Physical and Structural Basis for the Strong Interactions of the -ImPy- Central Pairing Motif in the Polyamide f-ImPyIm†

Karen L. Buchmueller,‡,§ Suzanna L. Bailey,‡ David A. Matthews,‡ Zarmeen T. Taherbhai,‡ Janna K. Register,∥ Zachary S. Davis,∥ Chrystal D. Bruce,∥ Caroline O'Hare,⊥ John A. Hartley,⊥ and Moses Lee*,‡,@

Department of Chemistry, Furman University, Greenville, South Carolina 29613, Department of Chemistry, Erskine College, Due West, South Carolina 29639, and Cancer Research UK Drug-DNA Interactions Research Group, Department of Oncology, University College London, London W1W 7BS, U.K.

Received June 22, 2006; Revised Manuscript Received September 13, 2006

ABSTRACT: The polyamide f-ImPyIm has a higher affinity for its cognate DNA than either the parent analogue, distamycin A (10-fold), or the structural isomer, f-PyImIm (250-fold), has for its respective cognate DNA sequence. These findings have led to the formulation of a two-letter polyamide "language" in which the -ImPy- central pairings associate more strongly with Watson−Crick DNA than -PyPy-, -PyIm-, and -ImIm-. Herein, we further characterize f-ImPyIm and f-PyImIm, and we report thermodynamic and structural differences between -ImPy- (f-ImPyIm) and -PyIm- (f-PyImIm) central pairings. DNase I footprinting studies confirmed that f-ImPyIm is a stronger binder than distamycin A and f-PyImIm and that f-ImPyIm preferentially binds CGCG over multiple competing sequences. The difference in the binding of f-ImPyIm and f-PyImIm to their cognate sequences was supported by the Na^+-dependent nature of DNA melting studies, in which significantly higher Na^+ concentrations were needed to match the ability of f-ImPyIm to stabilize CGCG with that of f-PyImIm stabilizing CCGG. The selectivity of f-ImPyIm beyond the four-base CGCG recognition site was tested by circular dichroism and isothermal titration microcalorimetry, which shows that f-ImPyIm has marginal selectivity for (A·T)CGCG(A·T) over (G·C)-CGCG(G·C). In addition, changes adjacent to this 6 bp binding site do not affect f-ImPyIm affinity. Calorimetric studies revealed that binding of f-ImPyIm, f-PyImIm, and distamycin A to their respective hairpin cognate sequences is exothermic; however, changes in enthalpy, entropy, and heat capacity (ΔC_p) contribute differently to formation of the 2:1 complexes for each triamide. Experimental and theoretical determinations of ΔC_p for binding of f-ImPyIm to CGCG were in good agreement (−142 and −177 cal mol^{-1} K^{-1}, respectively). ^1H NMR of f-ImPyIm and f-PyImIm complexed with their respective cognate DNAs confirmed positively cooperative formation of distinct 2:1 complexes. The NMR results also showed that these triamides bind in the DNA minor groove and that the oligonucleotide retains the B-form conformation. Using minimal distance restraints from the NMR experiments, molecular modeling and dynamics were used to illustrate the structural complementarity between f-ImPyIm and CGCG. Collectively, the NMR and ITC experiments show that formation of the 2:1 f-ImPyIm−CGCG complex achieves a structure more ordered and more thermodynamically favored than the structure of the 2:1 f-PyImIm−CCGG complex.

Pyrrole and imidazole-containing polyamides bind the minor groove of DNA and selectively target specific DNA sequences (*1*−*3*). Therefore, polyamides are being developed as therapeutics (*4*−*7*) and biosensors (*8*) for medicinal purposes. In addition, the interactions between polyamides and DNA provide a better understanding of the structure and function of DNA (*5*, *9*−*12*).

It is the composition of these polyamides, particularly the heterocyclic moieties, that can be programmed to recognize specific DNA sequences. For example, the pyrrole (Py) moieties in distamycin A (Figure 1A) stack with the pyrrole moieties in a second distamycin A molecule to form an antiparallel dimer in the DNA minor groove. These Py·Py pairs target A·T and T·A base pairs (*13*, *14*). In addition, imidazole (Im) can be paired opposite pyrrole (Im·Py) and selectively target a G·C base pair (*15*−*20*). Notably, recognition of A·T versus T·A base pairs can be achieved with 3-hydroxy-1*H*-pyrrole (Hp) and related moieties (*21*−*23*), but this coincides with a small loss of DNA affinity (*21*−*24*). Thus, to simplify this study, we have focused on pyrrole- and imidazole-containing polyamides. In summary, the pairing rules for pyrrole- and imidazole-containing polyamides are that A·T and T·A base pairs are recognized by Py·Py pairs, G·C base pairs by Im·Py pairs, and C·G base

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

* To whom correspondence should be addressed. Telephone: (616) 395-7190. Fax: (616) 395-7923. E-mail: lee@hope.edu.

‡ Furman University.

§ Current address: Department of Chemistry, Wake Forest University, Winston-Salem, NC 27109.

∥ Erskine College.

⊥ University College London.

@ Current address: Natural Sciences, Hope College, Holland, MI 49423.

10.1021/bi061245c CCC: $33.50
Published on Web 10/24/2006

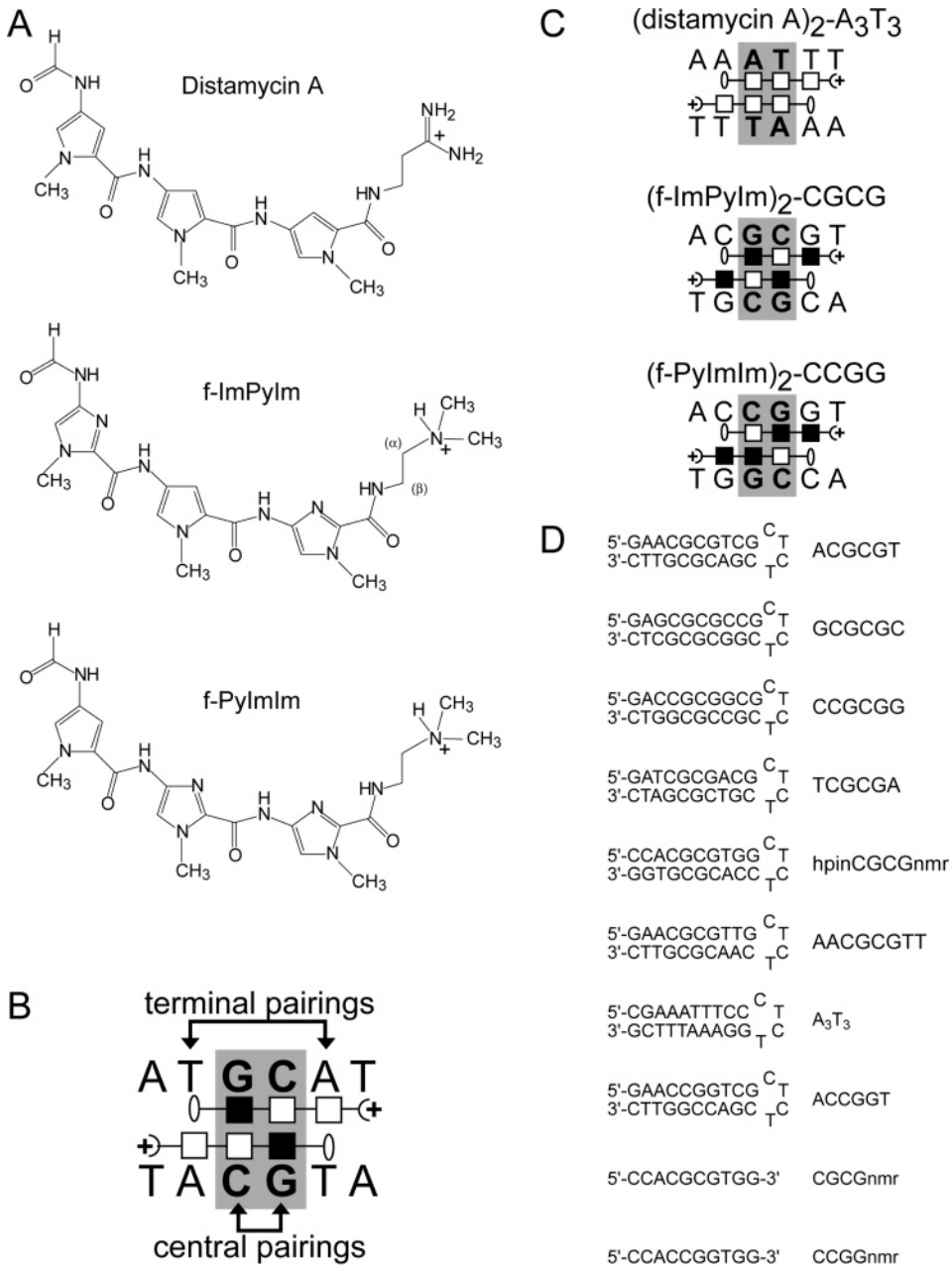

FIGURE 1: Polyamides, DNA sequences, and their expected recognition motifs. (A) Structures of naturally occurring distamycin A and its synthetic analogues, f-ImPyIm and f-PyImIm. (B) Triheterocyclic polyamide (triamide) binding diagram. Central pairings consist of heterocyclic dimers (Py and Im are white and black squares, respectively), and terminal pairings consist of formyl (oval)·heterocycle pairings. (C) Binding diagrams for triamides and their cognate DNA. (D) Synthetic 10 bp DNA hairpins and the palindromic 10-base oligomers used in the NMR studies.

pairs by Py·Im pairs (*13*−*20*). This recognition is analogous to a simple, single-letter language in which one heterocyclic pair reads one DNA base pair.

This single-letter code allows for easy design of polyamides for targeting specific (cognate) sequences; however, the affinity of closely related polyamide molecules for their cognate DNA can vary considerably. For example, the hairpin polyamide ImHpPyPy-(*R*)$^{H_2N}\gamma$-ImHpPyPy binds its cognate DNA 26- and 160-fold better than ImPyHpPy-(*R*)$^{H_2N}\gamma$-ImPyHpPy and ImPyPyPy-(*R*)$^{H_2N}\gamma$-ImHpHpPy bind their cognate sequences, respectively, even though these three polyamides each contain two Im·Py and two Hp·Py pairs (*25*). Another example is formyl (f)-ImPyIm (Figure 1A), which binds its cognate sequence, CGCG, with a 250-fold higher affinity than its structural isomer f-PyImIm (Figure 1A) binds its cognate sequence, CCGG (*26*, *27*). The investigation of triheterocyclic polyamides (triamides), which includes f-ImPyIm and f-PyImIm, led to the discovery of an expanded two-letter polyamide language (*27*). In the context of triamides, the central pairings (Figure 1B) dictate the cognate DNA sequence and also the affinity of the polyamide dimer for the cognate DNA. The central pairings rank from strongest to weakest affinity for their respective Watson−Crick cognate DNA (-ImPy- > -PyPy- ≫ -PyIm- ∼ -ImIm-), leading to formulation of the "central pairing rules" (*27*). The formyl group improves binding by at least 10-fold compared to triamides that lack the formyl group (*28*). However, the terminal pair (Figure 1B) provides little

benefit for sequence selectivity, with the exception of the f·Im pair in the f-ImPyIm dimer. In this context, the f·Im pairing is highly selective for the C·G base pair (*29*).

For its relatively small size, f-ImPyIm binds its cognate DNA very strongly ($K_{eq} = 1.9 \times 10^8$ M^{-1}). This interaction is 10-fold stronger than the affinity of its parent analogue, distamycin A, for A_3T_3 and 200-fold stronger than the affinity of its structural isomer f-PyImIm for CCGG (*26*, *27*). To better understand the two-letter language, we have further investigated the binding of f-ImPyIm, f-PyImIm, and distamycin A to their respective cognate sequences (Figure 1C).

EXPERIMENTAL PROCEDURES

f-ImPyIm, f-PyImIm, and Distamycin A. Synthesis of f-PyImIm was previously reported in ref *28*, and distamycin A was purchased from Sigma-Aldrich. The synthesis of *N*-[2-(dimethylamino)ethyl]-1-methyl-4-[1-methyl-4-(4-formamido-1-methylimidazole-2-carboxamido)pyrrole-2-carboxamido]imidazole-2-carboxamide (commonly named f-ImPyIm) is described below. A solution of *N*-[2-(dimethylamino)ethyl]-1-methyl-4-[1-methyl-4-(4-nitro-1-methylimidazole-2-carboxamido)pyrrole-2-carboxamido]imidazole-2-carboxamide (0.70 g, 1.44 mmol) in MeOH (30 mL) was hydrogenated over 5% Pd/C (200 mg) overnight. Upon removal of the catalyst, the filtrate was concentrated under reduced pressure and the residue was treated at 0 °C with freshly prepared acetic formic anhydride (9.0 mL). The reaction mixture was kept under a drying tube and was allowed to warm to room temperature overnight. The reaction was quenched with MeOH (~50 mL) and the mixture concentrated under reduced pressure. The residue was taken up in CH_2Cl_2 (50 mL) and washed with saturated $NaHCO_3$ (50 mL) followed by H_2O (50 mL). The aqueous layer was extracted with $CHCl_3$ (3 × 50 mL). The organic layers were collected, dried with Na_2SO_4, and concentrated under reduced pressure. The residue was purified on a silica gel column using a stepwise gradient from $CHCl_3$ to a 20% MeOH/$CHCl_3$ mixture. The desired fractions were collected and evaporated to dryness to yield a light yellow solid product (30 mg, 43%): mp 155 °C; TLC (10% MeOH/$CHCl_3$) R_f = 0.30; IR (Nujol) 3080, 1661, 1581, 1545, 1532, 1204, 1133, 1071, 664 cm^{-1}; ^{1}H NMR (DMSO-d_6) 10.39 (s, 1H), 10.37 (s, 1H), 10.50 (s, 1H), 8.22 (d, 1.5, 1H), 7.74 (t, 5.5, 1H), 7.51 (s, 1H), 7.49 (s, 1H), 7.37 (d, 2.0, 1H), 7.18 (d, 2.0, 1H), 3.96 (s, 3H), 3.95 (s, 3H), 3.86 (s, 3H), 3.34 (q, 6.0, 2H), 2.39 (t, 6.0, 2H), 2.18 (s, 6H); FAB-MS *m*/*z* (relative intensity) 485 ($M + H^+$, 3); HRMS (FAB) for $C_{21}H_{29}N_{10}O_4$ ($M + H^+$) calculated 485.2373, observed 485.2376; ϵ_{307} = 24 000 M^{-1} cm^{-1} in aqueous solution.

DNA Sequences. DNA molecules (10 bp) containing CTCT hairpin loops (Figure 1D and Table 1A) were designed for the isothermal titration microcalorimetry, circular dichroism, and thermal melting experiments. These DNA were chemically synthesized and desalted by Qiagen Inc. and resuspended in either 10 mM phosphate, 13 mM Na^+, and 1 mM EDTA (pH 6.2) ($PO_4$0) or 10 mM phosphate, 200 mM Na^+, and 1 mM EDTA (pH 6.2) ($PO_4$20). Nonhairpin, palindromic DNA sequences, 5′-CCACGCGTGG (CGCGnmr) and 5′-CCACCGGTGG (CCGGnmr), were used for the NMR experiments and the associated control experiments (Figure 1D). The DNA designed for footprinting is described below.

DNase I Footprinting. A 132 bp fragment of DNA was synthesized (Figure 2B) and amplified by PCR using the forward primer (5′-CTCCAGAAAGCCGGCACTCAG) and ^{32}P-radiolabeled reverse primer (5′-GTCGGTTAGGAGAGCTCCACTTG). DNase I cleavage reactions were conducted with distamycin A, f-PyImIm, or f-ImPyIm. The labeled DNA fragment (300 cps) was incubated for 30 min at room temperature in 50 mM KCl, 1 mM $MgCl_2$, 0.5 mM EDTA, 0.5 mM DTT, and 20 mM HEPES (pH 7.9) (50 μL total). DNA was then cleaved by the addition of ~10 units/μL DNase I (final concentration) in 83 mM $MgCl_2$, 83 mM $CaCl_2$, and 5 mM Tris (pH 7.5). Digestion was limited to less than 30% of the starting material to minimize the incidence of multiple strand cuts. After 3 min, digestion was quenched by addition of 2 volumes of 200 mM NaCl, 30 mM EDTA, and 1% SDS (pH 8.1). Finally, DNA was precipitated with ethanol and resuspended in 4 μL of loading dye (80% formamide, tracking dyes bromophenol and cyanol blue). The samples were denatured at 90 °C and rapidly cooled on ice prior to being loaded onto a denaturing 10% polyacrylamide gel. After electrophoresis (150 min, 70 W, 1800 V, at 55 °C in TBE buffer), the gel was transferred onto Whatman 3 MM paper, dried under vacuum, and exposed overnight to X-ray film at −80 °C.

Isothermal Titration Microcalorimetry. Experiments were performed on a VP-ITC microcalorimeter (MicroCal). The triamide solutions were prepared from the solid by resuspension in either $PO_4$0 or $PO_4$20, as indicated. The isothermal titration microcalorimetry (ITC) system was equilibrated at the noted temperature, and after an initial delay of 300 s, the triamide (50 μM) was titrated 50 times (3 μL for 7.2 s, repeated every 240 s) into 2 μM DNA. The solution was constantly mixed at 300 revolutions per minute. The reference cell contained $PO_4$0 or $PO_4$20 as appropriate for the experiment. The data were imported into Origin 7.0, and the area under the curve was integrated as a function of time. The integrated heat generated after DNA saturation was subject to a linear fit. This fit was then subtracted from the reaction integrations to normalize for nonspecific heat components, including the heat of dilution. The data were fit by the two-sets-of-sites, nonsequential model using the MicroCal version of Origin 7.0 (*30*). ΔG was calculated using eq 1

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

where R is 1.987 cal mol^{-1} K^{-1} and T is represented in units of kelvin for the appropriate temperature.

Each experiment was repeated at least twice, and error values reflect the standard deviation among different runs. The standard deviation, even when as many as five independent experiments were performed, is consistently less than 10%, which indicates they are high-quality data. When available, binding constants previously determined by surface plasmon resonance (*26*, *27*) were used to help refine the binding isotherms; however, the final values were not fixed. The isotherms were well-defined, and false local minima in the fits were avoided by careful probing of the fitting parameters.

Circular Dichroism and DNA Thermal Denaturation Studies. Circular dichroism (CD) studies were performed on a JASCO J-710 instrument with initial DNA concentrations

Table 1

(A) Thermodynamic Properties of Binding of f-ImPyIm to CGCG with Altered Flanking Sequences

DNA sequence[a]	DNA name	ITC[d] no. of A·T base pairs	ITC[d] ΔT_M[b] (°C)	ITC[d] T_M[c] (°C)	ITC[d] K_{eq} (M^{-1})	ΔG_{eq} (kcal/mol)	ΔH (kcal/mol)	$T\Delta S$ (kcal/mol)
5′-**GAA** CGCG **T**CG	ACGCGT	3	9[f]	78.5	6.9×10^7	−10.7	−7.6	3.1
5′-G**AT** CGCG **A**CG	TCGCGA	3	8	76.0	6.1×10^7	−10.6	−7.2	3.4
5′-G**A**G CGCG CCG	GCGCGC	1	0	78.0	1.4×10^7	−9.7	−5.4	4.3
5′-G**A**C CGCG GCG	CCGCGG	1	0	78.0	1.3×10^7	−9.7	−3.1	6.6
5′-G**AA** CGCG **TT**G	AACGCGTT	4	12	77.5	6.5×10^7	−10.6	−8.0	2.6
5′-G**TA** CGCG **T**CG		3	10	78.5				
5′-G**AA** CGCG **T**GG		3	10	78.0				
5′-CC**A** CGCG **T**GG[e]	hpinCGCGnmr	2	8	80.5	7.6×10^7	−10.7	−8.4	2.3
5′-GG**A** CGCG **T**CG		2	8	80.0				

(B) Thermodynamic Comparison of Binding of f-ImPyIm, f-PyImIm, and Distamycin A to Cognate and Noncognate DNA Hairpins

		SPR[h,i]	ITC[d]											
	ΔT_M[b,g] (°C)	K_{eq} (M^{-1})	K_1 (M^{-1})	K_2 (M^{-1})	K_{eq}[j] (M^{-1})	ΔG_1 (kcal/mol)	ΔG_2 (kcal/mol)	ΔG_{eq}[j] (kcal/mol)	ΔH_1 (kcal/mol)	ΔH_2 (kcal/mol)	ΔH_{eq}[k] (kcal/mol)	$T\Delta S_1$ (kcal/mol)	$T\Delta S_2$ (kcal/mol)	$T\Delta S_{eq}$[k] (kcal/mol)
f-ImPyIm														
ACGCGT[l]	9	1.9×10^8	9×10^6	5×10^8	7×10^7	−9.5	−11.9	−10.7	−7.7	−7.6	−7.6	1.8	4.3	3.1
ACCGGT	1	2.2×10^5	8×10^3	6×10^6	2×10^5	−5.3	−9.2	−7.2	3.9	−4.5	−0.3	9.2	4.7	7.0
AAATTT	1	5.3×10^4			$\ll 1 \times 10^5$									
f-PyImIm														
ACGCGT	0	3.2×10^3												
ACCGGT	3	8.5×10^5	4×10^5	9×10^6	2×10^6	−7.6	−9.5	−8.6	0.7	−5.2	−2.3	8.3	4.3	6.3
AAATTT		2.3×10^4												
distamycin A														
ACGCGT	0	3.2×10^3			$\ll 1 \times 10^5$									
ACCGGT	2	2.3×10^4												
AAATTT	14	1.7×10^7	2×10^8	4×10^5	3×10^7	−11.3	−9.0	−10.2	−4.7	−11.5	−8.1	6.6	−2.5	2.1

[a] Sequence of the top strand of hairpin DNA. All DNA molecules have a 5′-CTCT hairpin loop and a complementary bottom strand. [b] ΔT_M experiments were performed in $PO_4$0, which contains 13 mM Na^+. [c] T_M of the bound complex at a 3:1 f-ImPyIm:DNA hairpin ratio. [d] ITC experiments were performed at 25 °C in $PO_4$20, which contains 200 mM Na^+. [e] hpinCGCGnmr is the hairpin analogue to the nonhairpin DNA CCACGCGTGG, which was used for the NMR experiments. [f] From ref *27*. [g] From refs *27* and *28*. [h] SPR experiments were performed at 25 °C in $PO_4$20, which contains 200 mM Na^+.[i] From refs *26* and *27*. [j] $K_{eq} = (K_1K_2)^{1/2}$; $\Delta G_{eq} = (\Delta G_1 \Delta G_2)^{1/2}$. [k] $\Delta H_{eq} = (\Delta H_1 \Delta H_2)/2$; $T\Delta S_{eq} = (T\Delta S_1 \times T\Delta S_2)/2$. [l] Sequences shown in Figure 1D.

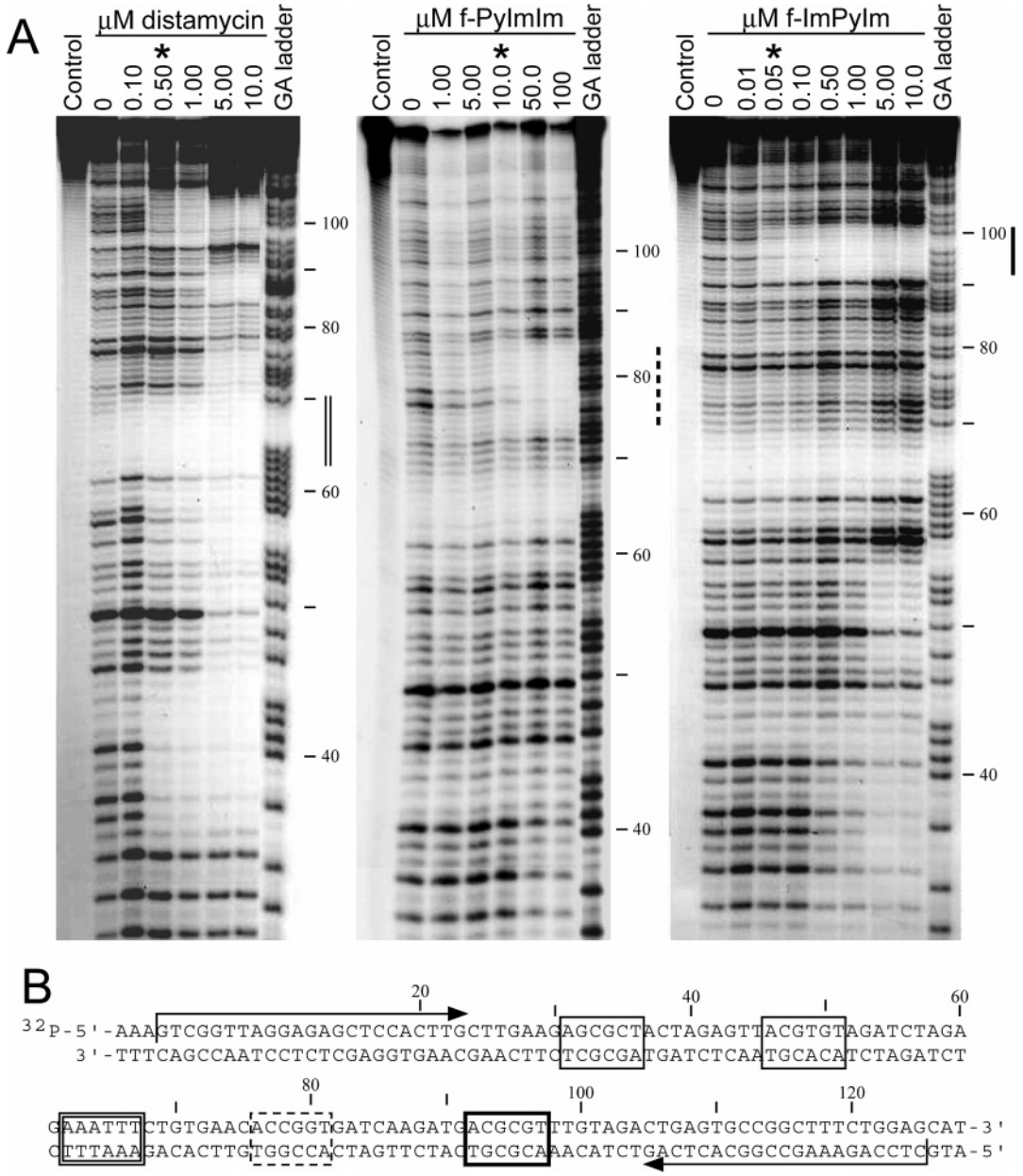

FIGURE 2: DNase I footprinting. (A) Concentrations of distamycin A, f-PyImIm, and f-ImPyIm. Asterisks denote the concentration of triamide at which protection is initially observed. The positions of the corresponding cognate sequences are highlighted by double, dashed, and thick lines for distamycin A, f-PyImIm, and f-ImPyIm, respectively. Control lanes lack DNase I, and the ladders denote the GA nucleotides. The lanes in each gel were internally normalized to nucleotides with a minimal variation in intensity (nucleotides 32, 46, and 46, respectively). (B) DNA used for the footprinting experiment. The top strand is 5′-labeled with ^{32}P. Arrows indicate the direction and length of the primers that were used. Cognate sequences are identified by double, dashed, and thick boxes for distamycin A, f-PyImIm, and f-ImPyIm, respectively. Thin boxes indicate sequences which f-ImPyIm targets with lower affinity than its cognate ACGCGT.

of 9 μM in $PO_4$20, as described previously (*28*). Thermal denaturation (T_M) studies were performed using the Cary 100 BioMelt system (Varian) with 1 μM DNA in $PO_4$0, as described previously (*28*).

Nuclear Magnetic Spectroscopy. NMR experiments were performed on a Varian Inova 500 MHz NMR spectrometer. The 10-base, palindromic DNAs (Figure 1D) were synthesized and desalted by Qiagen Inc.

Titration of Triamides to the DNA. Formation of the triamide−DNA complex was monitored by changes in the imino proton signals. f-ImPyIm, f-PyImIm, and the palindromic DNAs were independently dissolved in 10 mM sodium phosphate (pH 7.2) and a 90% H_2O/10% D_2O mixture. The triamides were titrated in 0.5 molar ratio increments to 2:1 f-ImPyIm:$(CGCGnmr)_2$ ratios and 2.5:1 f-PyImIm:$(CCGGnmr)_2$ ratios. The imino ^{1}H NMR spectra were collected at 5 °C. The binomial 1−3−3−1 pulse sequence suppressed the HDO signal, which was set at 4.60 ppm (*31*). The spectral width was set at 10 000 Hz; the relaxation delay was set at 10 s, and 64 FIDs were collected for each experiment. The data were processed with a line broadening factor of 2.

Nonexchangeable NMR Spectra for the DNA−Triamide Complexes. The DNA−triamide solutions from the experiment described above were lyophilized and redissolved in 10 mM sodium phosphate (pH 7.2) in 99.9% D_2O. The sample was repeatedly lyophilized and redissolved in 99.96% D_2O until the HDO peak was minimized in the one-

dimensional (1D) ^{1}H NMR spectrum. The experiments were carried out at 25 °C, and the sweep width was set at 8000 Hz; 1024 scans were collected for each experiment, and the FIDs were processed with a line broadening of 2. At 25 °C, the HDO signal was set at 4.80 ppm (*32*).

The T_1 relaxation times were determined using a nonselective inversion recovery method. Experimental conditions were as follows: sweep width of 8000 Hz, 180° pulse of 19.8 μs, 90° pulse of 9.9 μs, variable delay "τ" values of 0.1, 0.2, 0.3, 0.4, 0.5, and 0.75 s, relaxation delay of 10 s after each scan, temperature of 25 °C, and 32 scans collected for each experiment. The FIDs were processed with a line broadening of 0.5.

1D NOE difference experiments with the DNA−triamide complexes were conducted at 25 °C. An irradiation time of 0.2 s was used, and 256 FIDs were collected for each experiment. A sweep width of 8000 Hz was employed, and the results were processed with a line broadening of 5.

The two-dimensional (2D) correlated experiments were performed using the COSY program with 64 increments. The acquisition time was 0.21 s, the pulse width 9.9 μs for 90°, the temperature 25°, and the relaxation delay 2 s. The data were processed with a line broadening of 2 for both dimensions, and the contour plot was symmetrized.

NMR Peak Assignments. The nonexchangeable base protons were identified for free DNA and the polyamide−DNA complexes using conventional 1D ^{1}H NMR, T_1 relaxation, COSY, NOESY, and NOE experiments (described above). Differences in buffer conditions between those previously reported for $(CGCGnmr)_2$ (*33*) account for slight differences in peak assignments reported herein (Table S1A of the Supporting Information).

Determination of Distance Constraints. In the $(f\text{-}ImPyIm)_2$−$(CGCGnmr)_2$ complex, the distance between the CH_2 groups α and β to the dimethylamine moiety and A_3-H2 were determined to be 3.3 (±15%) and 3.8 Å (±15%), respectively. These distances were determined using eq 2:

$$r_{ij}^{\,6} = (-5.68 \times 10^{10} \times \tau_C)/(N_{ij}\tau_{null}/\ln 2) \qquad (2)$$

where the steady-state NOE (N_{ij}) is the ratio of the induced over excited signal integrations, τ_{null} is the midpoint of signal relaxation, and the normalized τ_C for the system is based on an r_{ij} of 2.7 Å for the T_8CH_3−T_8H6 NOE signal (*34*).

NMR Control Experiments. Samples of the 2:1 f-ImPyIm−$(CGCGnmr)_2$ and 2:1 f-PyImIm−$(CCGGnmr)_2$ complexes were taken directly from the NMR tube after the experiments were completed. The solutions were diluted in 10 mM phosphate, 50 mM Na^+, and 1 mM EDTA (pH 7.0) ($PO_4$5) for CD and T_M experiments. In addition, CD and T_M experiments were performed using fresh samples of f-ImPyIm, f-PyImIm, and the nonhairpin, palindromic DNA molecules.

Molecular Dynamics Simulations. Two f-ImPyIm polyamide molecules were docked into the minor groove of a DNA decamer representing its cognate sequence, $(CGCGnmr)_2$. The DNA decamer coordinates were obtained from the most representative conformation of this DNA (PDB entry 1KKV) determined by NMR analysis (*33*). The f-ImPyIm dimer was modeled by modifying the f-ImImIm dimer structure from PDB entry 1CYZ (*10*). The f-ImPyIm dimer was docked into $(CGCGnmr)_2$ by distance and electrostatic energy monitoring using SYBYL 7.0. A 1000-step minimization was then run using SYBYL 7.0 with six layers of water to eliminate unfavorable van der Waals contacts. Using AMBER 7.0 [*xleap* subprogram (*35*)], 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 3372 TIP3P waters (*36*), and the initial system had approximate dimensions of 50 Å × 50 Å × 60 Å. The system was equilibrated, and 25 kcal mol^{-1} Å^{-2} constraints were placed on the DNA, the f-ImPyIm, and the counterions. In addition, NMR analysis provided distance restraints of 3.3 Å (±15%) between the α CH_2 of f-ImPyIm and A_3H2 of $(CGCGnmr)_2$. The temperature of the system was gradually increased to 300 K, and a series of constant-volume simulations were conducted while gradually releasing the energetic constraints and distance restraints (*37*−*39*). Distance restraints on the DNA termini were maintained to ensure helical integrity. A 75 ps constant-pressure simulation was conducted, and then the 3 ns production run was initiated.

Solvent Accessible Surface Area from MD Simulations. The solvent accessible surface area (SASA) calculations were conducted using GRASP (*40*) with probe radii of 1.7683 Å (*41*), as previously described (*42*). In brief, carbon, carbon-bound hydrogen, and phosphorus atoms are considered nonpolar and all others are polar. The change in SASA (ΔA) was computed using eq 3.

$$\Delta A = A_{complex} - (A_{free-DNA} + A_{free-ligand}) \qquad (3)$$

The structure of the free DNA, $d(CCACGCGTGG)_2$, was obtained from the best representative structure from an ensemble of 20 NMR conformers from PDB entry 1KKV (*33*). The heat capacity change was calculated using three previously derived equations (eq 4 from refs *43*−*46*, eq 5 from refs *47*−*51*, and eq 6 from ref *52*).

$$\Delta C_{pSASA} = (0.32 \pm 0.04)\Delta A_{np} - (0.14 \pm 0.04)\Delta A_p \qquad (4)$$

$$\Delta C_{pSASA} = (0.45 \pm 0.02)\Delta A_{np} - (0.26 \pm 0.03)\Delta A_p + (0.17 \pm 0.07)\Delta A_{OH} \qquad (5)$$

$$\Delta C_{pSASA} = (0.382 \pm 0.026)\Delta A_{np} - (0.121 \pm 0.077)\Delta A_p \qquad (6)$$

The two OH groups at the 3′-ends of the DNA duplex are not involved in complex formation and make an insignificant contribution to the total area change. Thus, only eq 4 includes these groups in the polar term (ΔA_{OH}).

RESULTS AND DISCUSSION

DNase I Footprinting: Sequence Selectivity of f-ImPyIm. Binding of distamycin A, f-PyImIm, and f-ImPyIm to a 125 bp DNA was investigated by DNase I footprinting. DNase I does not cleave well at the A·T-rich site, but cleavage protection upon addition of 0.5 μM distamycin A is visible around nucleotide 64, which corresponds to A_3T_3 (double lines in Figure 2). Additional protected regions that center on A·T-rich sequences are visible for distamycin A. For example, nucleotides 35−46 are protected from cleavage at 0.50 μM, and this 11 bp segment contains eight A·T base pairs.

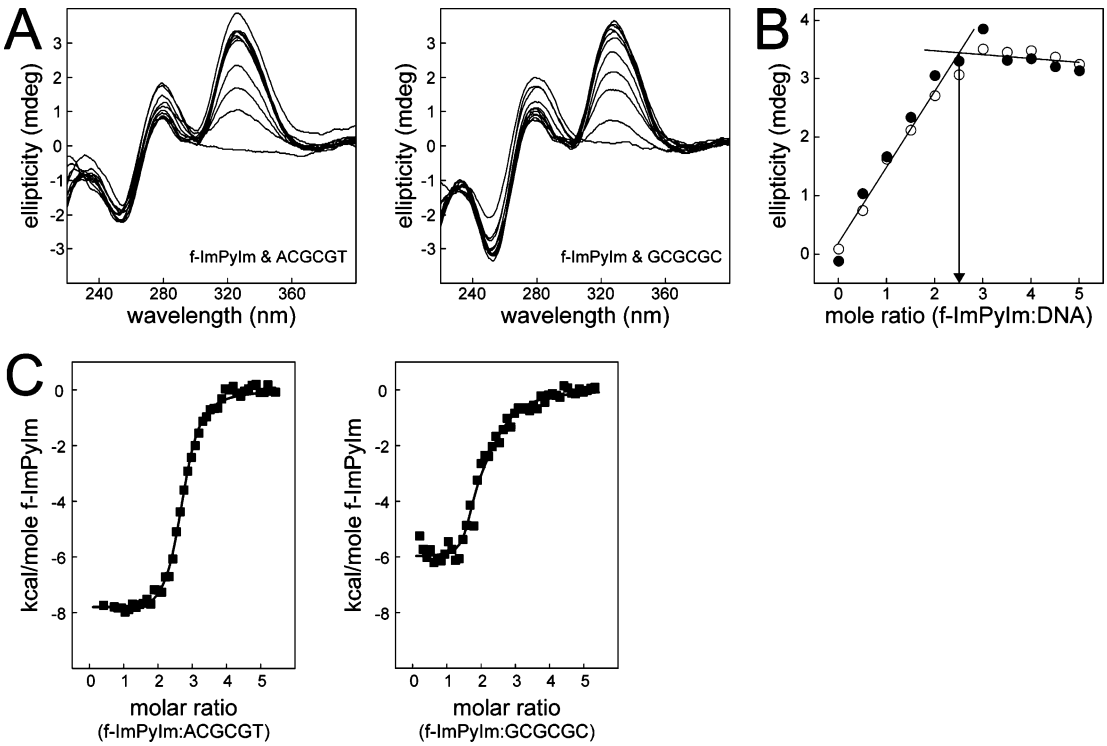

FIGURE 3: Selectivity of f-ImPyIm for changes in flanking sequence identity. (A) CD spectra of f-ImPyIm titrated into ACGCGT and GCGCGC (left and right panels, respectively). (B) Induced peak ellipticity as a function of f-ImPyIm:DNA ratio. Solid circles are data for ACGCGT and open circles data for GCGCGC. The arrow indicates the ratio at which the DNA is saturated. (C) Integrated ITC titrations of f-ImPyIm into ACGCGT and GCGCGC at 200 mM Na^+ and 25 °C.

At 10 μM f-PyImIm, modest protection of ACGCGT is exhibited, and this region is completely protected at 50 μM f-PyImIm. The region centering on nucleotide 79 is the only region visibly protected from DNase I cleavage (dashed lines in Figure 2). Alternatively, only 0.05 μM f-ImPyIm is required to protect ACGCGT (thick lines in Figure 2).

Surface plasmon resonance (SPR) has provided reliable binding constants for well-defined target sequences (*26*−*29*, *53*, *54*). Herein, the ability of ligands to protect DNA from cleavage by DNase I is utilized to probe the relative affinity and not to quantify binding constants. f-ImPyIm protects its cognate sequence, ACGCGT, at a concentration 10-fold lower than that at which distamycin A protects A_3T_3. This difference corresponds well to SPR-determined binding affinities for f-ImPyIm and distamycin A binding short DNA hairpins that contain ACGCGT (1.9×10^8 M^{-1}) and A_3T_3 (1.7×10^7 M^{-1}), respectively (*26*, *27*). f-ImPyIm protects ACGCGT from DNase I cleavage at a concentration 200-fold lower than that at which f-PyImIm protects ACCGGT. This difference agrees with previous SPR results where f-PyImIm binds a DNA containing the sequence ACCGGT with an affinity of 8.5×10^5 M^{-1} (*26*). Thus, as judged by DNase I footprinting, f-ImPyIm binds its cognate sequence better than distamycin A and f-PyImIm bind their cognate DNA, and this observation agrees well with SPR-determined binding affinities (*26*, *27*).

Footprinting assesses the sequence preference of the triamide when multiple DNA sequences are available in a single solution mixture; 20−200-fold higher concentrations of f-ImPyIm are necessary to protect the G·C-rich sites, AGCGCT and ACGTGT, respectively (thin boxes in Figure 2B). Interestingly, CCGG is not protected by f-ImPyIm. The SPR-determined affinities of f-ImPyIm for ACACGT and CCGG are 5×10^5 and 2×10^5 M^{-1}, respectively, significantly lower than the affinity of f-ImPyIm for ACGCGT (*27*). Thus, the footprinting results show that f-ImPyIm is selective for its cognate DNA when other DNA sequences are equally available for binding. Under these competitive conditions, the relative affinities of these triamides for their cognate sequences agree well with the SPR experiments.

Selectivity of f-ImPyIm for Sequences Flanking the CGCG Binding Site. The footprinting and SPR (*27*, *29*) results show that f-ImPyIm is selective for the CGCG sequence over related and unrelated DNA sequences. The binding preference of f-ImPyIm for DNA sites that comprise the CGCG site but have different flanking sequences has been probed. A series of DNA hairpins (10 bp) were designed to probe the flanking sequence selectivity (sequences of the top strand are shown in Table 1A). Sequences discussed in further detail are given shorter names (Figure 1D and Table 1A). The ability of f-ImPyIm (3:1 excess ratio of ligand to DNA hairpin) to stabilize the melting for each of these DNA sequences is monitored by the change in transition midpoint (ΔT_M). The ΔT_M values range from 0 to 12 °C for the sequences that were tested, with A·T-rich DNA exhibiting larger ΔT_M values than G·C-rich DNA (Table 1A). In contrast, the T_M values for the complexes are clustered at 78.5 ± 1 °C. Therefore, there is an apparent T_M maximum for the stabilization of f-ImPyIm bound to CGCG-containing DNA molecules.

CD studies of f-ImPyIm titrated into ACGCGT and GCGCGC are shown in Figure 3A. The resulting spectra, including maxima and intensities of the induced peaks, are nearly identical. The ellipticities of the induced peaks are plotted against the molar ratio of f-ImPyIm to DNA (Figure 3B), and both DNA sequences are indistinguishable and

saturated at 2.5:1 f-ImPyIm:DNA ratios. Thus, as judged by CD, f-ImPyIm binds both sequences well.

Several of these DNA sequences were studied by isothermal titration microcalorimetry (ITC). The ΔG_{eq}, ΔH_{eq}, and $T\Delta S_{eq}$ values for binding of the f-ImPyIm dimer to ACGCGT, TCGCGA, CCGCGG, and GCGCGC DNA hairpins that differ only in the identity of the base pairs that immediately flank the CGCG site are given in Figure 1D. The ΔG_{eq} values for GCGCGC and CCGCGG are only 1 kcal/mol lower than with the other DNA sequences. Binding of f-ImPyIm to GCGCGC and CCGCGG has smaller favorable enthalpic contributions (Figure 3C) and larger favorable entropic contributions than f-ImPyIm exhibits in binding to the other CGCG-containing DNA sequences. Previous structural studies (*10*, *55*), and the structural studies discussed below, show that dimethylamine groups interact well with A·T base pairs. Placement of G·C base pairs incurs a slight enthalpic penalty, which is mostly recouped by favorable entropic interactions (Table 1A). Thus, f-ImPyIm has a slightly stronger, but not very significant, preference for A·T base pairs immediately adjacent to the CGCG sequence.

The flanking sequences of hpinCGCGnmr and AACGCGTT DNA differ from the DNA discussed above by several base pairs (Figure 1D). The ΔG of f-ImPyIm is neither enhanced nor reduced as a consequence of the larger changes in flanking sequence (Table 1A). Therefore, f-ImPyIm has selectivity for only the 4 bp sequence CGCG, with a very slight binding preference for (A_T)CGCG(A_T) over (G_C)-CGCG(G_C).

The ΔT_M data (Table 1A) showed a significant preference for A·T-rich DNA, which is misleading. The A·T content (bold in Table 1A) appears to affect the stability of the complex. The ITC experiments showed that the 0 °C ΔT_M value is not the result of weak binding; it is rather the result of the inability of f-ImPyIm to further stabilize GCGCGC against thermal melting. However, it is the lowered T_M of A·T-rich (78 °C), unbound DNA that artificially implies that A·T-rich flanking sequences are preferred. As a comparison, the ΔT_M values for f-ImPyIm bound to the noncognate DNA sequences containing the sequences A_3T_3, ACCGGT, and CACG were ≤1 °C, and the T_M values for the DNA complexed with f-ImPyIm are 62, 70, and 71 °C, respectively (data not shown). SPR (*27*), DNase I footprinting (Figure 2), and ITC (Table 1B) studies have confirmed the poor binding of f-ImPyIm to noncognate DNA sequences. These findings suggest that monitoring ΔT_M values does not provide an accurate reflection of the relative binding affinities of f-ImPyIm for cognate DNA sequences. ΔT_M values obtained from thermal melting experiments are often utilized to screen for relative differences in binding affinities; however, these experiments intrinsically monitor binding affinity at higher temperatures and do not necessarily correlate to the affinities at considerably lower temperatures (i.e., room temperature). Therefore, as shown by these experiments, thermal melts can result in the inaccurate determination of relative binding affinities.

Isothermal Titration Microcalorimetry: Thermodynamic Profiles of f-ImPyIm, f-PyImIm, and Distamycin A. Isothermal titration microcalorimetry (ITC) measures the change in heat when f-ImPyIm, f-PyImIm, and distamycin A bind to their cognate sequences, ACGCGT, ACCGGT, and A_3T_3, respectively (Figure 4A and Table 1B). The binding of all three triamides to their cognate DNA sequences is exothermic; however, f-PyImIm releases less heat than the other two triamides upon binding their respective cognate DNA sequences. The lower magnitude of heat released correlates to the significantly weaker binding observed for f-PyImIm compared to that of both f-ImPyIm and distamycin A (*26*, *27*). The integrated data are best fit by a two-sets-of-sites and nonsequential 2:1 binding isotherm, and formation of the 2:1 complex is consistent with previous results (*26*, *27*). The fit data provide thermodynamic constants for formation of the entire complex (subscript eq) and binding of each molecule of the triamide to the DNA (subscripts 1 and 2). Table 1B compares the K_{eq} values determined by SPR and ITC. The two-sets-of-sites, nonsequential model implies that the oligonucleotide presents two binding sites for the triamides within the self-complementary hairpin DNA. The triamides would saturate the sites independently to form the stacked 2:1 complexes. Since the one-site model and the sequential model could not fit the data, it further suggests that the triamides would not be likely to bind its cognate sequence in a single step as a predimerized unit. It is also unlikely that the triamides would sequentially saturate one site before beginning to fill the second complementary site.

f-ImPyIm, f-PyImIm, and distamycin A each exhibit positive $T\Delta S$ values for formation of the 2:1 complex, which are consistent with the recently proposed thermodynamic signature for minor groove binders (*56*). The binding of intercalators is typically less favored or disfavored entropically due to the significant structural changes in DNA upon binding by the intercalator. Minor groove binders do not alter DNA structure as significantly and consequently do not pay the same entropic penalty as intercalators. This binding mode is supported by the NMR studies discussed below.

f-ImPyIm and distamycin A (Figure 4B,C) each exhibit negatively sloped, temperature-dependent thermodynamic profiles. The ΔG_{eq} values (data not shown) are independent of temperature over the range that was tested. However, increased experimental temperatures result in an increase in the enthalpy magnitudes and a concurrent decrease in the favorable entropic components (Figure 4B,C). The temperature-dependent slopes for ΔH and $T\Delta S$ are parallel. Therefore, changes in one compensate changes in the other, which results in the constant ΔG value observed within the temperature range that was tested.

The individual thermodynamic parameters (e.g., ΔH_1 and ΔH_2) are reported for each experiment. Distinct values were determined for binding of two separate molecules; however, the ITC experiment alone cannot distinguish between which values correspond to binding of the first and second polyamide molecules. This problem was circumvented through the use of results from both SPR (*26*, *27*) and NMR experiments (Figure 5A; *13*, *14*), which show that binding of f-ImPyIm and f-PyImIm to their cognate DNA is positively cooperative and that distamycin binding is negatively cooperative. Thus, the thermodynamic parameters were assigned accordingly.

f-ImPyIm exhibits essentially identical ΔH_1 and ΔH_2 values (Figure 4B). The positive cooperativity ($K_2 > K_1$) of binding of f-ImPyIm to ACGCGT [as observed by SPR (*27*)] appears to be the result of differences in the entropic contributions (Table 1B). $T\Delta S_2$ is more favorable, by 2.5 kcal/mol, than $T\Delta S_1$ at 25 °C. As a result, K_2 is greater than

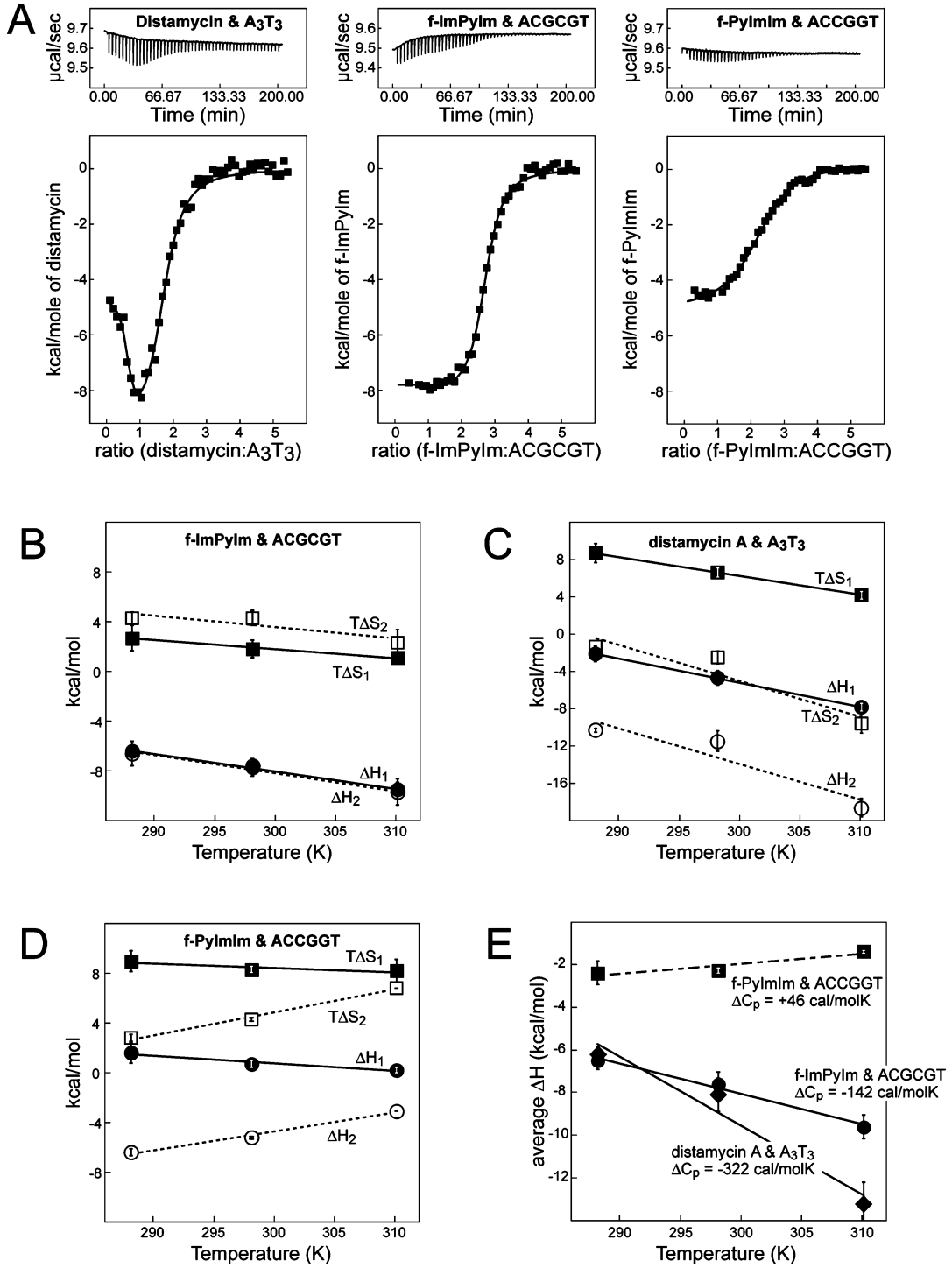

FIGURE 4: Isothermal titration microcalorimetry and thermodynamic parameters. (A) Sample thermograms for the three triamides and the resulting integrated data and data fits. The first data point was eliminated in the data fit. (B−D) Effects of temperature on the enthalpic (ΔH) and entropic ($T\Delta S$) parameters for binding of f-ImPyIm to ACGCGT, binding of distamycin A to A_3T_3, and binding of f-PyImIm to ACCGGT. Filled squares, solid lines, and the subscript 1 denote binding of the first triamide in the dimer. Empty squares, dashed lines, and the subscript 2 denote binding of the second molecule. (E) The change in heat capacity (ΔC_p) is the slope of the line generated by plotting ΔH_{eq} as a function of temperature. ΔH_{eq} values for f-ImPyIm, distamycin A, and f-PyImIm are shown as circles, diamonds, and squares, respectively. Error bars are generated from the standard deviation of multiple experiments.

K_1 by 60-fold (ITC), which is lower than the ~3000-fold difference observed by SPR (*27*). This technique-dependent discrepancy in K_2/K_1 may be a result of the higher ligand concentrations (50 μM) necessary for the ITC experiments to be performed compared to the low concentrations in SPR (<0.2 μM) or the result of physical differences in the two techniques. Regardless of the magnitude difference, f-ImPyIm binds its cognate DNA with positive cooperativity, which is further supported by NMR studies discussed below. It appears that binding of both molecules of f-ImPyIm to ACGCGT results in formation of favorable structural interactions (negative ΔH) and that complex formation is further improved by favorable changes in entropy (positive $T\Delta S$). The favorable entropic contributions may include release of restrained water molecules from the $(\text{f-ImPyIm})_1$−ACGCGT complex upon binding of the second f-ImPyIm molecule.

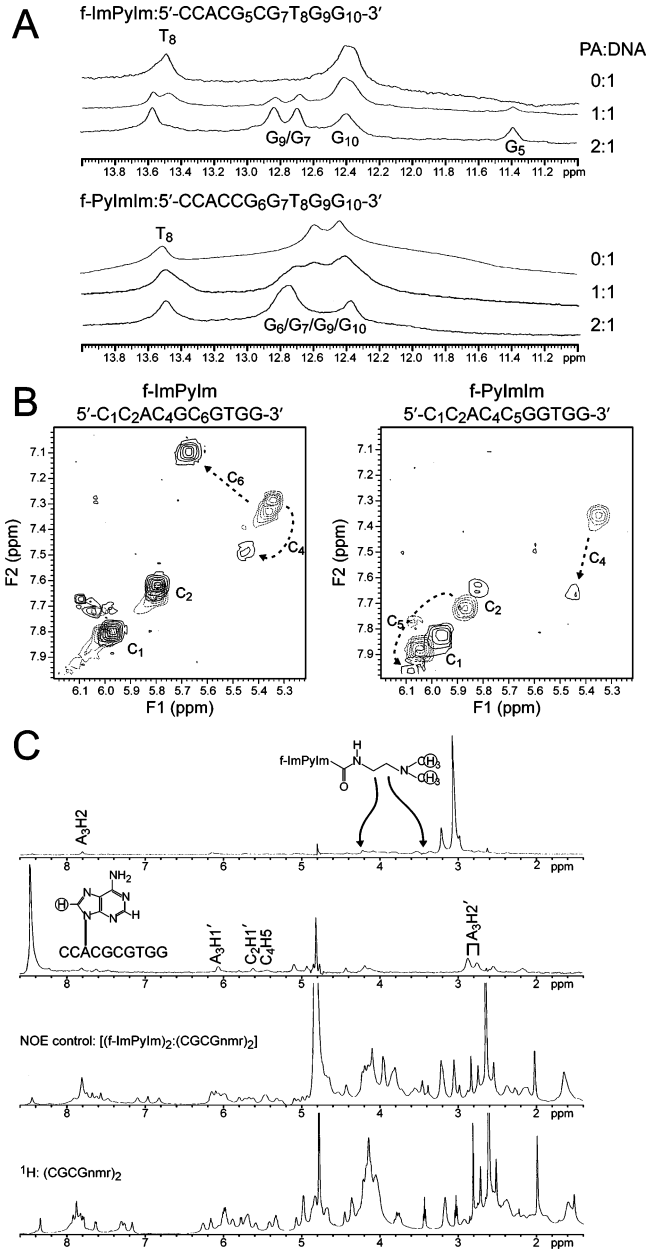

FIGURE 5: Assessment of stoichiometry, cooperativity, and binding site by NMR spectroscopy. (A) ^{1}H NMR spectra for the imino protons. Titration of f-ImPyIm and f-PyImIm into $(CGCGnmr)_2$ and $(CCGGnmr)_2$ (top and bottom panels, respectively). The HDO signal was referenced at 4.60 ppm. G_9/G_7 indicates that a general assignment of these imino protons to the two adjacent peaks in the top panel. G_6, G_7, G_9, and G_{10} are nonspecifically assigned to one of the two peaks at 12.76 and 12.35 ppm in the bottom panel. Nucleotide numbering is indicated in the sequence written at the top of each panel. (B) Expanded COSY spectra for the DNA−triamide complexes. C_5 and C_6 cross-peaks in COSY spectra in the absence (speckled) and presence (solid) of f-ImPyIm (left) or f-PyImIm (right). Shifts of the central cytosines are indicated by the dashed arrows. (C) Sample NOE spectra show the irradiated proton (circled) and the corresponding excited protons. The top two panels show irradiation of the polyamide C-terminus and A_3-H2 of the DNA. The bottom panels show the NOE control of the 2:1 f-ImPyIm−$(CGCGnmr)_2$ complex irradiated at 12.67 ppm and the ^{1}H NMR of the free DNA. The peaks at 4.80 ppm are HDO.

The binding affinity of f-ImPyIm for the G·C-rich, noncognate DNA, ACCGGT, was determined by ITC to be 2×10^5 M^{-1}, which is comparable to the K_{eq} observed by SPR (Table 1B). Binding to this sequence is predominantly entropically driven ($T\Delta S_{eq} = 7.0$ kcal/mol) with minimal favorable enthalpic contributions ($\Delta H_{eq} = -0.3$ kcal/mol). Binding is positively cooperative, which is mostly due to an unfavorable ΔH_1 and favorable ΔH_2 and $T\Delta S_2$ values. Titration of f-ImPyIm into A_3T_3 produced no detectable heat; therefore, the affinity is reported to be $<1 \times 10^5$ M^{-1}.

The ITC thermogram of binding of distamycin A to A_3T_3 (Figure 4A) shows the initial heat evolution and then decrease in the magnitude of heat evolved with sequential titrations. The integrated data (Figure 4A) exhibit maximum heat emission at a 1:1 distamycin A:A_3T_3 ratio, and the DNA is saturated between triamide:DNA molar ratios of 2:1 and 3:1. Thus, when the DNA is approximately half-saturated, it has switched from one mode of binding to another. This titration pattern is consistent with the negatively cooperative binding of distamycin A to A_3T_3, as shown previously by SPR (*26*), CD (*57*), and NMR (*14*). Binding of the first molecule of distamycin A is entropically driven, and binding of the second molecule is significantly disfavored entropically ($T\Delta S_2$ is negative at each temperature that was studied; Figure 4C). ΔH_1 is marginally favored, but ΔH_2 has a large negative value. Together, these findings show that the disfavored $T\Delta S_2$ value contributes significantly to the observed negative cooperativity. The ITC data reported herein agree with reported results for distamycin A (*58*). Thus, binding of the first molecule of distamycin A, which is entropically driven, may involve significant release of bound water and an increase in the level of water disorder. The highly negative ΔH_2, which promotes dimer formation, may involve favorable interactions between distamycin A and A_3T_3 that are not present in the 1:1 complex.

Binding of distamycin A to the C·G-rich DNA, ACGCGT, was also studied by ITC. The addition of distamycin A to ACGCGT produced no heat; therefore, the binding affinity is reported to be $<1 \times 10^5$ M^{-1}. This weak binding is supported by previous SPR data (Table 1B and ref *28*).

Binding of f-PyImIm to ACCGGT is considerably more complex than that of the other two triamides discussed above (Figure 4D and Table 1B). The ΔG values are constant at the temperatures that were tested; however, binding of the first and second molecules of f-PyImIm shows very different temperature profiles. ΔH_1 and $T\Delta S_1$ exhibit slightly parallel, negative slopes as a function of the increasing temperature. Binding of the first molecule is not favored enthalpically, but this binding is driven by favorable entropic interactions that are comparable in magnitude to $T\Delta S_1$ for binding of distamycin A to A_3T_3. Binding of the first molecule may involve significant dissociation of water from the DNA (positive $T\Delta S_1$), but net accumulation of favorable contacts does not occur between the triamide and the DNA (positive ΔH_1). In contrast to binding of the first molecule, ΔH_2 and $T\Delta S_2$ exhibit parallel positive slopes as a function of temperature. This positive slope exhibits a temperature-dependent loss of favorable enthalpic interactions coinciding with a gain in favorable entropic interactions. Nonetheless, the second molecule has both favorable ΔH_2 and $T\Delta S_2$ components. Thus, positive cooperativity for f-PyImIm is the result of disfavored changes in enthalpic interactions upon binding of the first molecule to DNA (ΔH_1). It is binding of the second molecule of f-PyImIm that allows for favorable interactions with the DNA and the polyamide in the 1:1 f-PyImIm−DNA complex (negative ΔH_2). K_2 is 20-fold

greater than K_1 as measured by both ITC (Table 1B) and SPR (*26*). The opposing temperature profiles for binding of the two molecules of f-PyImIm to its cognate sequence may correlate to the overall poor binding affinity of this triamide to ACCGGT.

Heat Capacity and Solvent Accessible Surface Area Calculations. The changes in heat capacity (ΔC_p) of (f-ImPyIm)$_2$−ACGCGT, (f-PyImIm)$_2$−ACCGGT, and (distamycin A)$_2$−A_3T_3 complexes were determined from the slope of ΔH_{eq} as a function of temperature (Figure 4E). Both distamycin A and f-ImPyIm exhibit negative ΔC_p values, −322 and −142 cal mol^{-1} K^{-1}, respectively. In contrast, f-PyImIm demonstrates a positive ΔC_p value of 46 cal mol^{-1} K^{-1}. The opposite signs of ΔC_p observed for f-ImPyIm and f-PyImIm are additional evidence of the fundamental difference between these two structural isomers and their ability to bind their cognate DNA sequences.

Negative ΔC_p values have been observed using ITC for a variety of polyamides and small molecules binding DNA (*42*, *52*, *53*, *58*, *59*). The change in the solvent accessible surface area (SASA) has been shown to be a large component of ΔC_p; however, it has become clear that SASA is not the sole contributor to ΔC_p (*42*). Other physical contributions include changes in protonation, covalent bond vibrations, and covalent bond stretching (*42*). In the simplest terms, the negative ΔC_p values relate to favorable changes in SASA upon formation of the complex. SASA can be theoretically determined from the three-dimensional model of the (f-ImPyIm)$_2$−CCACGCGTGG complex (modeling will be described below). The Cornell et al. (*41*) probe radii (1.7683 Å) were used to calculate the ΔC_p from the solvent accessible surface area, and values ranging from −177 to −222 cal mol^{-1} K^{-1} were determined according to eqs 4−6 (Table S2 of the Supporting Information). The calculated and experimental (−142 cal mol^{-1} K^{-1}) ΔC_p values are in reasonably good agreement. In addition to the other physical contributions to ΔC_p that have been described, the DNA molecules were different for experimental (hairpin) and theoretical (nonhairpin) studies.

Nuclear Magnetic Resonance: Stoichiometry, Cooperativity, and Structure of f-ImPyIm− and f-PyImIm−DNA Complexes. ^{1}H NMR studies were performed in an effort to understand the global structures and dynamics of DNA complexed with f-ImPyIm and f-PyImIm. These studies have provided a framework for better understanding the differences between the -ImPy- and -PyIm- central pairings.

Exchangeable Imino Proton NMR Studies. The (f-ImPyIm)$_2$−(CGCGnmr)$_2$ and (f-PyImIm)$_2$−(CCGGnmr)$_2$ complexes were formed in a 90% H_2O/10% D_2O solution, and the imino proton signals were monitored as a function of the mole ratio of triamide to DNA (Figure 5A). In the absence of f-ImPyIm, (CGCGnmr)$_2$ has two peaks between 14.0 and 11.0 ppm that correlate to imino protons of the guanosine and thymidine residues (Figure 5A, top panel). The peak at 13.49 ppm is T_8, and the other peak corresponds to all four guanosine imino protons. Peaks corresponding to all five possible imino protons (due to duplex symmetry) are visible upon addition of 2 molar ratio equivalents of f-ImPyIm to (CGCGnmr)$_2$. T_8 shifted downfield by 0.07 ppm. It is likely that G_{10} of (CGCGnmr)$_2$ was least affected by binding of f-ImPyIm and has therefore been assigned to 12.38 ppm. Two new peaks appeared at 12.82 and 12.68 ppm and correspond to G_7 and G_9, but these two peaks have not been further assigned. One guanine imino proton shifted significantly upfield to 11.38 ppm. This peak is most affected upon binding of the triamide and is, therefore, assigned to the central base in the duplex, G_5. This assignment is corroborated by the assignment of a similarly shifted peak to the central G of the GAACT*G*GTTC duplex upon binding by the f-ImImIm dimer. T·G mismatches (denoted with an asterisks) are bound strongly by Im·Im pairings (*10*, *59*). This G_5 base is involved in DNA recognition of the central pairing motif, -ImPy-, by which two imidazoles of the two central Im·Py pairs bind each G_5 in the duplex DNA (Figure 1C).

f-ImPyIm was titrated into (CGCGnmr)$_2$ in 0.5:1 molar ratio increments (0.5:1 and 1.5:1 not shown). The disappearance of the free DNA peaks requires titration of a 2:1 molar ratio of f-ImPyIm to (CGCGnmr)$_2$ (Figure 5A, top panel). Spectra of the 1:1 complex retain free DNA signals and do have distinct signals that correspond to the 2:1 complex. There are no peaks present in the 1:1 titration that are unique to this spectra. Thus, at the 1:1 ratio, the DNA exists both in an unbound state and as a 2:1 complex, which implies that there is no formation of a distinct (f-ImPyIm)$_1$−(CGCGnmr)$_2$ complex. This result affirms that f-ImPyIm binds with positive cooperativity to its cognate DNA that agrees well with prior SPR data (*27*) and the ITC data discussed above. In addition, the observation of five distinct and reasonable sharp peaks for the 2:1 complex indicates a well-structured and less dynamic DNA compared to the free DNA.

Titration of f-PyImIm into (CCGGnmr)$_2$ was also observed in a 90% H_2O/10% D_2O solution (Figure 5A, bottom panel). Prior to titration of f-PyImIm, (CCGGnmr)$_2$ has three peaks corresponding to the five possible imino peaks. The peak at 13.50 ppm is T_8, and the other two peaks at 12.59 and 12.43 ppm correspond to the four guanosine imino protons. Upon addition of 2 molar ratio equivalents of f-PyImIm, the peak corresponding to T_8 has undergone a shift of less than 0.01 ppm. After saturation, only two peaks are visible for the four guanine imino protons (12.76 and 12.35 ppm). Integration of the three peaks indicates that the peak at 12.76 ppm corresponds to three of the imino protons and the peak at 12.35 ppm to one imino proton. Unlike formation of the (f-ImPyIm)$_2$−(CGCGnmr)$_2$ complex (Figure 5A, top panel), there is no upfield shift observed for the imino proton signal for CG core sequence to the region of 11.4 ± 0.4 ppm. Thus, this lack of an upfield shift provides strong evidence that a strong correlation exists between the upfield shifting of the central guanine imino proton signal and binding by a central pairing, such as -ImPy-. In contrast, weak central pairings (-PyIm-) do not exhibit this shift.

The imino ^{1}H NMR spectra for the 1:1 titration of f-PyImIm to (CCGGnmr)$_2$ show a mixture of 0:1 and 2:1 complexes (Figure 5A, bottom panel). No additional peaks are observed in the 1:1 ratio that do not correspond to the 0:1 or 2:1 spectra. f-PyImIm was also titrated at 0.5:1 and 1.5:1, and the spectra provide evidence of mixtures between 0:1 and 2:1 complexes (data not shown). These results provide evidence that f-PyImIm binds (CCGGnmr)$_2$ with positive cooperativity, which supports the data from the ITC data discussed above and SPR results (*26*). f-PyImIm was titrated to a ratio of 2.5:1, and no difference is observed

between the spectra of the 2:1 and 2.5:1 ratios (data not shown). Thus, the 0.5 molar ratio excess of f-PyImIm to $(CGCGnmr)_2$ is free in solution, and the compound forms a 2:1 complex with the DNA. Only three, instead of five, relatively broad distinct peaks are observed for the 2:1 f-PyImIm−$(CGCGnmr)_2$ complex, suggesting that within the NMR time scale this complex is less structured and probably more dynamic than the 2:1 f-ImPyIm−$(CGCGnmr)_2$ complex.

Direct Evidence of Cognate Sequence Binding by f-ImPyIm and f-PyImIm. Nonexchangeable Proton NMR Studies. Two-dimensional COSY proton NMR experiments were performed in 99.96% D_2O with the free DNA sequences, $(CGCGnmr)_2$ and $(CCGGnmr)_2$, and with the DNA sequences in 2:1 complexes with f-ImPyIm and f-PyImIm, respectively (Figure 5B). There is a strong correlation between H5 and H6 within cytosine bases, and this correlation is easily monitored by 2D COSY NMR. The behaviors of these cytosine residues were monitored in identifying the f-ImPyIm and f-PyImIm binding sites.

The H6−H5 cross-peaks for C_1 and C_2 of $(CGCGnmr)_2$ undergo negligible shifts upon binding of the DNA by f-ImPyIm (Figure 5B, left panel). In the absence of f-ImPyIm, the C_4 and C_6 cross-peaks overlap, but C_4 is located at 7.28/5.35 ppm and C_6 at 7.32/5.37 ppm (Table S1A of the Supporting Information). When f-ImPyIm binds, the cross-peaks corresponding to C_4 and C_6 are oppositely shifted (dashed arrow, Figure 5B, left panel). These shifts are due to changes in the local environment of these bases. Thus, the sequence $C_4G_5C_6G_7$ is highly affected by binding of f-ImPyIm, whereas the termini of the DNA (C_1 and C_2 and their complementary bases, G_9 and G_{10}) are minimally affected by binding of f-ImPyIm. Thus, f-ImPyIm binds in the center of the CGCGnmr DNA. The C_6 cross-peak corresponds to the two base pairs that constitute the predicted binding site for the -ImPy- central pairing. Interestingly, the C_6 cross-peak undergoes a larger shift than C_4, which provides further evidence that links the importance of the $G_5{\cdot}C_6$ base pairs and the -ImPy- central pairing.

Binding of f-PyImIm to $(CCGGnmr)_2$ results in minimal shifts for the cross-peaks corresponding to the terminal cytosines (C_1 and C_2) and large shifts for C_4 and C_5 (Figure 5B, right panel; Table S1B of the Supporting Information). These results indicate that f-PyImIm does bind its proposed target sequence, $C_4C_5G_6G_7$. Changes in the C_1 and C_2 cross-peaks of $(CCGGnmr)_2$ upon binding f-PyImIm are slightly larger than that observed for C_1 and C_2 of $(CGCGnmr)_2$ and f-ImPyIm. This indicates that the local environments for these cytosines changed significantly upon binding of f-PyImIm compared to binding of f-ImPyIm to their respective cognate DNA sequences. This observation reflects another difference between f-ImPyIm and f-PyImIm. The CCGGnmr duplex DNA may require larger structural changes to accommodate f-PyImIm, whereas $(CGCGnmr)_2$ could be better structured to accept binding of f-ImPyIm.

Minor Groove Binding of f-ImPyIm and Retention of B-Form DNA Structure. The solutions of 2:1 f-ImPyIm−$(CGCGnmr)_2$ and f-PyImIm−$(CCGGnmr)_2$ complexes were subjected to NOE difference studies. Figure 5C shows sample NOE difference spectra, the control NOE of the 2:1 f-ImPyIm−$(CGCGnmr)_2$ complex, and ^{1}H NMR of the unbound CGCGnmr DNA. The control NOE spectrum exhibits relatively well resolved peaks for the 2:1 complex, indicating that the structure is well-defined. In contrast, the NMR spectrum for the 2:1 f-PyImIm−$(CCGGnmr)_2$ complex (data not shown) is less resolved and the peaks are broad, suggesting that the complex is less defined and undergoes more dynamic exchange processes, within the NMR time scale, than the 2:1 f-ImPyIm−$(CGCGnmr)_2$ complex.

In the top two panels of Figure 5C, the irradiated protons are circled in the corresponding f-ImPyIm or DNA structures. In the top panel of Figure 5C, the protons corresponding to the dimethylamine group (3.04 ppm) were irradiated, resulting in excitation of the peak corresponding to A_3H2. The A_3H2 proton and ligand signals were identified by their slow spin−lattice relaxation times as indicated by a nonselective inversion−recovery (T_1) experiment described in Experimental Procedures (data not shown) (*34*, *60*). Irradiation of the dimethylamine tail also enhances signals for the CH_2 protons α and β to the dimethyl amine tail. In the second panel of Figure 5C, irradiation of A_3H8 does not enhance any peaks that correspond to f-ImPyIm, and these results indicate that f-ImPyIm does not bind in the major groove. Irradiation of A_3H2 results in the simultaneous irradiation of overlapping signals (Table S1 of the Supporting Information), which does not result in the clear enhancement of NOE signals tied to A_3H2. Thus, the irradiation of A_3H8 shows that f-ImPyIm is bound in the minor groove and not the major groove. Irradiation of A_3H8 enhanced the A_3H1', A_3H2', C_2-$H1'$, and C_4H5 peaks. These NOE signals indicate that A_3 is in the *C2′-endo* conformation, which is indicative that complexed DNA is in the B-conformation (*60*, *61*).

Control Experiments for NMR Studies. Fresh solutions of $(CGCGnmr)_2$ and $(CCGGnmr)_2$ were studied by circular dichroism (CD), and f-ImPyIm and f-PyImIm were titrated into their respective cognate DNAs (Figure S1A of the Supporting Information). Induced peaks were observed upon addition of the triamides, which indicate binding in the minor groove. The DNAs were saturated at 2:1 triamide:duplex DNA ratios (Figure S1B of the Supporting Information). CD spectra of 2:1 f-ImPyIm−$(CGCGnmr)_2$ and f-PyImIm−$(CCGGnmr)_2$ samples from the concluded NMR experiments are indistinguishable from the titrated, non-NMR samples. In addition, the DNA alone exhibits a very broad melting profile with no distinct peaks, but addition of fresh f-ImPyIm (2:1) and f-PyImIm (2.5:1, which corresponds to the 0.5 molar ratio excess used in the NMR experiments) results in sharp melting transitions at 42 and 55 °C, respectively. The transition midpoint for the thermal denaturation of the NMR samples was indistinguishable from those of the fresh samples (Figure S1C of the Supporting Information).

(f-ImPyIm)$_2$−(CGCGnmr)$_2$ Model. A molecular model and molecular dynamic simulations were performed on the (f-ImPyIm)$_2$−$(CGCGnmr)_2$ complex. The f-ImPyIm molecules were docked into the minor groove at the 5′-ACGCGT site of the NMR-derived structure using SYBYL 7.0, and the CH_2 α−A_3H2 distance (3.3 ± 0.6 Å) was used to constrain the polyamide−DNA binding site. The structure of the complex was optimized using molecular mechanics with six layers of water to eliminate unfavorable van der Waals contacts. The resulting structure was thermodynamically optimized by AMBER 7.0, and the dynamics simulation was conducted for a 3 ns production time. Comparison of space filling models of the free DNA (Figure 6A) and the bound

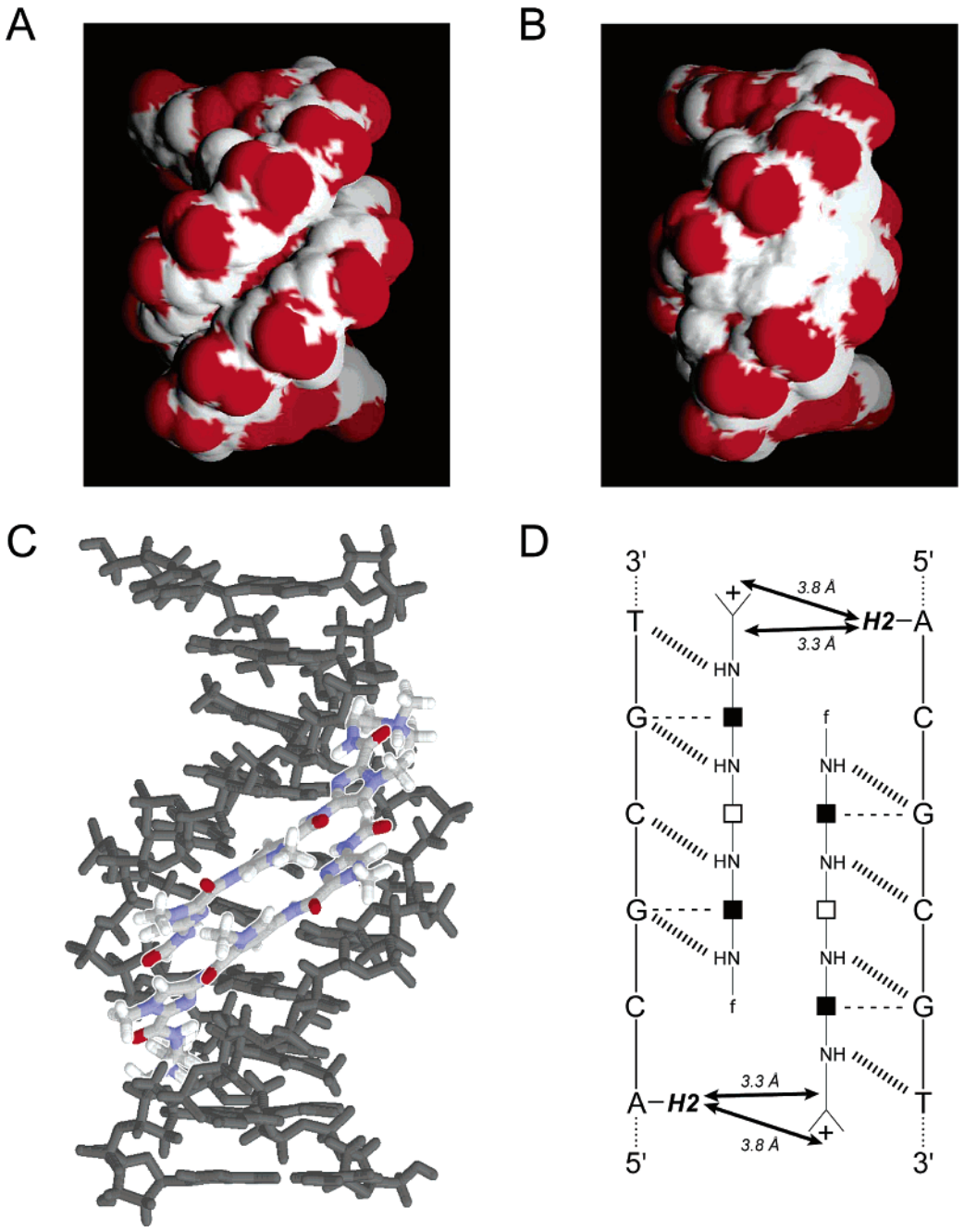

FIGURE 6: Three-dimensional models of the 5′-CCACGCGTGG homoduplex (A) and the DNA in complex with the f-ImPyIm dimer (B). The solvent accessible surface of the free DNA and the $(f\text{-ImPyIm})_2$−DNA complex generated using a probe radius of 1.7683 Å is shown (*41*). Red and white indicate polar and nonpolar surfaces, respectively. (C) Wireframe model of the complex with f-ImPyIm shown in CPK coloring with the DNA colored gray. (D) Schematic representation and alignment of the f-ImPyIm dimer to 5′-ACGCGT visually generated from the three-dimensional complex in panel C. Arrows indicate distances determined from NOE difference spectroscopy.

DNA (Figure 6B) shows that f-ImPyIm fits snuggly into the minor groove of the DNA, but the backbone of the f-ImPyIm dimer is still exposed to the solvent. The wireframe model in Figure 6C shows the slight twisting of f-ImPyIm away from a perfectly planar molecule. As seen with other polyamide dimers (*10*, *55*), the heterocyclic dimers are slightly offset to better align the amide hydrogens and the imidazole N3 groups into the minor groove of the DNA. Thus, as would be expected from the strong binding affinity and favorable thermodynamic contributions discussed above, f-ImPyIm fits very well when docked into the CGCGnmr duplex model.

CONCLUSIONS

The elucidation of the central pairing rules (-ImPy- > -PyPy- > -PyIm- ∼ -ImIm-) resulted in a better understanding of the relationship between the strength of binding and the heterocyclic sequence of triamides (*27*). These rules extend the essential pairing rules (*13*−*20*) and allow for better selection and design of target DNA sequences. These rules allow for the selection of the ideal DNA sequence within or immediately adjacent to the target sequence and then synthesis of the best triamide for binding this sequence. There are almost certainly similar rules for more complex polyamides, such as tetraheterocyclic and hairpin polyamides,

and more work needs to be done to identify these rules. Determination of these rules, and the biophysical interactions that govern these rules, will provide considerable guidance to the design of novel polyamides and related small molecules. Our hope is that these rules will reduce or even eliminate the wasteful synthesis of compounds that exhibit weak DNA binding even to their cognate sequences. The work presented herein is the first biophysical description of the central pairing rules.

The differences between the central pairings -ImPy- and -PyIm- are now better understood in the DNase I footprinting, thermodynamic, and structural studies of f-ImPyIm and f-PyImIm. Perhaps the most striking difference between these two compounds is the heat capacity. f-ImPyIm exhibits a significant, negative heat capacity (-142 cal mol^{-1} K^{-1}), whereas f-PyImIm has a nominal, positive heat capacity (46 cal mol^{-1} K^{-1}). Most compounds exhibit negative ΔC_p values for binding to DNA (*42*, *52*, *53*, *58*, *59*). Contributions of the thermodynamic factors (for example, hydration, protonation, and bond vibrations) that contribute to heat capacity measurements are very difficult to experimentally determine. Often ΔC_p is approximated by theoretically calculating the burial of hydrophobic solvent accessible surface area upon complex formation. However, the ΔC_p from this theoretical method and the experimentally determined ΔC_p do not match. As both calorimetric and theoretical experimental techniques are improved, it is clear that considering only the burial of hydrophobic residues is not a good estimation of ΔC_p. The structural similarity between f-ImPyIm and f-PyImIm make it hard to believe that f-ImPyIm binds through burial of hydrophobic regions and the other through a very different mechanism. It is more likely that the difference in ΔC_p between these two molecules is due to other thermodynamic components, such as water uptake versus release or water- and salt-mediated interactions between the polyamide and the DNA.

Finally, the mechanism by which polyamides bind as dimers to DNA is not totally clear. However, the data suggest that the process might occur via a two-site model, in which two molecules bind independently to the DNA to form the 2:1 complex.

ACKNOWLEDGMENT

We thank David Wilson and Binh Nguyen (Georgia State University, Atlanta, GA) and Jeff Petty (Furman University) for helpful discussions. Peter Spielmann (University of Kentucky, Lexington, KY) is acknowledged for information about the CAACGCGTTG duplex.

SUPPORTING INFORMATION AVAILABLE

CD and T_M control experiments for the NMR experiments (Figure S1), assignment of DNA base and polyamine protons (Table S1), and solvent accessible surface areas calculated with GRASP (Table S2). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wemmer, D. E. (2000) Designed sequence-specific minor groove ligands, *Annu. Rev. Biophys. Biomol. Struct. 29*, 439−461.
2. Bailly, C., and Chaires, J. B. (1998) Sequence-specific DNA minor groove binders. Design and synthesis of netropsin and distamycin analogues, *Bioconjugate Chem. 9*, 513−538.
3. Dervan, P. B. (2001) Molecular recognition of DNA by small molecules, *Bioorg. Med. Chem. 9*, 2215−2235.
4. Uil, T. G., Haisma, H. J., and Rots, M. G. (2003) Therapeutic modulation of endogenous gene function by agents with designed DNA-sequence specificities, *Nucleic Acids Res. 31*, 6064−6078.
5. Dervan, P. B., and Edelson, B. S. (2003) Recognition of the DNA minor groove by pyrrole-imidazole polyamides, *Curr. Opin. Struct. Biol. 13*, 284−299.
6. Melander, C., Burnett, R., and Gottesfeld, J. M. (2004) Regulation of gene expression with pyrrole-imidazole polyamides, *J. Biotechnol. 112*, 195−220.
7. Philips, B. J., Chang, A. Y., Dervan, P. B., and Beerman, T. A. (2005) DNA damage effects of a polyamide-CBI conjugate in SV40 virions, *Mol. Pharmacol. 67*, 877−882.
8. Fechter, E. J., Olenyuk, B., and Dervan, P. B. (2005) Sequence-specific fluorescence detection of DNA by polyamide-thiazole orange conjugates, *J. Am. Chem. Soc. 127*, 16685−16691.
9. Yang, X.-L., and Wang, A. H.-J. (1999) Structural studies of atom-specific anticancer drugs acting on DNA, *Pharmacol. Ther. 83*, 181−215.
10. Yang, X.-L., Hubbard, R. B., Lee, M., Tao, Z.-F., Sugiyama, H., and Wang, A. H.-J. (1999) Imidazole-imidazole pair as a minor groove recognition motif for T:G mismatched base pairs, *Nucleic Acids Res. 27*, 4183−4190.
11. Edayathumangalam, R. S., Weyermann, P., Gottesfeld, J. M., Dervan, P. B., and Luger, K. (2004) Molecular recognition of the nucleosomal "supergroove", *Proc. Natl. Acad. Sci. U.S.A. 101*, 6864−6869.
12. Edayathumangalam, R. S., Weyermann, P., Dervan, P. B., Gottesfeld, J. M., and Luger, K. (2005) Nucleosomes in solution exist as a mixture of twist-defect states, *J. Mol. Biol. 345*, 103−114.
13. Pelton, J. G., and Wemmer, D. E. (1989) Structural characterization of a 2:1 distamycin A·d(CGCAAATTGGC) complex by two-dimensional NMR, *Proc. Natl. Acad. Sci. U.S.A. 86*, 5723−5727.
14. Pelton, J. G., and Wemmer, D. E. (1990) Binding modes of distamycin A with d(CGCAAATTTGCG)$_2$ determined by two-dimensional NMR, *J. Am. Chem. Soc. 112*, 1393−1399.
15. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P., and Dickerson, R. E. (1985) The molecular origin of DNA-drug specificity in netropsin and distamycin, *Proc. Natl. Acad. Sci. U.S.A. 82*, 1376−1380.
16. Mrksich, M., Wade, W. S., Dwyer, T. J., Geierstanger, B. H., Wemmer, D. E., and Dervan, P. B. (1992) Antiparallel side-by-side dimeric motif for sequence-specific recognition in the minor groove of DNA by the designed peptide 1-methylimidazole-2-carboxamide netropsin, *Proc. Natl. Acad. Sci. U.S.A. 89*, 7586−7590.
17. Dwyer, T. J., Geierstanger, B. H., Bathini, Y., Lown, J. W., and Wemmer, D. E. (1992) Design and binding of a distamycin A analog to d(CGCAAGTTGGC)·d(GCCAACTTGCG): Synthesis, NMR studies, and implications for the design of sequence-specific minor groove binding oligopeptides, *J. Am. Chem. Soc. 114*, 5911−5919.
18. Wade, W. S., Mrksich, M., and Dervan, P. B. (1992) Design of peptides that bind in the minor groove of DNA at 5′-(A,T)G-(A,T)C(A,T)-3′ sequences by a dimeric side-by-side motif, *J. Am. Chem. Soc. 114*, 8783−8794.
19. Lown, J. W., Krowicki, K., Bhat, U. G., Skorobogaty, A., Ward, B., and Dabrowiak, J. C. (1986) Molecular recognition between oligopeptides and nucleic acids: Novel imidazole-containing oligopeptides related to netropsin that exhibit altered DNA sequence specificity, *Biochemistry 25*, 7408−7416.
20. Kissinger, K., Krowicki, K., Dabrowiak, J. C., and Lown, J. W. (1987) Molecular recognition between oligopeptides and nucleic acids. Monocationic imidazole lexitropsins that display enhanced GC sequence dependent DNA binding, *Biochemistry 26*, 5590−5595.
21. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. (1998) Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands, *Nature 391*, 468−471.
22. Marques, M. A., Doss, R. M., Foister, S., and Dervan, P. B. (2004) Expanding the repertoire of heterocycle ring pairs for programmable minor groove DNA recognition, *J. Am. Chem. Soc. 126*, 10339−10349.
23. Renneberg, D., and Dervan, P. B. (2003) Imidazopyridine/Pyrrole and Hydroxybenzimidazole/Pyrrole Pairs for DNA Minor Groove Recognition, *J. Am. Chem. Soc. 125*, 5707−5716.

24. Wellenzohn, B., Loferer, M. J., Trieb, M., Rauch, C., Winger, R. H., Mayer, E., and Liedl, K. R. (2003) Hydration of hydroxypyrrole influences binding of ImHpPyPy-β-Dp polyamide to DNA, *J. Am. Chem. Soc. 125*, 1088−1095.
25. Melander, C., Herman, D. M., and Dervan, P. B. (2000) Discrimination of A/T sequences in the minor groove of DNA within a cyclic polyamide motif, *Chem.—Eur. J. 6*, 4487−4497.
26. Lacy, E. R., Cox, K. K., Wilson, W. D., and Lee, M. (2002) Recognition of T*G mismatched base pairs in DNA by stacked imidazole-containing polyamides: Surface plasmon resonance and circular dichroism studies, *Nucleic Acids Res. 30*, 1834−1841.
27. 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., and Lee, M. (2005) Extending the Language of DNA Molecular Recognition by Polyamides: Unexpected Influence of Imidazole and Pyrrole Arrangement on Binding Affinity and Specificity, *J. Am. Chem. Soc. 127*, 742−750.
28. Lacy, E. R., Le, N. M., Price, C. A., Lee, M., and Wilson, W. D. (2002) Influence of a Terminal Formamido Group on the Sequence Recognition of DNA by Polyamides, *J. Am. Chem. Soc. 124*, 2153−2163.
29. 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., and Lee, M. (2005) Molecular recognition of DNA base pairs by the formamido/pyrrole and formamido/imidazole pairings in stacked polyamides, *Nucleic Acids Res. 33*, 912−921.
30. Indyk, L., and Fisher, H. F. (1998) Theoretical aspects of isothermal titration calorimetry, *Methods Enzymol. 295*, 350−364.
31. Hore, P. J. (1983) Solvent suppression in Fourier-transform nuclear magnetic-resonance, *J. Magn. Reson. 55*, 283−300.
32. Cheatham, S. (1989) Nuclear magnetic resonance spectroscopy in biochemistry, *J. Chem. Educ. 66*, 111−117.
33. Isaacs, R. J., Rayens, W. S., and Spielmann, H. P. (2002) Structural differences in the NOE-derived structure of G-T mismatched DNA relative to normal DNA are correlated with differences in ^{13}C relaxation-based internal dynamics, *J. Mol. Biol. 319*, 191−207.
34. Patel, D. J., Kozlowski, S. A., and Bhatt, R. (1983) Sequence dependence of base-pair stacking in right-handed DNA in solution: Proton nuclear Overhauser effect NMR measurements, *Proc. Natl. Acad. Sci. U.S.A. 80*, 3908−3912.
35. Case, D. A., Cheatham, T. E., III, Darden, T., Gohlke, H., Luo, R., Merz, K. M., Jr., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *J. Comput. Chem. 26*, 1668−1688.
36. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys. 79*, 926−935.
37. Wellenzohn, B., Flader, W., Winger, R. H., Hallbrucker, A., Mayer, E., and Liedl, K. R. (2001) Complex of B-DNA with Polyamides Freezes DNA Backbone Flexibility, *J. Am. Chem. Soc. 123*, 5044−5049.
38. Wellenzohn, B., Flader, W., Winger, R. H., Hallbrucker, A., Mayer, E., and Liedl, K. R. (2001) Structural Flexibility of the $d(CCAGTACTGG)_2$ B-DNA Decamer and Its Complex with Two Polyamides, *J. Phys. Chem. B 105*, 3135−3142.
39. Hamelberg, D., Williams, L. D., and Wilson, W. D. (2001) Influence of the Dynamic Positions of Cations on the Structure of the DNA Minor Groove: Sequence-Dependent Effects, *J. Am. Chem. Soc. 123*, 7745−7755.
40. Nicholls, A., Sharp, K. A., and Honig, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−296.
41. 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., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
42. Nguyen, B., Stanek, J., and Wilson, W. D. (2006) Binding-linked protonation of a DNA minor-groove agent, *Biophys. J. 90*, 1319−1328.
43. Livingstone, J. R., Spolar, R. S., and Record, M. T., Jr. (1991) Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area, *Biochemistry 30*, 4237−4244.
44. Spolar, R. S., Ha, J.-H., and Record, M. T., Jr. (1989) Hydrophobic effect in protein folding and other noncovalent processes involving proteins, *Proc. Natl. Acad. Sci. U.S.A. 86*, 8382−8385.
45. Spolar, R. S., Livingstone, J. R., and Record, M. T., Jr. (1992) Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water, *Biochemistry 31*, 3947−3955.
46. Spolar, R. S., and Record, M. T., Jr. (1994) Coupling of local folding to site-specific binding of proteins to DNA, *Science 263*, 777−784.
47. Gomez, J., and Freire, E. (1995) Thermodynamic mapping of the inhibitor site of the aspartic protease endothiapepsin, *J. Mol. Biol. 252*, 337−350.
48. Gomez, J., Hilser, V. J., Xie, D., and Freire, E. (1995) The heat capacity of proteins, *Proteins 22*, 404−412.
49. Murphy, K. P., and Freire, E. (1992) Thermodynamics of structural stability and cooperative folding behavior in proteins, *Adv. Protein Chem. 43*, 313−361.
50. Habermann, S. M., and Murphy, K. P. (1996) Energetics of hydrogen bonding in proteins: A model compound study, *Protein Sci. 5* 1229−1239.
51. Murphy, K. P., Bhakuni, V., Xie, D., and Freire, E. (1992) Molecular basis of co-operativity in protein folding. III. Structural identification of cooperative folding units and folding intermediates, *J. Mol. Biol. 227*, 293−306.
52. Ren, J., Jenkins, T. C., and Chaires, J. B. (2000) Energetics of DNA intercalation reactions, *Biochemistry 39*, 8439−8447.
53. Mazur, S., Tanious, F. A., Ding, D., Kumar, A., Boykin, D. W., Simpson, I. J., Neidle, S., and Wilson, W. D. (2000) A thermodynamic and structural analysis of DNA minor-groove complex formation, *J. Mol. Biol. 300*, 321−337.
54. Rich, R. L., and Myszka, D. G. (2000) Advances in surface plasmon resonance biosensor analysis, *Curr. Opin. Biotechnol. 11*, 54−61.
55. Yang, X.-L., Kaenzig, C., Lee, M., and Wang, A. H.-J. (1999) Binding of AR-1-144, a tri-imidazole DNA minor groove binder, to CCGG sequence analyzed by NMR spectroscopy, *Eur. J. Biochem. 263*, 646−655.
56. Chaires, J. B. (2006) A thermodynamic signature for drug-DNA binding modes, *Arch. Biochem. Biophys. 453*, 24−29.
57. Chen, F.-M., and Sha, F. (1998) Circular Dichroic and Kinetic Differentiation of DNA Binding Modes of Distamycin, *Biochemistry 37*, 11143−11151.
58. Haq, I. (2002) Thermodynamics of drug-DNA interactions, *Arch. Biochem. Biophys. 403*, 1−15.
59. Lacy, E. R., Nguyen, B., Le, M., Cox, K. K., O'Hare, C., Hartley, J. A., Lee, M., and Wilson, W. D. (2004) Energetic basis for selective recognition of T*G mismatched base pairs in DNA by imidazole-rich polyamides, *Nucleic Acids Res. 32*, 2000−2007.
60. Lee, M., Krowicki, K., Hartley, J. A., Pon, R. T., and Lown, J. W. (1988) Molecular recognition between oligopeptides and nucleic acids: Influence of van der Waals contacts in determining the 3′-terminus of DNA sequences read by monocationic lexitropsins, *J. Am. Chem. Soc. 110*, 3641−3649.
61. Gronenborn, A. M., and Clore, G. M. (1985) Investigation of the solution structures of short nucleic acid fragments by means of nuclear overhauser enhancement measurements, *Prog. Nucl. Magn. Reson. Spectrosc. 17*, 1−32.

BI061245C