isothermal titration microcalorimetry (ITC). ITC analysis was performed using a VP_ITC micromalorimeter (MicroCal). Compound **4** was dissolved in $PO_4$20 and the instrument equilibrated at the noted temperature. After an initial delay of 300 s, compound **4** (100 μM) was titrated, via 50–60 injections (3 μL for

7.2 s, repeated every 240 s), into 2 μM DNA ($PO_4$5). The data were analyzed using the same method as previously reported.[1] Origin 7.0 was used and the area under the curve integrated as a function of time. A linear fit was then employed and this was subtracted from the reaction integrations to normalize for non-specific heat components. The two-sets-of-sites, non-sequential model using the MicroCal version of Origin 7.0 was used.[18] ΔG was calculated from

$$\Delta G = -RT \ln K_{eq} \quad (1)$$

where R is 1.987 cal mol^{-1} K^{-1} and T is measured in K. The DNA used in this experiment was ACGCGT, AAATTT, and ACCGGT.

4.3.5. Solvent accessible surface areas (SASA). The SASA calculation procedures have been previously described.[1,19] In brief, ions were removed and carbon, carbon-bound hydrogen, and phosphorus atoms were assigned as non-polar and the rest polar. The PIPI/CGCG complex (a total of 780 atoms: 630 from DNA, 150 from ligands) consists of 284 polar (p) atoms (244 from DNA, 40 from ligands) and 496 non-polar (np) atoms (386 from DNA, 110 from ligands). The solvent accessible surface area (SASA) was calculated with GRASP[20] using a probe radius of 1.7683 Å and Cornell et al. radii.[21]

The heat capacity change (cal/(mol K)) arisen from the polar/non-polar area change was calculated with three different models Eq. 2,[22] Eq. 3,[23] and Eq. 4.[24]

$$\Delta A = A_{complex} - (A_{free\_dna} + A_{ligand\_1} + A_{ligand\_2})$$

$$\Delta C_{p\text{-}SASA} = (0.32 \pm 0.04)\Delta A_{np} - (0.14 \pm 0.04)\Delta A_p \quad (2)$$

$$\Delta C_{p\text{-}SASA} = (0.45 \pm 0.02)\Delta A_{np} - (0.26 \pm 0.03)\Delta A_p + (0.17 \pm 0.07)\Delta A_{OH} \quad (3)$$

$$\Delta C_{p\text{-}SASA} = (0.382 \pm 0.026)\Delta A_{np} - (0.121 \pm 0.077)\Delta A_p \quad (4)$$

Per mole of ligand: $\Delta C_{p\text{-}SASA-1} = \Delta C_{p\text{-}SASA}/2$

This value ($\Delta C_{SASA}/2$) is then compared to the total heat capacity change ($\Delta C_{p-total}$) observed with calorimetric titration.

4.3.6. Surface plasmon resonance. Biosensor chip surface preparations and biotinated DNA immobilizations were conducted as previously described.[25] Biotin labeled DNA hairpins (5′-biotin-GA**ACGCGT**CCTCTG**ACGCGT**TC-3′, 5′-biotin-GA**ACCGGT**CCTCTG**ACCGGT**TC-3′, and 5′-biotin-CG**AAATTT**CCTCTG**AAATTT**CG-3′). The experiments were conducted in a phosphate buffer at 200 mM Na^+ and 0.0005% P20 surfactant. In a typical experiment, 250 μL samples at different concentrations were injected onto the chip surface with a flow rate of 10 μL/min and 1500-s dissociation period. The surface was regenerated with a glycine, pH 2.5, solution and multiple buffer injections. Steady-state analyses were conducted and the response units were converted to mole of ligand per mole of compound as previously described.[3a] The data were fitted with Eq. (5) to obtain macroscopic binding constants.

$$r = (K_1 \times C_{free} + 2 \times K_1 \times K_2 \times C_{free}^2)/(1 + K_1 \times C_{free} + K_1 \times K_2 \times C_{free}^2) \quad (5)$$

4.3.7. Molecular modeling. Two PyImPyIm (**4**) polyamide molecules were docked into the minor-groove of a DNA decamer $(CCACGCGTGG)_2$. The DNA decamer coordinates were obtained from the most representative conformation of this DNA (pdbid:1KKV) determined by NMR analysis.[26] The PyImPyIm (**4**) dimer was modeled by modifying the f-ImPyIm (**1**) dimer from Buchmueller et al.[1] The PyImPyIm (**4**) dimer was docked into $(CCACGCGTGG)_2$ by monitoring the distance and electrostatic energy and conducting torsional minimization of the DNA and PyImPyIm (**4**) using SYBYL 7.0. Using AMBER 7.0 (*xleap* subprogram),[27] 30 Na^+ ions and 14 Cl^- ions were added to neutralize the system, resulting in 0.15 M NaCl, which correspond to the experimental conditions. The complex was solvated with 3,546 TIP3P waters[28] and the initial system had approximate dimensions of 50 × 50 × 56 Å. The system was equilibrated and 25 kcal mol^{-1} Å^{-2} constraints were placed on the DNA, the PyImPyIm (**4**) and the counterions. In addition, NMR analysis provided distance restraints of 3.3 Å (±15%) between the αCH_2 of PyImPyIm (**4**) and A_3H2 of $(CGCGnmr)_2$.[1] The temperature of the system was gradually raised to 300 K and a series of constant volume simulations were conducted while gradually releasing the energetic constraints and distance restraints.[29] Distance restraints on the DNA termini were added to ensure helical integrity during the final phases of structure optimization. A 45 ps constant pressure simulation was conducted and then the 5 ns production run was initiated with no restraints.

4.3.8. DNase I footprinting. A radiolabeled DNA fragment of 131 base pairs was generated by polymerase chain reaction as described previously by O'Hare *et al.*[30] The resulting labeled fragment was purified on a Bio-Gel P-6 column (Bio-Rad) followed by agarose gel electrophoresis and isolated using a Geneclean II kit (VWR) according to the manufacturer's instruction. DNase I footprint reactions were performed by incubating polyamides with 1000 cps of 5′-single end-labeled fragment in 10 mM Tris, pH 7.0, 1 mM EDTA, 50 mM KCl, 1 mM $MgCl_2$, 0.5 mM DTT, 40 mM Hepes, pH 7.9, at room temperature for 30 min in a total volume of 50 μL. Cleavage by DNase I was carried out at room temperature and initiated by the addition of 2 μL (0.5 U) DNase I diluted in ice-cold 10 mM Tris, pH 7.0, from a stock solution (1U/μL, Promega) and 1 μL of a solution of 250 mM $MgCl_2$ and $CaCl_2$. The reactions were stopped after 3 min by the addition of 100 μL of a stop mix containing 200 mM NaCl, 30 mM EDTA, pH 8, 1% SDS. The cleavage products were precipitated in the presence of 1 μL of glycogen (20 mg/ml, Roche Diagnostics), washed once in 80% ethanol and dried. The samples were resuspended in formamide loading buffer (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) denatured for 5 min at 90 °C, cooled on ice and electrophoresed at 1,500 V for 2 h on a 10% acrylamide dena-

turing gel (Sequagel, National Diagnostics). The gels were dried under vacuum at 80 °C and exposed to film for 24 h (Super RX, Fuji).

Acknowledgments

The authors thank Hope College, Furman University, the NSF (CHE-0550992 and (C2259/A3083), Cancer Research UK, the National Science Foundation sponsored Research Site for Educators in Chemistry at UT-Knoxville and the Erskine College Bell Grant program for support.

References and notes

1. Buchmueller, K. L.; Bailey, S. L.; Matthews, D. A.; Taherbhai, Z. T.; Register, J. K.; Davis, Z. S.; Bruce, C. D.; O'Hare, C.; Hartley, J. A.; Lee, M. *Biochemistry* **2006**, *45*, 13551.
2. (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci.* **1989**, *86*, 5723; (b) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393.
3. (a) Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2153; (b) 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; (c) 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; (d) Lee, M.; Rhodes, A. L.; Wyatt, M. D.; Forrow, S.; Hartley, J. A. *Biochemistry* **1993**, *32*, 4237.
4. Brown, T.; Taherbhai, Z.; Sexton, J.; Sutterfield, A.; Turlington, M.; Jones, J.; Stallings, L.; Stewart, M.; Buchmueller, K.; Mackay, H.; O'Hare, C.; Hartley, J.; Nguyen, B.; Wilson, D.; Lee, M. *Bioorg. Med. Chem.* **2007**, *15*, 474.
5. Buchmueller, K. L.; Taherbhai, Z.; Howard, C. M.; Bailey, S. L.; Nguyen, B.; O'Hare, C.; Hochhauser, D.; Hartley, J. A.; Wilson, W. D.; Lee, M. *ChemBioChem* **2005**, *6*, 2305.
6. Zhan, Z.-Y. J.; Dervan, P. B. *Bioorg. Med. Chem.* **2000**, *8*, 2467.
7. Lee, M.; Krowicki, K.; Shea, R. G.; Lown, J. W.; Pon, R. T. *J. Mol. Recogn.* **1989**, *2*, 84.
8. Boger, D. L.; Johnson, D. S. *Proc. Natl. Acad. Sci.* **1995**, *92*, 3642.
9. Heckel, A.; Dervan, P. B. *Chem.—Eur. J.* **2003**, *9*, 3353.
10. (a) Lyng, R.; Rodger, A.; Norden, B. *Biopolymers* **1992**, *32*, 1201; (b) Lyng, R.; Rodger, A.; Norden, B. *Biopolymers* **1991**, *31*, 1709.
11. (a) Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *Science* **2004**, *226*, 646; (b) White, S.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 8756.
12. Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104.
13. Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Science* **1998**, *282*, 111.
14. Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1995**, *117*, 3325.
15. Swalley, S. E.; Baird, E. E.; Dervan, Peter. B. *J. Am. Chem. Soc.* **1997**, *119*, 6953.
16. Le, N. M.; Sielaff, A. M.; Cooper, A. J.; Mackay, H.; Brown, T.; Kotecha, M.; O'Hare, C.; Hochhauser, D.; Lee, M.; Hartley, J. A. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 6161.
17. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215.
18. Indyk, L.; Fisher, H. F. *Methods Enzymol.* **1998**, *295*, 350.
19. Nguyen, B.; Stanek, J.; Wilson, W. D. *Biophys. J.* **2006**, *90*, 1319.
20. Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins* **1991**, *11*, 281.
21. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M., jr.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
22. (a) Spolar, R. S.; Ha, J. H.; Record, M. T., Jr. *Proc. Natl. Acad. Sci.* **1989**, *86*, 8382; (b) Livingstone, J. R.; Spolar, R. S., ; Record, M. T., Jr. *Biochemistry* **1991**, *30*, 4237; (c) Spolar, R. S.; Livingstone, J. R.; Record, M. T., Jr. *Biochemistry* **1992**, *31*, 3947; (d) Spolar, R. S.; Record, M. T., Jr. *Science* **1994**, *263*, 777.
23. (a) Murphy, K. P.; Bhakuni, V.; Xie, D.; Freire, E. *J. Mol. Biol.* **1992**, *227*, 293; (b) Murphy, K. P.; Freire, E. *Adv. Protein Chem.* **1992**, *43*, 313; (c) Gomez, J.; Freire, E. *J. Mol. Biol.* **1995**, *252*, 337; (d) Gomez, J.; Hilser, V. J.; Xie, D.; Freire, E. *Proteins* **1995**, *22*, 404; (e) Habermann, S. M.; Murphy, K. P. *Protein Sci.* **1996**, *5*, 1229.
24. Ren, J.; Jenkins, T. C.; Chaires, J. B. *Biochemistry* **2000**, *39*, 8439.
25. Nguyen, B.; Tanious, F. A.; Wilson, W. D. *Methods* **2007**, *42*, 150.
26. Isaacs, R. J.; Rayens, W. S.; Spielmann, H. P. *J. Mol. Biol.* **2002**, *319*, 191.
27. Case, D. A.; Cheatham, T. E., III.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comp. Chem.* **2005**, *26*, 1668.
28. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
29. (a) Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *J. Am. Chem. Soc.* **2001**, *123*, 5044; (b) Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *J. Phys. Chem. B* **2001**, *105*, 3135; (c) Hamelberg, D.; Williams, L. D.; Wilson, W. D. *J. Am. Chem. Soc.* **2001**, *123*, 7745.
30. O'Hare, C. C.; Uthe, P.; Mackay, H.; Blackmon, K.; Jones, J.; Brown, T.; Nguyen, B.; Wilson, W. D.; Lee, M.; Hartley, J. A. *Biochemistry* **2007**, *46*, 11661.