Bioorganic & Medicinal Chemistry 16 (2008) 9145–9153

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Sequence specific and high affinity recognition of 5′-ACGCGT-3′ by rationally designed pyrrole-imidazole H-pin polyamides: Thermodynamic and structural studies

Hilary Mackay [a], Toni Brown [a], Peter B. Uthe [b], Laura Westrate [b], Alan Sielaff [b], Justin Jones [b], James P. Lajiness [a], Jerome Kluza [c], Caroline O'Hare [c], Binh Nguyen [d], Zach Davis [e], Chrystal Bruce [e], W. David Wilson [d], John A. Hartley [c], Moses Lee [a,b,*]

[a] Department of Chemistry, Hope College, 35 E. 12th Street, P.O. Box 9000, Holland, MI 49422, USA
[b] Department of Chemistry, Furman University, Greenville, SC 29613, USA
[c] Cancer Research UK Drug-DNA Interactions Research Group, UCL Cancer Institute, Paul O'Gorman Building, 72 Huntley Street, London WC1E 6BT, USA
[d] Department of Chemistry, Georgia State University, Atlanta, GA 30303, USA
[e] Department of Chemistry, Erskine College, Due West, SC 29639, USA

ARTICLE INFO

Article history:
Received 9 May 2008
Revised 29 August 2008
Accepted 10 September 2008
Available online 13 September 2008

Keywords:
Binding
Biophysical
DNA
H-pin
Sequence selectivity
Polyamide
SPR
ITC

ABSTRACT

Imidazole (Im) and Pyrrole (Py)-containing polyamides that can form stacked dimers can be programmed to target specific sequences in the minor groove of DNA and control gene expression. Even though various designs of polyamides have been thoroughly investigated for DNA sequence recognition, the use of H-pin polyamides (covalently cross-linked polyamides) has not received as much attention. Therefore, experiments were designed to systematically investigate the DNA recognition properties of two symmetrical H-pin polyamides composed of PyImPyIm (**5**) or f-ImPyIm (**3e**, f = formamido) tethered with an ethylene glycol linker. These compounds were created to recognize the cognate 5′-ACGCGT-3′ through an *overlapped* and *staggered binding motif*, respectively. Results from DNaseI footprinting, thermal denaturation, circular dichroism, surface plasmon resonance and isothermal titration microcalorimetry studies demonstrated that both H-pin polyamides bound with higher affinity than their respective monomers. The binding affinity of formamido-containing H-pin **3e** was more than a hundred times greater than that for the tetraamide H-pin **5**, demonstrating the importance of having a formamido group and the *staggered motif* in enhancing affinity. However, compared to H-pin **3e**, tetraamide H-pin **5** demonstrated superior binding preference for the cognate sequence over its non-cognates, ACCGGT and AAATTT. Data from SPR experiments yielded binding constants of $1.6 \times 10^8\ M^{-1}$ and $2.0 \times 10^{10}\ M^{-1}$ for PyImPyIm H-pin **5** and f-ImPyIm H-pin **3e**, respectively. Both H-pins bound with significantly higher affinity (ca. 100-fold) than their corresponding unlinked PyImPyIm **4** and f-ImPyIm **2** counterparts. ITC analyses revealed modest enthalpies of reactions at 298 K (ΔH of −3.3 and −1.0 kcal mol^{-1} for **5** and **3e**, respectively), indicating these were entropic-driven interactions. The heat capacities (ΔC_p) were determined to be −116 and −499 cal mol^{-1} K^{-1}, respectively. These results are in general agreement with ΔC_p values determined from changes in the solvent accessible surface areas using complexes of the H-pins bound to (5′-CC*ACGCGT*GG)$_2$. According to the models, the H-pins fit snugly in the minor groove and the linker comfortably holds both polyamide portions in place, with the oxygen atoms pointing into the solvent. In summary, the H-pin polyamide provides an important molecular design motif for the discovery of future generations of programmable small molecules capable of binding to target DNA sequences with high affinity and selectivity.

1. Introduction

The development of polyamide analogs of distamycin (**1**, Fig. 1) that can target specific sequences of DNA is an area of active research. Such compounds have the potential to be used as gene control agents by inhibition of native transcription factors at the target site.[1] Monomer polyamides, which bind DNA in a 2:1 (ligand:DNA) motif, have been shown to successfully target the desired sequence.[1c,2] A concern in this strategy is that the monomers may slip into extended staggered motifs, thus, creating a reading frame different to that intended.[3] The monomer may also bind in a mixed 2:1 and

* Corresponding author. Tel.: +1 616 395 7190; fax: +1 616 395 7923.
E-mail address: lee@hope.edu (M. Lee).

0968-0896/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2008.09.034

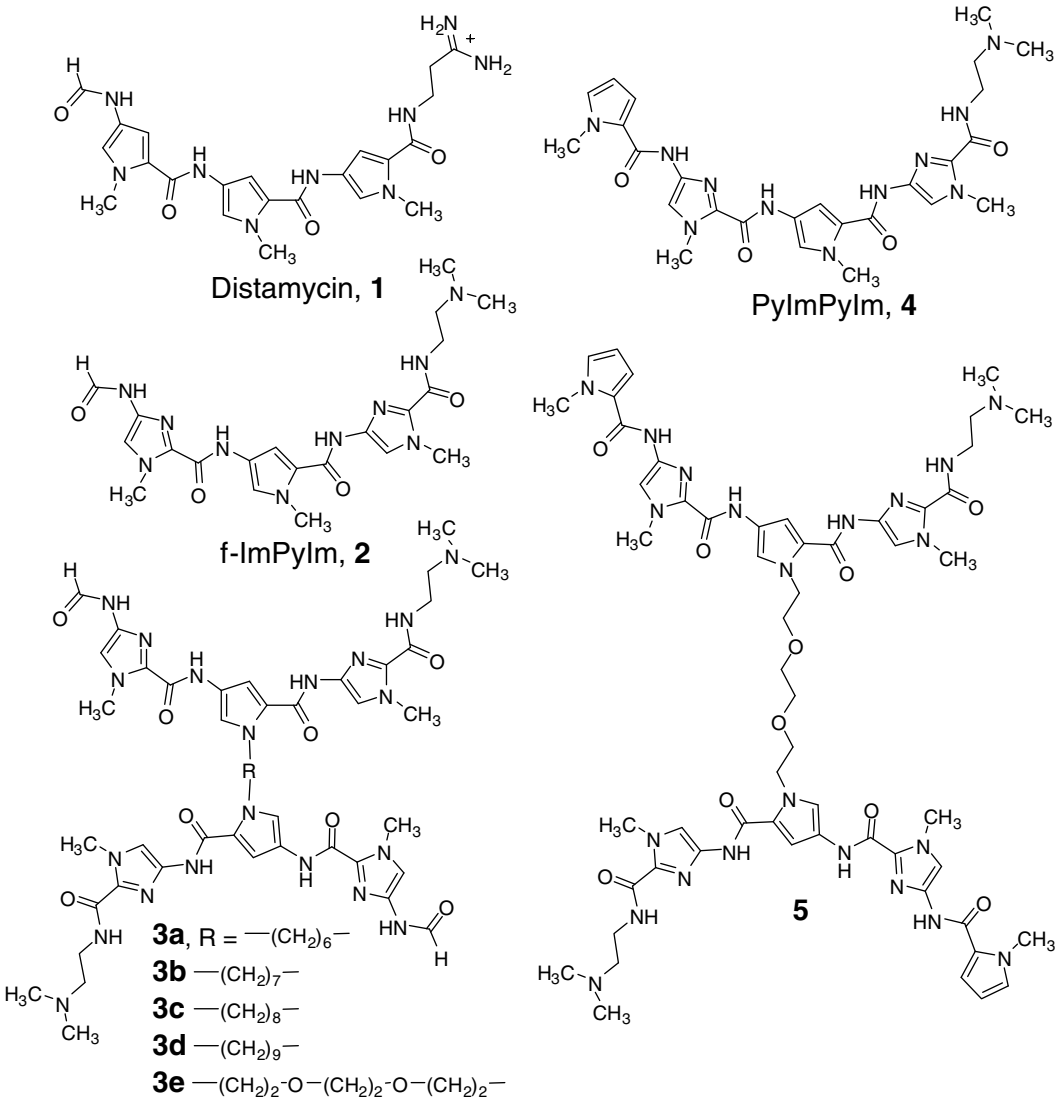


Figure 1. Structure of polyamides: distamycin **1**, f-ImPyIm **2**, f-ImPyIm H-pins **3a-e**, PyImPyIm **4**, and the polyethylene glycol PyImPyIm H-pin **5**.

1:1 fashion, thereby reducing the overall affinity. To overcome these issues, many strategies have been employed to tether both monomers together, such as the 'hairpin' design which links the monomers 'head'-to-'tail' using a flexible γ-butyric acid moiety.[1c,2,4] Hairpin polyamides have demonstrated excellent binding affinities and their binding characteristics have been extensively reviewed.[1,2] An alternative motif is the 'H-pin' which links the two monomers via the N1-nitrogen of the central heterocyclic units.[5] To date H-pins linked via alkyl chains of varying lengths (C3–C12) have been investigated as duplex DNA binders that interact with the minor-groove in the *overlapped motif* (the two polyamides units are stacked directly over each other and the linker directly joins two heterocycles across the polyamides, Fig. 2A).[5,6] However, weak binding affinities were reported and the compounds were not tested against DNAs that enable a different binding motif.[2,4–6]

The authors' laboratory has previously reported that the polyamide monomer formamido-imidazole-pyrrole-imidazole (f-ImPyIm, **2**, Fig. 1) binds as a dimer in *a staggered binding motif* (Fig. 2B), with exceptional affinity, to its cognate DNA sequence (5′-ACGCGT-3′).[7] In this motif, the polyamides units are stacked in an off-centered or staggered fashion and the linker joins two adjacent heterocycles in a diagonal arrangement. The 5′-ACGCGT-3′ sequence is of significant interest because it is present in the core sequence of the MluI cell-cycle box (MCB) transcriptional element found in the promoter of the human Dbf4 (huDbf4 or ASK , activator of S-phase kinase) gene. Dbf4 is the regulatory subunit of Cdc7 (cyclin dependent 7) kinase, and high levels of this kinase have been implicated for development of various cancers.[8] As part of our strategy to develop MCB-targeted polyamides, the synthesis of a series of H-pin polyamides composed of two f-ImPyIm monomers tethered via the central pyrrole using either a hydrocarbon

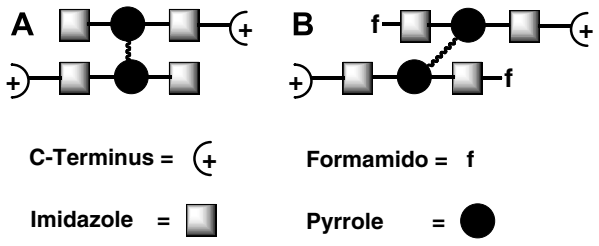


Figure 2. Diagram to show the 'overlapped' (A), and 'staggered' (B), H-pin motifs of polyamide minor-groove binding to duplex DNA. The wiggly line connecting the two polyamide units represents the linker.

chain of length C6–C9 or an ethylene glycol linker (**3a–e**, Fig. 1) was carried out.[9] These studies were designed to discover a suitable covalent linker on the H-pins that would enable the compounds to bind tightly and specifically to the cognate sequence through the *staggered binding motif*. It was determined via DNaseI footprinting, and analysis of molecular models, that the C8 and C9 hydrocarbon linked H-pins yielded the optimal length to allow binding in a staggered fashion. However, the alkyl linked compounds readily formed aggregates in aqueous solutions. The ethylene glycol linked H-pin **3e** overcame the solubility obstacle and was found to bind with exceptionally high affinity to the cognate sequence ($K = 2 \times 10^{10}\ M^{-1}$ by SPR). H-pin **3e** interacts with DNA in a similar manner to that of its monomer counterpart **2**, suggesting that the linker does not affect the DNA sequence recognition process.[9] The outstanding binding affinity of **3e** was somewhat compromised by the reduced selectivity of this compound in comparison to monomer **2**.[9] It must be noted that the Lown group had conducted some experiments on an H-pin related to **3c**; however, the H-pin was only tested against DNA through the *overlapped motif*.[5] The possibility that using such linkers would allow enough flexibility to bind in a *staggered binding motif* was not anticipated nor investigated.[2,4–6]

In an effort to design compounds with improved sequence specificity and to expand the repertoire of polyamide–DNA binding designs, the authors' laboratory reported the binding characteristics of a monomer tetraamide, PyImPyIm (**4**, Fig. 1).[10] The non-formamido compound **4** was anticipated to bind in an overlapped fashion, extending the core reading frame from two heterocycles (as with f-ImPyIm **2**) to four. Thus, compound **4** was expected to be more selective than **2** for the cognate sequence (5′-ACGCGT-3′). This hypothesis was verified by DNaseI footprinting which demonstrated no other footprints at non-cognate sites up to concentrations as high as 50 μM. However, the binding affinity of **4**, determined from SPR studies, was reduced by about 23-fold in comparison to its f-containing counterpart **2** ($1.9 \times 10^8\ M^{-1}$, SPR).[10] Thus, the gain in selectivity was at the expense of binding affinity.

The above studies have demonstrated the fine balance between high binding affinity and high selectivity. The present study aims to address this balance by combining the high selectivity of tetraamide monomers with the outstanding binding affinity of the H-pin motif. The target compound **5** (Fig. 1) consists of two PyImPyIm (**4**) monomers tethered via the central Py heterocycles with an ethylene glycol linker joining two pyrrole moieties in a diagonal arrangement (Fig. 2A). This structural design is different from the 'directly joined' H-pins, as in Figure 2A, reported previously by Lown's[5] and Dervan's[6] groups. The ethylene glycol linker will aid in the dissolution of the final molecule, as shown with H-pin **3e**,[9] and will allow each two polyamides to stack in the *overlapped motif* for recognition of the cognate sequence 5′-ACGCGT-3′. Accordingly, the synthesis of H-pins **3e** and **5** as well as their DNA binding properties, minus those previously reported for **3e**,[9] are reported herein. The cognate and non-cognate oligonucleotides utilized in the biophysical chemistry studies are given in Figure 3.

2. Results and discussion

2.1. Synthesis

H-pin **5** was formed in 28% yield by the reaction of the central core **6a**[10] and PyIm carboxylic acid **7** using a carbodiimide coupling procedure with EDCI (Scheme 1). PyIm-acid **7** was obtained by reacting pyrrole-2-carbonyl-chloride[10] with intermediate **8b** (formed by reduction of **8a** using Pd-C catalyzed hydrogenation) in a Schotten-Bauman coupling. The acid **7** was obtained after hydrolysis of intermediate **9**. H-pin **3e** was formed in 22% yield by following a similar strategy from the nitro-containing compound **10** with 4-formamido-1-methylimidazole-2-carboxylic acid **11**.[9,11]

CGCG-non-hairpin: 5'-C C A **C G C G** T G G-3'
3'-G G T **G C G C** A C C-5'

ACGCGT: 5'- G A A **C G C G** T C G C T
3'- C T T **G C G C** A G C T C

ACCGGT: 5'- G A A **C C G G** T C G C T
3'- C T T **G G C C** A G C T C

AAATTT: 5'- C G A **A A T T** T C C C T
3'- G C T **T T A A** A G G T C

Figure 3. The DNA sequences used in this study.

2.2. DNaseI footprinting

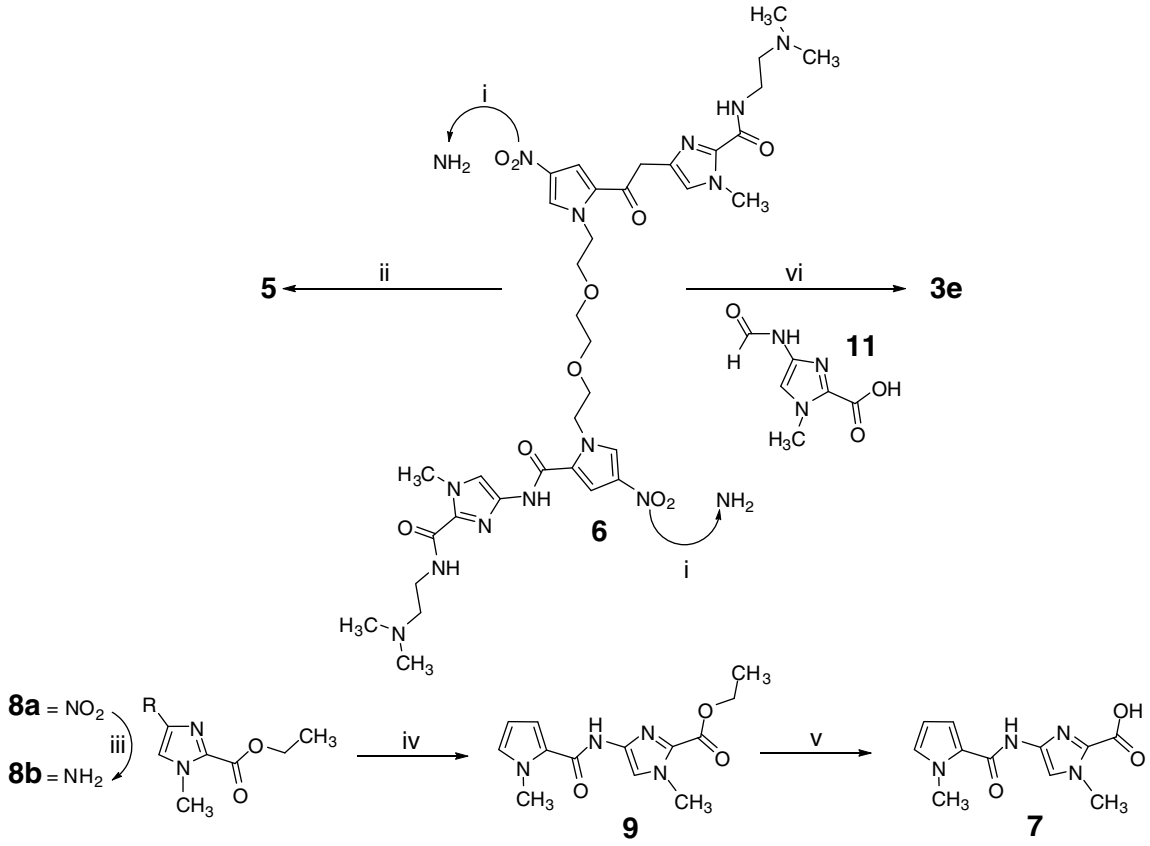
To investigate the sequence selectivity of monomer **4** and H-pin **5**, DNaseI footprinting experiments were performed using a 130 bp 5′-[^{32}P]-radiolabeled DNA fragment containing the following sequences 5′-ACGCGT-3′ (a); 5′-ACCGGT-3′ (b); 5′-ACGTGT-3′ (c); 5′-AGCGCT-3′ (d). The autoradiogram given in Figure 4 shows that a footprint appears at the 5′-ACGCGT-3′ cognate site from 0.5 μM for monomer **4** and 0.1 μM for H-pin **5**. The only other protection site observed with both compounds was at 5′-ACGTGT-3′. Here binding occurred from 5 μM, with H-pin **5** showing stronger binding than monomer **4** at this site. Some affinity for 5′-ACGTGT-3′ was expected in view of its high homology with the 5′-ACGCGT-3′. These results demonstrated that linkage of the two monomers did not affect selectivity and increased affinity for the cognate site.

2.3. Thermal denaturation

Thermal denaturation analyses were performed to support the selectivity of PyImPyIm H-pin **5** for its cognate DNA sequence (5′-ACGCGT-3′). The data shown in Table 1 indicate that **5** has incredible affinity for the cognate sequence with a ΔT_m of 16 °C. The two non-cognate sequences ACCGGT and AAATTT show no melt whatsoever, strongly corroborating the footprinting results. This ΔT_m value is less than that previously reported for H-pin **3e** which showed a value >30 °C,[9] an anticipated result due to the absence of the formamido groups. For comparison, the binding affinity of distamycin to its cognate AAATTT was found to be 14 °C.

2.4. Circular dichroism

Circular dichroism studies were carried out to provide evidence for the binding mode and to further probe selectivity for the target sequence. H-pin **5** was found to bind in the minor groove of its cognate sequence ACGCGT, as indicated by the strong induced CD band at 335 nm (Fig. 5A).[7,12] The appearance of an isodichroic point at ~315 nm provided evidence that the H-pin binds to the oligonucleotide via a single mechanism, presumably interacting in the minor groove as a covalently tethered and stacked dimer.



Scheme 1. Reagents and conditions (i) 10% Pd/C, cold MeOH, H_2, 55 PSI, 18 h, RT; (ii) **7**, EDCI, DMAP, dry DMF, 6 days, RT; (iii) 10% Pd/C, cold MeOH, H_2, 18 h, RT; (iv) pyrrole-2-carbonyl-chloride, dry DCM, dry TEA, 18 h, 0 °C-RT; (v) NaOH (2 M), H_2O, 24 h, reflux; (vi) **11**, EDCI, DMAP, dry DMF, 50 °C.

By contrast, titration of H-pin **5** to the non-cognate ACCGGT and AAATTT sequences did not produce an appreciable induced band (Fig. 5B and C). Consistent with the binding of f-ImPyIm (**2**),[7] PyIm-PyIm (**4**)[10] and H-pins (**3a–e**)[9] to ACGCGT, absence of a defined induced band or an isodichroic point often suggests that binding of the polyamide to these oligonucleotides is weak, and non-selective at best. It is worthy to note that an alternative binding mode begins to emerge on addition of **5** to the non-cognate sequences as evidenced by shifting of the isodichroic point to the right as the concentration of ligand is increased. Overall, the CD results corroborate both the footprinting and thermal denaturation results; that H-pin **5** binds selectively to the cognate sequence.

CD experiments were also conducted using H-pin **3e** and ACGCGT and a non-cognate ACCGGT with varying concentrations of added salt (12.5, 25, 50, 75, 100, 200, and 300 mM). In all cases salt concentration had virtually no effect on the signal intensity and shape of the induced band for the DNA/**3e** complex (data not shown), indicating insignificant changes in the conformation of the complex. This suggests that the electrostatic effects on the binding of compound **3e** and the DNA appear as secondary and that the binding is mainly driven by non-electrostatic factors similar to that observed for distamycin.[13]

2.5. Surface plasmon resonance

To obtain a more accurate measure of binding affinity, selectivity and to probe the stoichiometry of binding, SPR-biosensor experiments were performed. From the sensorgrams in Figure 6, it can be observed that H-pin **5** (Fig. 6C) retains excellent selectivity for the cognate sequence. No binding to either non-cognate AAATTT or ACCGGT sequence (Fig. 6A and B, respectively) was observed at concentrations up to 400 nM. The SPR binding profile of PyIm-PyIm H-pin is identical to that of the monomer PyImPyIm **4**,[10] supporting earlier evidence that the linker played a minor role in DNA sequence selectivity.[9] The linker was crucial, however, in enhancing the binding affinity. A binding constant (K) for PyImPyIm H-pin **5** for ACGCGT was determined to be $1.6 \times 10^8\ M^{-1}$ and the data best fit a 1:1 stoichiometry. Even though H-pin **5** bound with about 100-fold lower affinity than the f-ImPyIm H-pin **3e** to the same sequence ($K = 2.0 \times 10^{10}\ M^{-1}$), it represented a 25-fold greater binding constant ($K = 7.1 \times 10^6\ M^{-1}$) than the PyImPyIm monomer **4**. The results are summarized in Table 1. Judging from the sensorgram for the binding of PyImPyIm H-pin **5** to ACGCGT (Fig. 6) and H-pin **3e** with the same DNA sequence,[9] it is evident that the enhanced binding affinity of the H-pins over their monomer counterparts is largely driven by a slow dissociation rate of the ligand from DNA. Achieving this result is beneficial from a biological perspective because compounds with this DNA binding property are generally more effective in interfering the binding of proteins/transcriptional factors to promoter sites over compounds with rapid off-rates.[14]

2.6. Isothermal titration calorimetry

ITC experiments were performed at 25 °C to probe the thermodynamics of the H-pins binding to DNA. The H-pins again demonstrated binding to the cognate sequence as evidenced by the exothermic enthalpy of binding of H-pin **5** to ACGCGT (Fig. 7A); the same result observed with f-ImPyIm (**2**) and f-ImPyIm H-pin **3e** (Fig. 7B).[9] No heat change was observed to either ACCGGT or AAATTT. The ITC experiment was repeated at elevated temperatures (30 to 50 °C) in order to discern the heat capacity (ΔC_p) of the binding

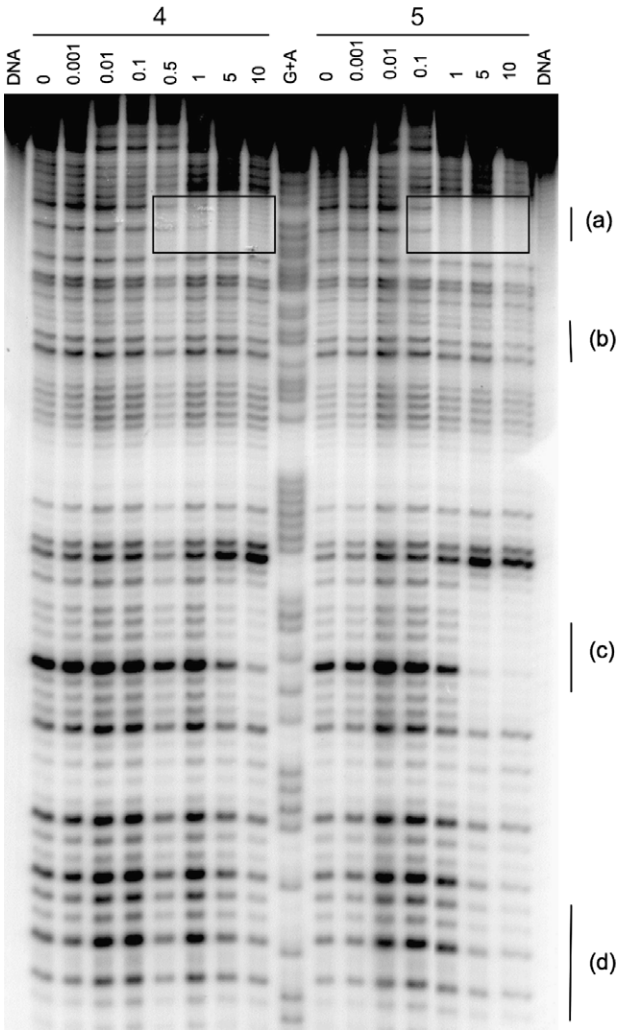


Figure 4. DNase I footprinting of monomer 4 and H-pin 5 using a 130 bp 5′-[^{32}P]-radiolabeled DNA fragment showing the sites 5′-ACGCGT-3′ (a), 5′-ACCGGT-3′ (b), 5′-ACGTGT-3′ (c), 5′-AGCGCT-3′ (d). DNA is undigested control and *G+A* is the sequencing lane. Thin boxes indicate the DNaseI cleavage protection for the cognate site.

Table 1
ΔT_m (°C), ΔH (kcal mol^{-1}) and SPR derived binding constants, K_{eq} (M^{-1}) for compounds **2**, **3e**, **4**, and **5** with the three DNA sequences[c]

		ACGCGT	ACCGGT	AAATTT
PyImPyIm H-pin (**5**)	ΔT_m (°C)	16.0 ± 0.3	0	0
	ΔH (kcal mol^{-1}) at 25 °C	−3.3	—	—
	K (M^{-1})[a]	1.6×10^8	nd[b]	nd
PyImPyIm (**4**) (ref.10)	ΔT_m	3.0	0.1	0.1
	ΔH	−3.2	—	—
	K	7.1×10^6	nd	nd
f-ImPyIm (**2**) (ref.7c)	ΔT_m	7.8	1.1	0.9
	ΔH	7.6	—	—
	K	1.9×10^8	2.2×10^5	5.3×10^4
f-ImPyIm H-pin (**3e**)	ΔT_m	> 30	14.1 ± 0.3	9.3 ± 0.3
	ΔH	−1.0	—	—
	K	2.0×10^{10}	1.1×10^8	5×10^5

[a] The uncertainty is estimated to be 10%.[7a]
[b] nd, not determined because no response units (RU) observed.
[c] The detection limit for SPR in our hands is about 1×10^2 M^{-1}.

reaction. From Figure 7C, ΔC_p for PyImPyIm H-pin **5** was determined from the slope of the line to be −116 cal mol^{-1} K^{-1}, compared to that observed for PyImPyIm **4** and f-ImPyIm H-pin **3e** (−116 and −499 cal mol^{-1} K^{-1}, respectively).

Using the binding constant of H-pin **5** for ACGCGT of 1.6×10^8 M^{-1} (determined by SPR at 25 °C), and the ΔH value of −3.3 kcal mol^{-1} from ITC, the ΔG and $T\Delta S$ at 298 K were calculated to be −11.2 and 7.9 kcal mol^{-1}, respectively (Table 1). This finding is consistent with the binding of f-ImPyIm H-Pin **3e** to ACGCGT in which the binding event is also mainly driven by entropy ($T\Delta S$ = 13.0 kcal mol^{-1}, ΔH = −1.0 kcal mol^{-1}). The results could also be explained by considering the attribution and additivity of binding energies of the interaction of small molecules to large biomolecules.[15] In this instance, A–B is a ligand containing two identical units of A and B covalently linked together. The relationship between binding free energy of A–B with DNA, ΔG_{A-B} would therefore be the sum of ΔG_A, ΔG_B, and ΔG_C, in which ΔG_C is the 'cooperative free energy.' For the linked A–B, the free energy gain should be favorable, since the lower entropic cost is less (comparing with unlinked ligands).

2.7. Molecular modeling

Based on the aforementioned results on the binding of PyImPyIm H-pin 5 in the minor groove of 5′-ACGCGT-3′, in a similar manner as that reported for the stacked dimer formed from unlinked PyImPyIm 4,[10] a representation of the complex is depicted in Figure 8A. A computer generated model of the complex of PyImPyIm H-pin 5 bound to a self-complementary duplex (5′-dCCACGCGTGG-3′)$_2$ was subsequently produced and shown in Figure 8B. The complex was generated using SYBYL and the model was structurally optimized. The model provides several insights on the complex formation. In common with the DNA binding properties of f-ImPyIm H-pin 3e,[9] the cross-linked and stacked polyamide 5 fits snugly within the minor groove and forms specific contacts with the floor and wall of the groove. The ethylene glycol linker provides sufficient length in order for the polyamides to stack in a staggered manner. Finally, the oxygen atoms in the linker point out from the complex to maximize interactions with water molecules, thereby improving solubility and binding affinity.

With the models of H-pin **5** and **3e** bound to 5′-ACGCGT-3′, the heat capacity (ΔC_p) arising from changes in hydrophobicity could be estimated using solvent accessible surface area (SASA) calculations as previously reported.[10] The calculated values of ΔC_p for H-pins **5** and **3e** were determined using Eqs. (2)–(4). For H-pin **5**, the results obtained from these calculations were −481 ± 85, −649 ± 47, −595 ± 78 cal mol^{-1}K^{-1}. Similarly, for H-pin **3e**, the values were −369 ± 82, −480 ± 47, −470 ± 85 cal mol^{-1} K^{-1}, respectively. The calculated values are generally close to the experimental values providing support for the structure and conformation of the complexes of PIPI H-pin **5** and f-IPI H-pin **3e** and ACGCGT. The SASA calculated data of the H-pins with CCACGCGTGG are given in Figure 8C and D, respectively.

3. Conclusions

A successful strategy for systematically developing polyamides that exhibit an excellent balance between sequence selectivity and binding affinity is reported herein. This strategy utilizes the H-pin molecular design that combines the high selectivity of tetraamide monomers with the strong binding affinity of the H-pin motif. As a result, we have produced molecules, e.g., H-pins **3e** and **5**, which showed excellent affinity and sequence specificity. The development of the next generation of H-pins is currently underway and the goal is to produce molecules with even higher affinity while simultaneously addressing cellular uptake/nuclear localization is-

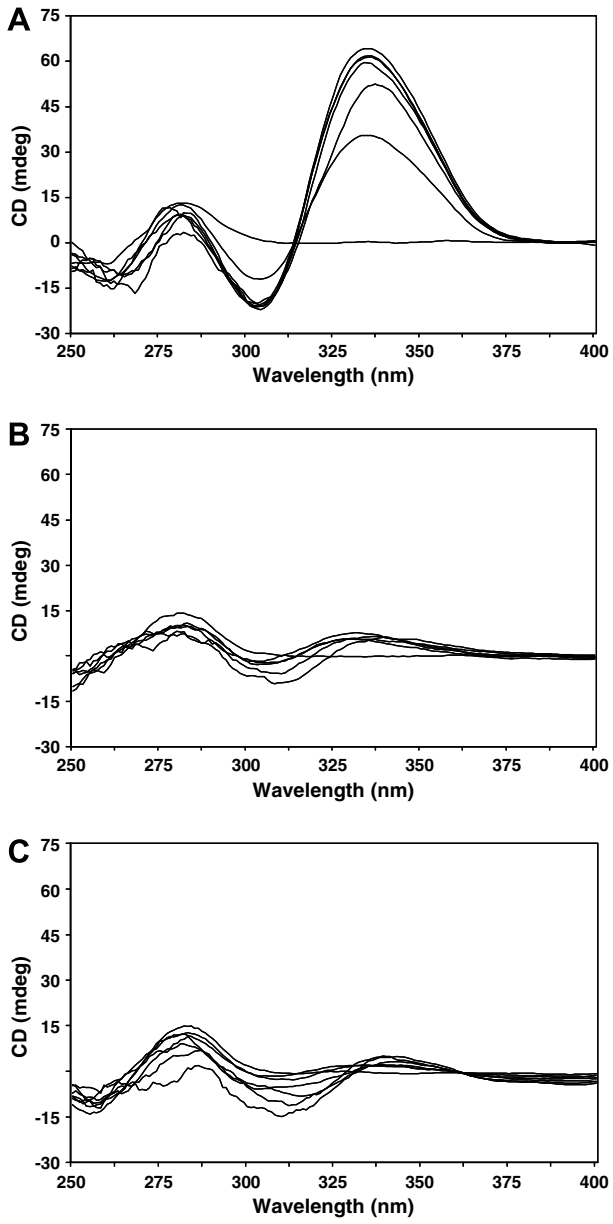


Figure 5. CD spectra of H-pin **5** with (A) cognate sequence ACGCGT and non-cognate sequences ACCGGT (B) and AAATTT (C).

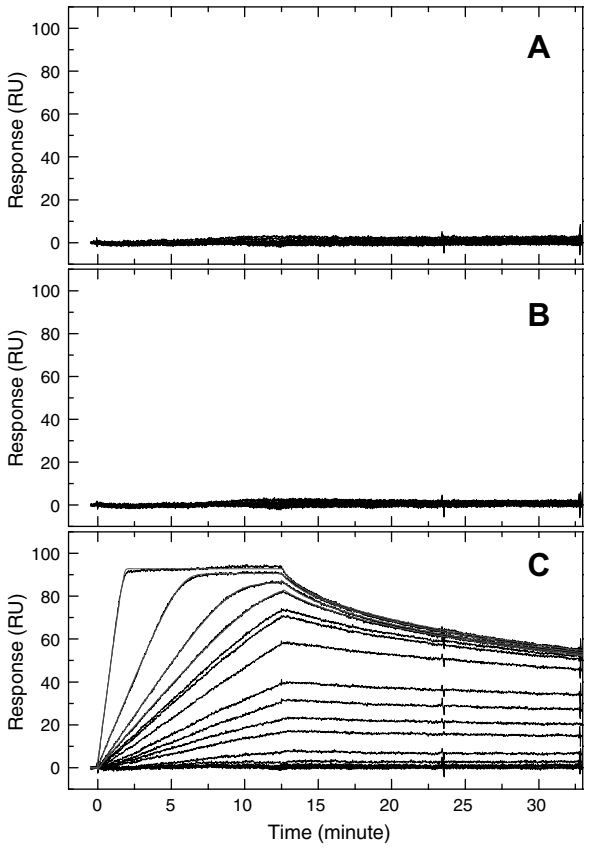


Figure 6. SPR thermograms of H-pin **5** with non-cognate sequences AAATTT (A) and ACCGGT (B) and cognate sequence ACGCGT (C). The ligand concentrations from the lowest to highest response are 0.1, 0.2, 0.4, 0.6, 0.8, 1, 2, 4, 6, 8, 10, 20, 40, 60, 80, 100, 200, and 400 nM. All sensorgrams were plotted in the same scale and concentration range for comparison. The binding constant with ACGCGT ($1.6 \times 10^8\ M^{-1}$) was obtained from kinetic fits of the 4 highest concentration sensorgrams (four fit curves are overlaid with the top four sensorgrams).

sues, which are key stumbling blocks in the development of polyamides as potential therapeutic agents.[16]

4. Experimental

4.1. 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 Midiac FT-IR instrument as films on NaCl discs. ^{1}H-NMR spectra were obtained using a Varian Unity Inova 400 and 500 instruments. 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.

4.1.1. Synthesis of 9

N-Methylpyrrole-2-carbonyl chloride[10,11] (2 mmol) was dissolved in DCM (anhydrous, 2 mL) and added drop-wise to a solution of **8**[17] (1 mmol) in DCM (2 mL) and TEA (0.28 mL, 2 mmol) and cooled to 0 °C with stirring. The reaction was allowed to warm to RT overnight. Reaction mixture transferred to sep. funnel and water (~15 mL) added. The aqueous layer was then basified with NaOH, and the organic layer collected. Aqueous layer extracted with DCM (2 × 25 mL) and organic layers collected, dried (Na_2SO_4) and solvent removed in vacuo. The residue was then purified by column chromatography (silica, gradient, 100:0–96:4 % $CHCl_3$/MeOH) to yield **9** as an off-white solid (74 mg, 24 %), mp. 182.8–186.2 °C: R_f 0.35 (99:1 % v/v $CHCl_3$/MeOH); IR ν (neat) 730, 754, 1025, 1058, 1116, 1176, 1245, 1275, 1320, 1375, 1411, 1471, 1545, 1649, 1704, 2362, 2927 cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ 8.19 (s, 1H), 7.56 (s, 1H), 6.79 (t, J = 2 Hz, 1H), 6.66 (dd, J = 1.6, 4 Hz, 1H), 6.15 (dd, J = 2.4, 2 Hz, 1H), 4.44 (q, J = 7.1 Hz, 2H), 4.02 (s, 3H), 3.99 (s, 3H), 1.45 (t, J = 7.2 Hz, 3H) ppm; MS (ES^+) m/z (rel. intensity) 277 ([M+H], 100%), 299 ([M+Na], 55%).

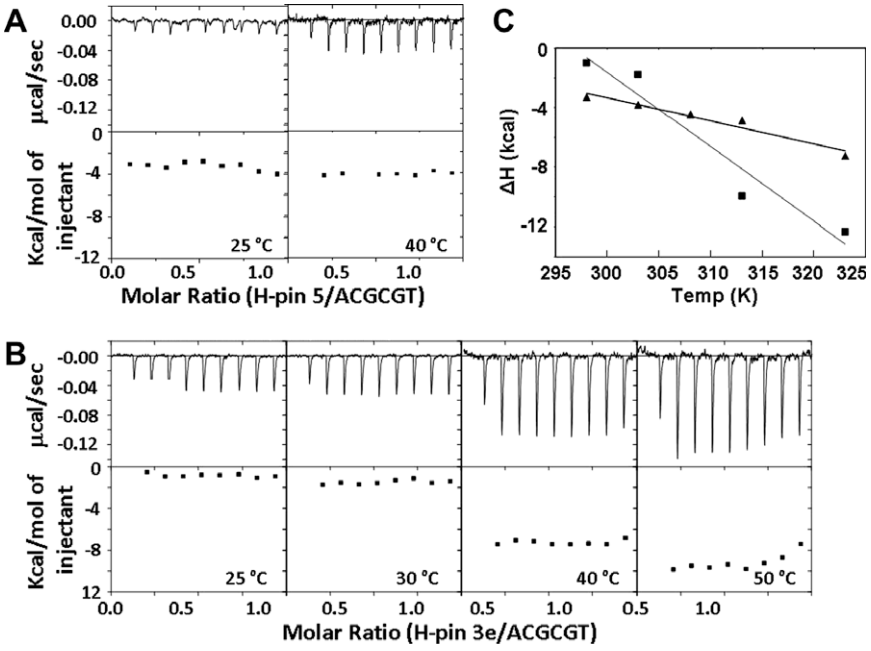


Figure 7. (A) Thermogram on the binding of H-pin **5** with ACGCGT at 25 and 40 °C. (B) ITC studies on the binding of H-pin **3e** to ACGCGT at 25, 30, 40, and 50 °C. (C) Plots of ΔH on the binding of **5** (▲) and **3e** (■) to ACGCGT over a temperature range for determining values of ΔC_p.

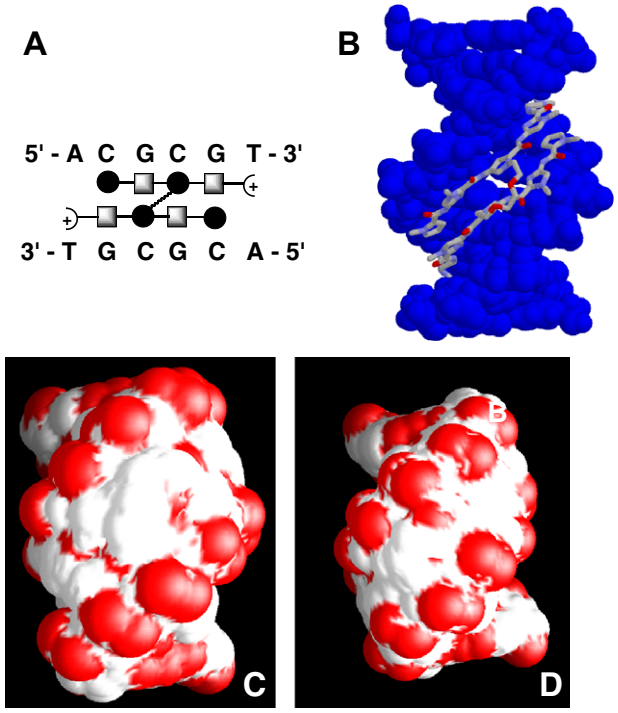


Figure 8. (A) A model of the complex of H-pin **5** with 5′-ACGCGT-3′. (B) Computer generated depiction of the complex of H-pin **5** with 5′-ACGCGT-3′ using SYBYL following a 1000-step structural optimization. (C) Solvent accessible surface area of PyImPyIm H-pin **5** bound to $(dCCACGCGTGG)_2$. Red is polar and white is non-polar. (D) SASA of f-ImPyIm H-pin **3e** bound to $(dCCACGCGTGG)_2$.

4.1.2. Synthesis of 7

PyIm-ethyl ester (**9**, 74 mg, 0.27 mmol) was dissolved in MeOH (3 mL). NaOH (0.16 mL, 0.32 mmol) and water (3 mL) were then added and the reaction mixture heated under reflux with stirring for 24 h. The cooled solution was then washed with $CHCl_3$ (20 mL) and the aqueous layer acidified with HCl_{aq} (6 M) until the product precipitated. The suspension was filtered to yield compound **7** as an off-white solid (59 mg, 88 %), mp. 141.2–141.8 °C: R_f 0.17 (50:10:1, $CHCl_3$:MeOH:NH_4OH); IR ν (neat) 745, 847, 926, 964, 1038, 1064, 1121, 1190, 1367, 1440, 1456, 1646, 1680, 2888, 2960, 3308 cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ 7.46 (s, 1H), 6.90 (m, 1H), 6.71 (s, 1H), 6.04–6.02 (m, 1H), 3.95 (s, 3H), 3.84 (s, 3H) ppm; MS (ES^+) *m/z* (rel. intensity) 249 ([M+H], 100%), 271 ([M+Na], 18%); HRMS calcd for $C_{11}H_{12}N_4O_3$, 249.0987; obsd, 249.0982.

4.1.3. Synthesis of PyImPyIm H-Pin 5

Glycol linked intermediate **6** (79 mg, 0.095 mmol) was dissolved in MeOH (75 mL) and reduced for ~18 h using 10% Pd/C in the presence of H_2. PyIm-acid **7** (59 mg, 0.24 mmol, 2.5 equiv) EDCI (55 mg, 0.29 mmol), DMAP (5.8 mg, 0.05 mmol) and dry DMF (2 mL) were reacted with the resulting amine, stirring at RT for 6 days, protected from light in an Ar atmosphere. The DMF was removed via Kugelröhr distillation (0.005 atm.) and the residue purified via column chromatography (silica, gradient, 0:100–100:0 % v/v $CHCl_3$/MeOH then 100:10:1–70:10:1, $CHCl_3$:MeOH:NH_4OH) to yield H-pin **5** as a pale brown solid (32 mg, 28 %), mp. 127 °C: R_f 0.62 ($CHCl_3$:MeOH:NH_4OH, 50:10:1); IR ν (neat) 1685, 1536, 1468, 1410, 1248, 1118, 752 cm^{-1}; ^1H NMR (400 MHz, DMSO) 10.40 (br s, 2H), 10.39 (br s, 2H), 9.83 (br s, 2H), 7.74 (br t, J = 5.6 Hz, 2H), 7.55 (s, 2H), 7.50 (s, 2H), 7.47 (s, 2H), 7.14 (s, 2H), 7.12 (m, 2H), 6.99 (s, 2H), 6.05 (t, J = 3.2 Hz, 2H), 4.46 (t, J = 3.2 Hz, 4H), 3.97 (s, 6H), 3.93 (s, 6H), 3.89 (s, 6H) 3.67–3.66 (m, 4H), 3.48–3.47 (m, 4H), 2.40 (s, 4H), 2.18 (s, 12H) ppm; MS (ES^+) *m/z* (rel. intensity) 1214 ([M+H], 40%), 607 (100%), 405 (60%); HRMS calcd for $C_{56}H_{73}N_{22}O_{10}$, 1213.5880; obsd, 1213.5837.

4.1.4. Synthesis of f-ImPyIm H-pin 3e

The procedure was similar to that used for the synthesis of H-Pin **5**, except nitro-starting material **6** (63 mg, 0.077 mmol) and f-Im-acid **11** [11] (28 mg, 0.16 mmol) were used. The product was isolated as a grey/brown solid (19 mg, 22%), mp. 142–145 °C: R_f 0.37 (79:20:1 v/v $CHCl_3$/MeOH/NH_4OH); IR ν (neat) 2924, 2862, 1655, 1535, 1463, 1437, 1380, 1256 cm^{-1}; ^1H NMR

(DMSO-d_6) δ 1.69 (t, J = 6.5 Hz, 6H), 2.25 (s, 12H), 3.03 (q, J = 6.5 Hz, 4H), 3.47 (s, 4H), 3.65 (t, J = 5.2 Hz, 4H), 3.93 (s, 6H), 3.94 (s, 6H), 4.43 (t, J = 6.9 Hz, 4H), 7.17 (s, 2H), 7.43 (s, 2H), 7.48 (s, 2H), 7.50 (s, 2H), 8.21 (s, 2H), 10.04 (s, 2H), 10.32 (s, 2H), 10.32 (s, 2H); MS (ES^+) m/z (rel. intensity) 1056 ([M+H], 50%), 529 ([M+2H]$^{2+}$, 20%).

4.2. DNaseI footprinting

4.2.1. Preparation of the DNA substrate, radiolabeling and purification

A 130 bp fragment was amplified by using PCR as fellow. The forward primer 5′-GTCGTTAGGAGAGCTCACTTG-3′ (4 ng) was radioactively labeled by treatment with γ-[^{32}P] (3 μL) and 1 μL T4 polynucleotide kinase (Invitrogen) following the standard protocols. PCR was performed in thermophilic DNA polybuffer containing dNTPs (50 μL, 125 μM), $MgCl_2$ (1 mM), Flexi Taq polymerase (1 U) and ^{32}P-labeled forward primer, reverse primer 5′-CTCCAGAAAGCCGGCAACTCAG-3′ and the templates 5′-ATGCTC CAGAAAGCCGGCACTCAGTCTACAAACGCGTCATCTTGATCACCGGTG TTCACAGAAATTTCTCTAGATCTACACGTAACTCTAGTAGCGCTCTTCA AGCAAGTGGAGCTCTCCTAACCGACTTT-3′ (20 ng) and 5′-AAAGTCG GTTAGGAGAGCTCCACTTGCTTGAAGAGCGCTACTAGAGTTACGTGTA GATCTAGAGAAATTTCTGTGAACACCGGTGATCAAGATGACGCGTTT GTAGACTGAGTGCCGGCTTTCTGGAGCAT-3′ (20 ng). Polymerase chain reaction was carried out as follows: an initial denaturation step for 3 min at 95 °C and [1 min at 94 °C, 1 min at 63 °C, and 1 min at 72 °C] for 35 cycles. The PCR products were purified by 2% agarose gel electrophoresis. Finally, DNA was isolated by using the Mermaid Kit (Q-biogene) according the manufacturer's instructions.

4.2.2. DNase I footprinting experiments

DNase I digestions were conducted in a total volume of 8 μL. The labeled DNA fragment (200 μL, 200 count s^{-1}) was incubated 30 min in 4 μL TN binding buffer (10 mM Tris Base, 10 mM NaCl, pH 7) containing the desired drug concentration. Cleavage by DNase I was initiated by addition of 2 μL DNase I solution (2 μL, 20 mM NaCl, 2 mM $MgCl_2$, 2 mM $MnCl_2$, DNaseI 0.1 U mL^{-1}, pH 8) and was stopped after 3 min by cooling the samples on dry ice. The samples were then lyophilized dry. DNA was resuspended in 4 μL formamide loading dye (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% cyanol blue) and denatured by heating the sample to 90 °C and cooled on ice prior to loading onto a conventional denaturing polyacrylamide (10%) gel containing urea (7.5 M). Electrophoresis was performed for 2.5 hours (70 W, 50 °C) in TBE Buffer (89 mM Tris base, 89 mM boric acid, 2.5 mM Na_2EDTA, pH 8.3). The gel was transferred onto Whatman 3MM paper and dried under vacuum at 80 °C for 2 h. The gel was exposed overnight to X-Ray film (Super RX, Fuji).

4.3. Thermal denaturation

The synthetic DNA hairpins used in these studies were obtained from Operon (Huntsville, AL): **ACGCGT**, 5′-GAACGCGTCGCTCTCGA CGCGTTC; **ACCGGT**, 5′-GAACCGGTCGCTCTCGACCGGTTC; **AAATTT**, 5′-CGAAATTTCCCTCTGGAAA-TTTCG. Data was obtained using a Cary Bio 100 spectrophotometer and cells with a 10 mm pathlength. Experiments were performed in $PO_4$0 (10 mM sodium phosphate, 1 mM EDTA, pH 6.2) with 1 μM oligonucleotide and 3 μM ligand. Oligonucleotide samples were reannealed prior to denaturation studies by heating at 70 °C for 1 min then allowing cooling to RT. The temperature was programmed to ramp from 25 to 95 °C at a rate of 0.5 °C/min recording the Abs_{260} every 0.5 °C. The data was analyzed using KaleidaGraph (Synergy Software, Reading PA) and the T_m values determined as the maximum of the first derivative.

4.4. Circular dichroism

CD studies were performed on an OLIS DSM20 spectrophotometer using the three oligonucleotides detailed above. Experiments were conducted at ambient temperature in a 10 mm pathlength cuvette, using a 2.4 nm bandpass. A 9 μM solution of DNA in $PO_4$5 (10 mM phosphate, 50 mM Na^+, 1 mM EDTA, pH 6.2) was titrated with a solution of H-pin **5** (0.5 mM in water), in aliquots of 1 molar equivalent, past the point of saturation. CD response was recorded every 1 nm over a wavelength range of 250 to 400 nm.

4.5. Surface plasmon resonance

SPR-biosensor experiments were conducted in degassed MES buffer (200 mM Na^+, 10 mM [2-(*N*-morpholino)ethanesulfonic acid], 1 mM EDTA, 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 Reagents (Midland, TX) with HPLC purification with the following sequences: **5′-biotin-ACGCGT**, 5′-biotin-GAACGCGTCCTCTGACGC-GTTC; **5′-biotin-AAATTT**, 5′-biotin-CGAAATTTCCTCTGAAATTTCG; **5′-biotin-ACCGGT**, 5′-biotin-GAACCGGTCCTCTGACCGGTTC. The experiments were conducted using a BIACORE 2000 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 a flow rate of 1 μL min^{-1} until the response units reached ~400 RUs. Flow cell 1 was left blank, while flow cells 2, 3, and 4 were immobilized with the three different DNA hairpins. Typically, a series of different concentrations of ligand was injected onto the chip with a flow rate of 20 μL min^{-1} for a period of 12.5 min, followed by a dissociation period of 20 min. At the end of every cycle, the chip surface was regenerated with a 10 μL injection of 500 mM NaCl/25 mM NaOH aqueous solution, injection tube rinsing, and multiple 1 min buffer injections. For H-pin **5** and ACGCGT, the global kinetic fit (1:1) was conducted on four sensorgrams of highest concentrations (80, 100, 200, 400 nM), and the binding constant was reported as the ratio of k_a/k_d.

4.6. Isothermal titration microcalorimetry

ITC experiments were carried out on a MicroCal VP-ITC (Northampton, MA) at 25, 30, and 35 °C. DNA (2 μM in $PO_4$5) was titrated with H-pin **5** (0.5 mM in $PO_4$5) in 3 μL aliquots every 300 s (50 injections). Data was processed with MicroCal Origin 7.0 as previously described[7d,18] using a one-site model to fit the curve. A linear fit was then employed and this was subtracted from the reaction integrations to normalize for non-specific heat components. ΔG was calculated from Eq. 1 and using the binding constant of 1.6×10^8 M^{-1} that was obtained from SPR studies.

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

Where R is 1.987 cal mol^{-1} K^{-1} and T is measured in K.

4.7. Molecular modeling

The 3 D model depiction of f-ImPyIm-EG-8 (**3e**) /5′-CCAC GCGTGG-3′ complex was generated using SYBYL 7.0 on a Silicon Graphics workstation and reported previously.[9,19] The model for H-pin **5** with the same oligonucleotide was generated in the same manner. B-form double helical DNA was generated in SYBYL, and the complex was constructed using the published unlinked PyIm-PyIm (**4**)/CGCG complex[10] by building the ethylene glycol linker joining the two polyamide moieties through the central pyrrole group at N1. The complex in water was then structurally optimized in a 1000-step minimization.

4.8. Solvent accessible surface areas (SASA)

The SASA calculation procedures have been previously described.[7c,20] In brief, ions were removed and carbon, carbon-bound hydrogen, and phosphorus atoms were assigned as non-polar and the rest polar. The PyImPyIm/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[21] using a probe radius of 1.7683 Å and Cornell et al. radii.[22]

The heat capacity change (cal mol^{-1} K^{-1}) arisen from the polar/non-polar area change was calculated with three different models Eq. 2,[23] Eq. 3,[24] Eq. 4.[25]

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

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

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

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

Acknowledgments

Support from the National Science Foundation (CHE-0550992), the National Institutes of Health, (GM61587 to WDW), Glaxosmithkline summer undergraduate research fellowship (J.P.L.), Cancer Research UK (C2259/A3083), and Hope College is gratefully acknowledged.

References and notes

1. (a) Melander, C.; Burnett, R.; Gottesfeld, J. M. *J. Biotechnol.* **2004**, *112*, 195; (b) Baraldi, P. G.; Bovero, A.; Fruttarolo, F.; Preti, D.; Tabrizi, M. A.; Pavani, M. G.; Romagnoli, R. *Med. Res. Rev.* **2004**, *24*, 475; (c) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284; (d) Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 291; (e) Reddy, B. S.; Sharma, S. K.; Lown, J. W. *Curr. Med. Chem.* **2001**, *8*, 475.
2. Dervan, P. B.; Bürli, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 688.
3. (a) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393; (b) Chen, X.; Ramakrishana, B.; Rao, S. T.; Sundaralingam, M. *Nat. Struct. Biol.* **1994**, *1*, 169.
4. (a) Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215; (b) Dervan, P. B.; Doss, M. R.; Marques, M. A. *Curr. Med. Chem. Anti-Cancer Agents* **2005**, *5*, 373; (c) Walker, W. L.; Kopka, M. L.; Goodsell, D. S. *Biopolymers* **1997**, *44*, 323; (d) Jones, G.; Palumbo, B. M. *Advances in DNA Sequence-Specific Agents*; JAI Press Inc.: Greenwich, 1998.
5. (a) Al-Said, N.; Lown, J. W. *Synth. Commun.* **1995**, *25*, 1059; (b) Chen, Y-H.; Yang, Y.; Lown, J. W. *J. Biomol. Struct. Dyn.* **1996**, *14*, 341; (c) Chen, Y-H.; Liu, J-X.; Lown, J. W. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2223; (d) O'Hare, C. C.; Mack, D.; Tandon, M. S.; Sharma, K.; Lown, J. W.; Kopka, M. L.; Dickerson, R. E.; Hartley, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 72.
6. (a) Dwyer, T. J.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1993**, *115*, 9900; (b) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1993**, *115*, 9892; (c) Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 3663; (d) Olenyuk, B.; Jitianu, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 4741.
7. (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) Buchmueller, K. L.; Bailey, S. L.; Matthews, D. A.; Taherbhai, Z. T.; Register, J. K.; Davis, Z. S.; Bruce, C. D.; O'Hare, C. C.; Hartley, J. A.; Lee, M. *Biochemistry* **2006**, *45*, 13551.
8. (a) Verma, R.; Patapoutian, A.; Gordon, C. B.; Campbell, J. L. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 7155; (b) Wu, X.; Lee, H. *Oncogene* **2002**, *21*, 7786; (c) Yamada, M.; Sato, N.; Taniyama, C.; Ohtani, K.; Arai, K.; Masai, H. *J. Biol. Chem.* **2002**, *277*, 27668; (d) Hess, G. F.; Drong, R. F.; Weiland, K. L.; Slightom, J. L.; Sclafani, R. A.; Hollingsworth, R. E. *Gene* **1998**, *28*, 133.
9. O'Hare, C. C.; Uthe, P. B.; Mackay, H.; Blackmon, K. N.; Jones, J.; Brown, T.; Nguyen, B.; Wilson, W. D.; Lee, M.; Hartley, J. A. *Biochemistry* **2007**, *46*, 11661.
10. Brown, T.; Mackay, H.; Turlington, M.; Sutterfield, A.; Smith, T.; Sielaff, A.; Westrate, L.; Bruce, C.; Kluza, J.; O'Hare, C.; Hartley, J. A.; Nguyen, B.; Wilson, W. D.; Lee, M. *Bioorg. Med. Chem.* **2008**, *16*, 5266.
11. Mulder, K.; Brown, T.; Nolan, P.; Smith, T.; Lee, M. *Heterocycl. Commun.* **2008**, *14*, 137.
12. (a) Lyng, R.; Rodger, A.; Norden, B. *Biopolymers* **1992**, *32*, 1201; (b) Lyng, R.; Rodger, A.; Norden, B. *Biopolymers* **1991**, *31*, 1709.
13. Lah, J.; Vesnaver, G. *J. Mol. Biol.* **2004**, *3*, 73.
14. (a) Paal, K.; Baeuerle, P. A.; Schmitz, M. L. *Nucleic Acids Res.* **1997**, *25*, 1050; (b) Ptashne, M.; Gann, A. *Genes and Signals*, first ed.; Cold Spring Harbor Laboratory: NY, 2001.
15. Jencks, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4046.
16. (a) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *ACS Chem. Biol.* **2007**, *2*, 561; (b) Nickols, N. G.; Jacobs, C. S.; Farkas, M. E.; Dervan, P. B. *Nucleic Acids Res.* **2007**, *35*, 363; (c) Edelson, B. S.; Best, T. P.; Olenyuk, B.; Nickols, N. G.; Doss, R. M.; Foister, S.; Heckel, A.; Dervan, P. B. *Nucleic Acids Res.* **2004**, *20*, 2802; (d) Best, T. P.; Edelson, B. S.; Nickols, N. G.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12063; (e) Belitsky, J. M.; Leslie, S. J.; Arora, P. S.; Beerman, T. A.; Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 3313; (f) Anthony, N. G.; Fox, K. R.; Johnston, B. F.; Khalaf, A. I.; Mackay, S. P.; McGroarty, I. S.; Parkinson, J. A.; Skellern, G. G.; Suckling, C. J.; Waigh, R. D. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1353; (g) Anthony, N. G.; Johnston, B. F.; Khalaf, A. I.; Mackay, S. P.; Parkinson, J. A.; Suckling, C. J.; Waigh, R. D. *J. Am. Chem. Soc.* **2004**, *15*, 11338.
17. (a) Krowicki, K.; Lown, J. W. *J. Org. Chem.* **1987**, *52*, 3493; (b) Wurtz, N. R.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Org. Lett.* **2001**, *3*, 1201; (c) Nishiwaki, E.; Tanaka, S.; Lee, H.; Shibuya, M. *Heterocycles* **1988**, *27*, 1945.
18. Indyk, L.; Fisher, H. F. *Methods Enzymol.* **1998**, *295*, 350.
19. Isaacs, R. J.; Rayens, W. S.; Spielmann, H. P. *J. Mol. Biol.* **2002**, *319*, 191.
20. Nguyen, B.; Stanek, J.; Wilson, W. D. *Biophys. J.* **2006**, *90*, 1319.
21. Nicholls, A.; Sharp, K. A.; Honig, B. *Proteins* **1991**, *11*, 281.
22. 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.
23. (a) Spolar, R. S.; Ha, J. H.; Record, M. T., Jr. *Proc. Natl. Acad. Sci. U.S.A.* **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.
24. (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.
25. Ren, J.; Jenkins, T. C.; Chaires, J. B. *Biochemistry* **2000**, *39*, 8439.