Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 8306–8311, August 1996
Biochemistry

Binding of a hairpin polyamide in the minor groove of DNA: Sequence-specific enthalpic discrimination

(polyamide-DNA binding affinity/isothermal calorimetry/2:1 pyrrole-imidazole DNA binding motif/polyamide-base hydrogen bonds/imidazole-guanine interaction)

DANIEL S. PILCH*†, NATAŠA POKLAR*, CRAIG A. GELFAND*, SCOTT M. LAW*, KENNETH J. BRESLAUER*†‡, ELDON E. BAIRD§, AND PETER B. DERVAN§

*Department of Chemistry, Rutgers–The State University of New Jersey, New Brunswick, NJ 08903; †The Cancer Institute of New Jersey, New Brunswick, NJ 08901; and §Arnold and Mabel Beckman Laboratories of Chemical Synthesis, California Institute of Technology, Pasadena, CA 91125

Contributed by Peter B. Dervan, May 13, 1996

ABSTRACT **Hairpin polyamides are synthetic ligands for sequence-specific recognition in the minor groove of double-helical DNA. A thermodynamic characterization of the DNA-binding properties exhibited by a six-ring hairpin polyamide, ImPyPy-γ-PyPyPy-β-Dp (where Im = imidazole, Py = pyrrole, γ = γ-aminobutyric acid, β = β-alanine, and Dp = dimethylaminopropylamide), reveals an ≈1-2 kcal/mol greater affinity for the designated match site, 5′-TGTTA-3′, relative to the single base pair mismatch sites, 5′-TGGTA-3′ and 5′-TATTA-3′. The enthalpy and entropy data at 20°C reveal this sequence specificity to be entirely enthalpic in origin. Correlations between the thermodynamic driving forces underlying the sequence specificity exhibited by ImPyPy-γ-PyPyPy-β-Dp and the structural properties of the heterodimeric complex of PyPyPy and ImPyPy bound to the minor groove of DNA provide insight into the molecular forces that govern the affinity and specificity of pyrrole-imidazole polyamides.**

Pyrrole–imidazole (Py–Im) polyamide-DNA complexes (1–13) coupled to solid phase synthetic methods (14) provide a paradigm for the design of artificial molecules for the digital readout of double-helical DNA. Polyamides containing *N*-methylimidazole and *N*-methylpyrrole amino acids can be combined in antiparallel side-by-side dimeric complexes with the minor groove of DNA (1–3). The DNA sequence specificity of these small molecules can be controlled by the linear sequence of pyrrole and imidazole amino acids (1–3). An imidazole ring on one ligand complemented by a pyrrolecarboxamide ring on the second ligand recognizes a G·C base pair, while a pyrrolecarboxamide/imidazole combination targets a C·G base pair (1–3). A pyrrolecarboxamide/pyrrolecarboxamide pair is degenerate for A·T or T·A base pairs (1–5).

Covalently linking polyamide heterodimers and homodimers within the 2:1 motif has led to designed ligands with both increased affinity and specificity (10–13). A simple hairpin polyamide motif with γ-aminobutyric acid (γ) serving as a "turn monomer" provides a synthetically accessible method of linking polyamide units within the 2:1 motif (12–13). The six-ring polyamide ImPyPy-γ-PyPyPy-Dp was found to bind the 5-bp 5′-TGTTA-3′ site with high specificity and an ≈300-fold binding enhancement over the individual unlinked polyamides ImPyPy and PyPyPy (6, 7, 12). Addition of a C-terminal β-alanine residue recently has been found to enhance both the DNA binding affinity and sequence specificity of the hairpin polyamide, ImPyPy-γ-PyPyPy-β-Dp, relative to ImPyPy-γ-PyPyPy-Dp (13) (Fig. 1).

Footprinting and affinity cleaving studies have provided information regarding the orientations and specific affinities of dimeric polyamide complexes (1, 3, 7, 11–13). However, comparatively little is known about the thermodynamic properties that govern the binding events. This limits our understanding of the molecular forces that control the affinity and specificity of binding. We have used a combination of calorimetric and spectroscopic techniques to characterize the binding of the hairpin polyamide, ImPyPy-γ-PyPyPy-β-Dp, to three 11-mer DNA duplexes, whose base sequences are presented in Fig. 1. The central five base pair sequence of one of the three duplexes (duplex 1) is 5′-TGTTA-3′, the match site as defined by the pairing rules and footprinting studies (12, 13). The other two duplexes contain single base pair changes to produce 5′-TGGTA-3′ (duplex 2) and 5′-TATTA-3′ (duplex 3), which we designate as mismatch sites. Our studies at 20°C reveal that ImPyPy-γ-PyPyPy-β-Dp exhibits an ≈1-2 kcal/mol greater affinity for the 5′-TGTTA-3′ target site than for either the 5′-TATTA-3′ or the 5′-TGGTA-3′ single base pair mismatch site, with this enhanced affinity being entirely enthalpic in origin. We discuss possible correlations between this enthalpically driven binding preference and the NMR-derived structural properties of an unlinked heterodimeric 2:1 polyamide–DNA complex (8).

MATERIALS AND METHODS

Oligonucleotide Synthesis and Characterization. Oligomers were synthesized on a BioSearch 8600 synthesizer by standard cyanoethyl-phosphoramidite chemistry, followed by purification using reverse-phase high-pressure liquid chromatography (HPLC). Molar extinction coefficients (ε) for the single-stranded oligomers were determined by phosphate analysis (15). The following ε values [in units of (mol strand)/liter)$^{-1}$·cm^{-1}] at 260 nm and 25°C were so obtained: 96,900 for d(CATTGTTAGAC); 97,400 for d(GTCTAACAATG); 94,300 for d(CATTGGTAGAC); 91,500 for d(GTCTACCAATG); 92,000 for d(CATTATTAGAC); and 104,700 for d(GTCTAATAATG).

Hairpin Polyamide Synthesis and Characterization. The polyamide, ImPyPy-γ-PyPyPy-β-Dp (Fig. 2), was prepared by machine-assisted solid-phase protocols and characterized by a combination of ^{1}H NMR, analytical HPLC, and matrix-assisted laser desorption ionization–time of flight (MALDI–TOF) mass spectrometry, with details described elsewhere (14).

Buffer Conditions. All spectroscopic and calorimetric experiments were conducted in 10 mM sodium cacodylate (pH 6.9), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. These buffer conditions were chosen to match as closely as possible

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: Im, imidazole; Py, pyrrole; β, β-alanine; γ, γ-aminobutyric acid; Dp, dimethylaminopropylamide; CD, circular dichroism; HPLC, high-pressure liquid chromatography; DSC, differential scanning calorimetry; T_m, melting temperature; r_{Dup}, total-ligand-to-duplex ratio.

‡To whom reprint requests should be addressed.

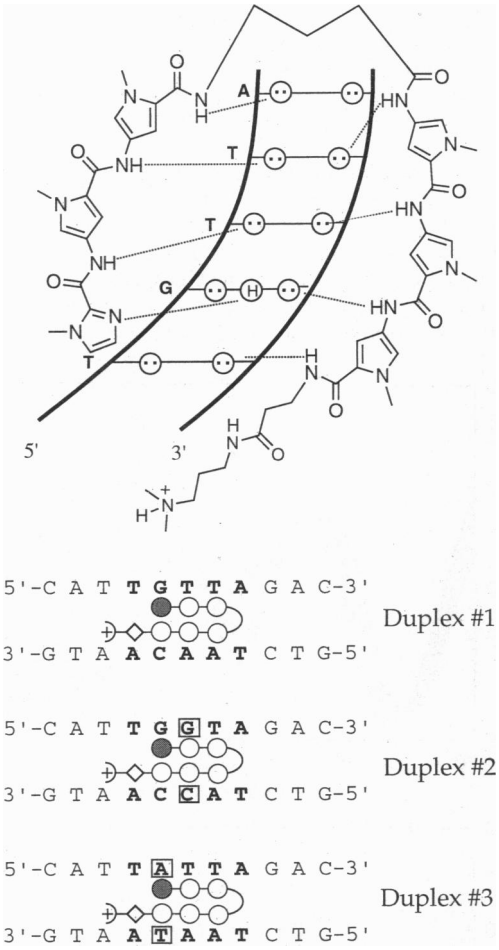
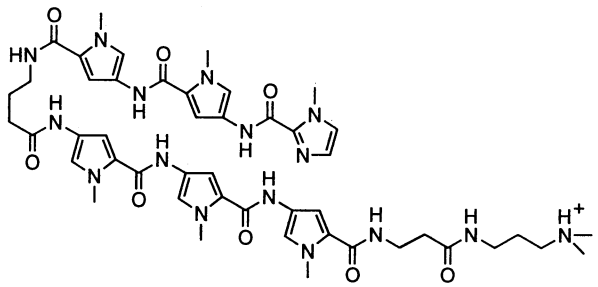


FIG. 1. (*Upper*) Binding model for the complex formed between ImPyPy-γ-PyPyPy-β-Dp and the 5′-TGTTA-3′ match site of duplex 1. Circles with dots represent lone electron pairs of either N3 of adenine, O2 of thymine, or O2 of cytosine. Circles containing an H represent the 2-amino hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. (*Lower*) Base sequences for the three 11-mer DNA duplexes used in this study (denoted as duplex 1, 2, and 3). Schematic binding model of putative complexes between the hairpin polyamide and each duplex. The imidazole and pyrrole rings are represented as shaded and unshaded spheres, respectively, while the β-alanine residue is represented as an unshaded diamond. The boxes indicate the base pairs in duplexes 2 and 3 that have been changed relative to the match duplex 1. The central five base pair binding sites of each duplex are presented in boldface type.

those used by Mrksich *et al.* (12) in their footprinting studies on the hairpin polyamide, ImPyPy-γ-PyPyPy-Dp. We used 10 mM sodium cacodylate in place of the 10 mM Tris·HCl



ImPyPy-γ-PyPyPy-β-Dp

FIG. 2. Structure of the hairpin polyamide, ImPyPy-γ-PyPyPy-β-Dp.

employed by Mrksich *et al.* (12), since the large temperature dependence of the pK_a for Tris·HCl ($-0.031\ \Delta pK_a/°C$) makes it poorly suited for thermal denaturation experiments. Significantly, however, control circular dichroism (CD) experiments in 10 mM sodium cacodylate were virtually identical to those in 10 mM Tris·HCl.

UV Absorption Spectrophotometry. Absorbance versus temperature profiles were measured at 260 nm on a Perkin–Elmer model λ4C spectrophotometer equipped with a thermoelectrically controlled cell holder and a cell path length of 1 cm. The heating rate in all experiments was 0.5°C/min. For each optically detected transition, the melting temperature (T_m) was determined as described (16, 17). The DNA concentration was 5 μM in duplex, while the ImPyPy-γ-PyPyPy-β-Dp concentration ranged from 0 to 5 μM.

CD Spectropolarimetry. All CD measurements were performed on an AVIV model 60DS spectropolarimeter (AVIV Associates; Lakewood, NJ) equipped with a thermoelectrically controlled cell holder and a cell path length of 1 cm. Isothermal ImPyPy-γ-PyPyPy-β-Dp titrations were performed at 20°C by incrementally adding 5 to 20 μl aliquots of 250–300 μM ImPyPy-γ-PyPyPy-β-Dp into a 2 ml solution of 5 μM duplex. After each addition, the CD spectrum was recorded from 220 to 380 nm, with an averaging time of 3 sec. The final CD spectra were normalized to reflect equimolar concentrations of duplex.

Isothermal Stopped-Flow Mixing Microcalorimetry. Isothermal calorimetric measurements were performed at 20°C using an all tantalum, differential, stopped-flow, heat conduction microcalorimeter (model DSFC-100, Commonwealth Technology, Alexandria, VA), developed by Mudd and Berger (18, 19). In a typical experiment, the reaction was initiated by a microprocessor-controlled stepping motor that activates a syringe drive that delivers, within 0.6 sec, 80 μl of each reagent (25 μM in both duplex and polyamide) into tantalum mixing chambers, with distilled water being used in the reference mixing chamber. A delay of 200 sec was used between each injection/reaction. Each reaction generated a heat burst curve (μjoule/sec versus sec), with the area under the curve being determined by integration to obtain the heat for that reaction, which ranged from 81 to 98 μjoules compared with polyamide dilution heats of 38 μjoules. The calorimeter was calibrated chemically by measuring the heat associated with a 1:2 dilution of 10 mM NaCl (20, 21).

Differential Scanning Calorimetry (DSC). The excess heat capacity (ΔC_p) versus temperature (T) profiles for the thermally induced transitions of the three ligand-free DNA duplexes were measured using a prototype model 5100 Nano calorimeter (Calorimetry Science, Provo, UT). In these experiments, the heating rate was 60°C/hr. Transition enthalpies ($\Delta H_{W\text{-}C}$) were calculated from the areas under the heat capacity curves using the ORIGIN version 1.16 software (MicroCal, Northampton, MA). The DNA solutions were 75 μM in duplex.

RESULTS AND DISCUSSION

The Hairpin Polyamide Binds to and Enhances the Thermal Stabilities of Each DNA Duplex in a Manner That Is Sensitive to Single Base Pair Changes in the Target Sequence. UV melting experiments were conducted in the absence and presence of ligand to assess the impact, if any, of ImPyPy-γ-PyPyPy-β-Dp on the thermal stabilities of the three 11-mer DNA duplexes studied here. The resulting melting profiles are shown in Fig. 3. As the total-ligand-to-duplex ratio (r_{Dup}) increases from 0 to 1.0, the thermal stabilities of all three host duplexes increase concomitantly. Higher polyamide-to-duplex ratios do not result in further increases in the T_m of either duplex 1 or duplex 2, while inducing only marginal increases in the T_m of duplex 3 (data not shown). This observation is

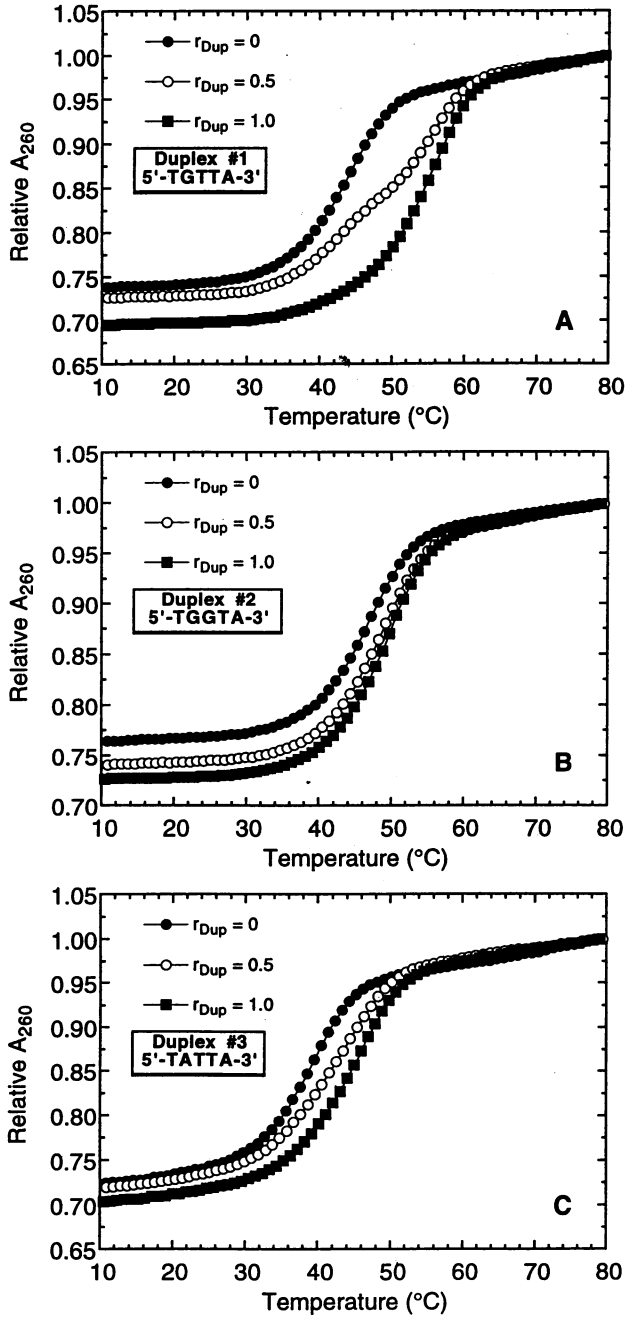


FIG. 3. UV melting profiles at 260 nm for duplexes 1 (*A*), 2 (*B*), and 3 (*C*) and their ImPyPy-γ-PyPyPy-β-Dp complexes at the indicated r_{Dup} values. Solution conditions are 10 mM sodium cacodylate (pH 6.9), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. For clarity of presentation, the melting curves for a given duplex and its ImPyPy-γ-PyPyPy-β-Dp complexes are normalized so as to produce identical absorbances at 80°C. As expected (22, 23), at r_{Dup} values below saturation, the melting curves of the complexes are not monophasic.

suggestive of secondary binding to duplex 3 at high polyamide concentrations. In this work, we will focus exclusively on the one-to-one complex, which is the one observed in footprinting (12, 13) and NMR studies (8, 10).

The polyamide-induced changes in duplex thermal stability noted above are consistent with ImPyPy-γ-PyPyPy-β-Dp binding to each duplex, with a preference for the duplex versus single-stranded state (22, 24, 25). Further inspection of Fig. 3 reveals the extent of ImPyPy-γ-PyPyPy-β-Dp-induced enhancement in duplex thermal stability to follow the hierarchy: duplex 1 > duplex 3 > duplex 2. Specifically, at a r_{Dup} ratio of 1.0, hairpin polyamide binding increases the thermal stabilities of duplex 1 (Fig. 3*A*), duplex 3 (Fig. 3*C*), and duplex 2 (Fig. 3*B*) by approximately 11, 6, and 2°C, respectively. Thus, as measured by differences in ΔT_m, the hairpin polyamide is able to distinguish between duplex targets that differ by only a single base pair.

DNA Duplex Binding Induces Chirality in the Hairpin Polyamide. In addition to the UV thermal denaturation studies described above, CD spectropolarimetry provides a second means for detecting and characterizing the DNA binding of the hairpin polyamide. Fig. 4 shows the CD spectra from 220 to 380 nm obtained by incremental titration of ImPyPy-γ-PyPyPy-β-Dp into a solution of either duplex 1 (Fig. 4*A*), duplex 2 (Fig.

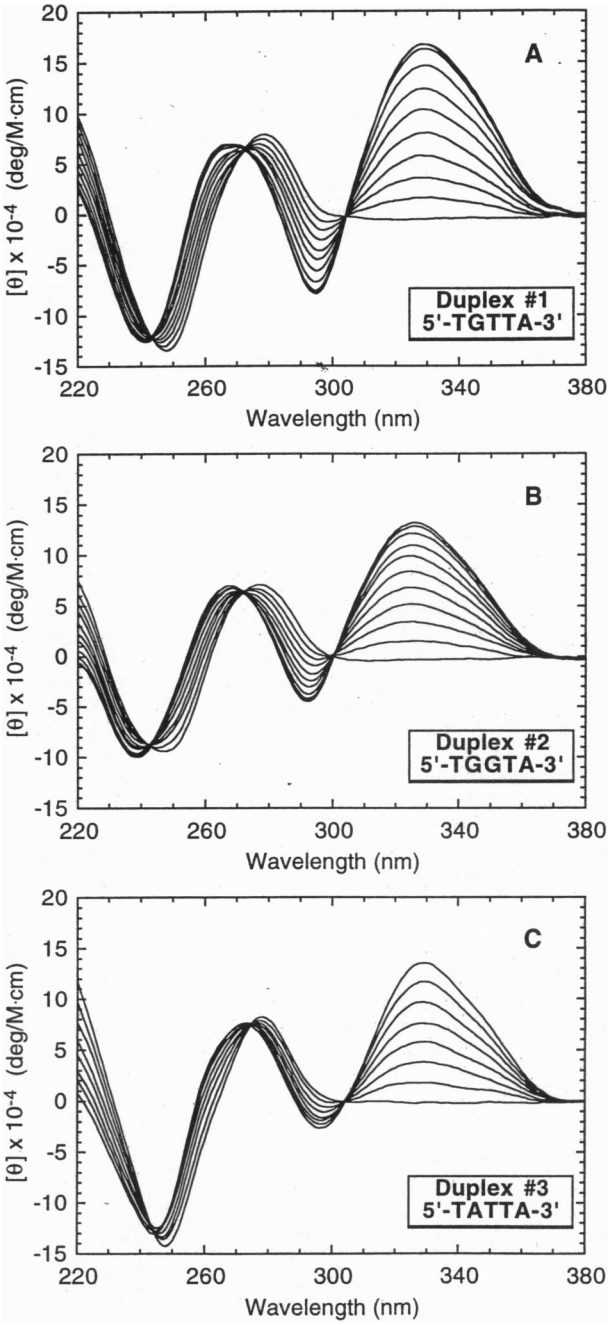


FIG. 4. CD titrations at 20°C of either duplex 1 (*A*), duplex 2 (*B*), or duplex 3 (*C*) with ImPyPy-γ-PyPyPy-β-Dp. From bottom to top at 325 nm, the CD spectra correspond to r_{Dup} values ranging from 0 to 1.4. Solution conditions are as described in the legend to Fig. 3. Molar ellipticities, [θ], are in units of deg/M·cm, where M refers to moles of DNA strand per liter.

4*B*), or duplex 3 (Fig. 4*C*). Neither free ImPyPy-γ-PyPyPy-β-Dp (spectrum not shown) nor any of the ligand-free duplexes exhibit CD signals between 300 and 380 nm. However, substantial CD signals arise in this wavelength range upon addition of ImPyPy-γ-PyPyPy-β-Dp to a solution of any one of the three duplexes (Fig. 4). These induced CD signals are indicative of interactions between ImPyPy-γ-PyPyPy-β-Dp, and each of the host DNA duplexes and can be used to detect and to monitor *CD-active* DNA binding mode(s).

Inspection of Fig. 4 *A–C* reveals that the magnitudes of the induced CD signals differ in a manner that depends on the host duplex. These differences in CD signal suggest that ImPyPy-γ-PyPyPy-β-Dp adopts different structural/electronic properties when bound to each duplex, a reasonable expectation given the differences in the binding sites. Note that the magnitude of the induced CD signal ($\Delta[\theta]$) follows a similar hierarchy to that defined above based on our ΔT_m data; namely, duplex 1 > duplex 3 ≳ duplex 2. This correlation between $\Delta[\theta]$ and ΔT_m-based trends also extends to the hierarchy of the ligand binding constants at 20°C (K_{20}), which are presented in a later section. This concordance between the $\Delta[\theta]$, ΔT_m, and K_{20} data suggests that the duplex binding strength of ImPyPy-γ-PyPyPy-β-Dp is correlated with its induced chirality and its ability to thermally stabilize the host duplex.

Hairpin Polyamide Binding to the 5′-TGTTA-3′ Match Site Is Enthalpically More Favorable Than Binding to Either the 5′-TGGTA-3′ or 5′-TATTA-3′ Mismatch Site. Isothermal, stopped-flow mixing calorimetry was used to measure the binding enthalpies (ΔH_b) for ImPyPy-γ-PyPyPy-β-Dp complexation with the three 11-mer DNA duplexes studied here. The resulting ΔH_b values are listed in Table 1. Inspection of these data reveals the enthalpy for hairpin polyamide binding to duplex 1, which contains the primary 5′-TGTTA-3′ match site as defined by footprinting (12), is −6.7 kcal/mol. By contrast, the enthalpies for hairpin polyamide binding to the 5′-TGGTA-3′ and 5′-TATTA-3′ mismatch sites are only −4.6 kcal/mol and −4.4 kcal/mol, respectively, values that are essentially indistinguishable. Thus, the enthalpy data are consistent with the observed hairpin polyamide binding preference for the 5′-TGTTA-3′ site (12).

The Hairpin Polyamide Binds to the 5′-TGTTA-3′ Match Site with a Greater Affinity Than it Binds to the 5′-TGGTA-3′ and 5′-TATTA-3′ Single Mismatch Sites. We used the ΔT_m approach described below to assess, by a single method, the relative strength of polyamide binding to all three duplexes, since the magnitude of the binding to duplex 1 precluded a Scatchard analysis of the optical data. Significantly, both the ΔT_m and Scatchard methods yield similar binding constants for duplex 2, thereby validating our use of the ΔT_m method for the systems studied here. This validation is consistent with previous reports in which the ΔT_m method was successfully used to determine ligand binding affinities for both oligomeric and polymeric host duplexes (25–28).

Table 1. Calorimetrically derived binding enthalpies (ΔH_b) for the interactions of 2-ImPyPy-γ-PyPyPy-β-Dp with the three 11-mer DNA duplexes at 20°C

Duplex	ΔH_b,* kcal/mol
1 (5′-TGTTA-3′)	−6.7 ± 0.6
2 (5′-TGGTA-3′)	−4.6 ± 0.8
3 (5′-TATTA-3′)	−4.4 ± 0.6

Solution conditions are 10 mM sodium cacodylate (pH 6.9), 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$.

*ΔH_b values were determined at r_{Dup} of 1.0, with the indicated uncertainties corresponding to the sum of the standard deviations from three separate mixing experiments (DNA–ligand, ligand–buffer, and buffer–buffer) of at least 18 independent injections each.

Measured ligand-induced changes in the thermal stabilities of the three 11-mer duplexes (see Fig. 3) were used in conjunction with a binding site size, n_{app}, defined by both footprinting studies on a virtually identical ligand (12) and NMR studies on a similar polyamide without the hairpin link (8), to estimate apparent ligand–duplex association constants at T_m (K_{Tm}) from the expression (22):

$$\frac{1}{T_m^\circ} - \frac{1}{T_m} = \frac{R}{n_{app}(\Delta H_{W\text{-}C})} \ln[1 + (K_{T_m})a_f], \qquad [1]$$

where T_m° and T_m are the melting temperatures of the ligand-free and ligand-saturated duplexes, respectively; $\Delta H_{W\text{-}C}$ is the enthalpy change for the melting of a Watson-Crick (W-C) base pair in the absence of bound ligand (values we determined independently for each of the three target duplexes using DSC); and a_f is the free ligand activity for a W-C transition, which we estimated as one half the total ligand concentration. We then used the binding enthalpies (ΔH_b) listed in Table 1 to extrapolate these calculated binding constants at T_m to a common reference temperature of 20°C. The resulting K_{20} values are listed in Table 2. Inspection of these data reveals that the apparent binding affinities of the hairpin polyamide follows the hierarchy: duplex 1 (5′-TGTTA-3′) > duplex 3 (5′-TATTA-3′) > duplex 2 (5′-TGGTA-3′). Note the agreement between this hierarchy and that noted above for binding-induced enhancement in duplex thermal stability. Thus, given the binding enthalpies listed in Table 1, the relative extent to which ImPyPy-γ-PyPyPy-β-Dp thermally stabilizes the target duplex is correlated with its relative binding affinity.

The ≈9- to 47-fold relative higher affinity of the hairpin polyamide for the 5′-TGTTA-3′ site observed here is in agreement with the footprinting results of Mrksich *et al.* (12), while the absolute binding affinity is roughly an order of magnitude lower than that determined by Mrksich *et al.* (12). The latter difference is not surprising given the short length (11 bp) of our DNA target relative to the 135-bp DNA fragment used in the footprinting studies (12), and the large differences (1000- to 10,000-fold) in the DNA concentrations used in footprinting relative to our optical/calorimetric studies. The significant feature is that both the biophysical and the footprinting studies independently reveal that the hairpin polyamide binds preferentially to the 5′-TGTTA-3′ site.

The Preferential Binding of the Hairpin Polyamide to the 5′-TGTTA-3′ Match Site Is Enthalpic in Origin. Armed with the binding constants listed in Table 2, we calculated the corresponding binding free energies (ΔG_b) using the standard relationship

Table 2. ΔT_m-derived binding affinities of ImPyPy-γ-PyPyPy-β-Dp for the three 11-mer DNA duplexes at 20°C

Duplex	T_m°,* °C	T_m,* °C	K_{20},† M^{-1}
1 (5′-TGTTA-3′)	42.7 ± 0.3	53.7 ± 0.7	$7.3 \pm 1.3 \times 10^6$
2 (5′-TGGTA-3′)	46.1 ± 0.3	48.2 ± 0.5	$1.6 \pm 0.7 \times 10^5$
3 (5′-TATTA-3′)	38.1 ± 0.3	43.8 ± 0.5	$8.6 \pm 0.9 \times 10^5$

Solution conditions are as described in Table 1.

*T_m values were derived from UV melting profiles at 5 μM duplex (D) in the absence (T_m°) and presence of ligand (L) at 1L:1D stoichiometric ratios. Each T_m value is an average derived from two independent experiments, with the indicated errors corresponding to the average deviation from the mean.

†Binding constants at 20°C (K_{20}) were determined using Eq. **1**, a footprinting-derived apparent binding site size (n_{app}) of 5 bp/ligand (see ref. 12), the appropriate values of ΔH_b listed in Table 1, and the following calorimetrically determined duplex-to-single strand transition enthalpies ($\Delta H_{W\text{-}C}$) for the three host duplexes: 77.1 kcal/mol for duplex 1, 73.5 kcal/mol for duplex 2, and 60.1 kcal/mol for duplex 3. The indicated uncertainties reflect the maximum errors in K_{20} that result from the corresponding uncertainties noted above in T_m and T_m°, as propagated through Eq. **1**.

$$\Delta G_b = -RT\ln K. \quad [2]$$

These binding free energies, coupled with our calorimetrically determined binding enthalpies, also allowed us to calculate the corresponding binding entropies (ΔS_b) using

$$\Delta S_b = \frac{\Delta H_b - \Delta G_b}{T}. \quad [3]$$

These calculations enabled us to generate complete thermodynamic profiles for the binding of ImPyPy-γ-PyPyPy-β-Dp to each of the three 11-mer duplexes studied here. These profiles are summarized in Table 3. Inspection of these data reveals that, at 20°C, the preferential binding of the hairpin polyamide to duplex 1 (5′-TGTTA-3′) is primarily (≈73%) enthalpically driven, while the reduced binding to either duplex 2 (5′-TGGTA-3′) or duplex 3 (5′-TATTA-3′) is due to less favorable binding enthalpies. In fact, relative to duplex 1, the reduced binding to duplex 3 occurs despite a favorable entropic contribution to binding [$\Delta(T\Delta S)$], which is overcompensated by the enthalpy loss. This favorable entropic contribution may reflect binding-induced desolvation of the all-AT minor groove that is present only in duplex 3 (29–32). Thus, the preferential binding to the 5′-TGTTA-3′ match site is enthalpic in origin.

Single Base Pair Changes in the High-Affinity 5′-TGTTA-3′ Site Reduce the Hairpin Polyamide Binding Affinity by ≈1-2 kcal/mol. The data listed in Table 3 allow us to evaluate the thermodynamic consequences on hairpin polyamide binding of single base pair changes in the high-affinity 5′-TGTTA-3′ site (Table 4). Inspection of these data reveals that the single base pair changes that produce duplex 2 (T·A to G·C at position 3) and duplex 3 (G·C to A·T at position 2) result in losses of 2.1 and 2.3 kcal/mol of binding enthalpy, respectively, while resulting in entropy changes that depend on the nature of the alteration. These losses in binding enthalpy, coupled with the corresponding entropy changes, translate into losses in binding free energy of 2.2 and 1.2 kcal/mol, respectively, which reflect an ≈9- to 47-fold binding preference for duplex 1 relative to duplexes 2 and 3. Thus, the ≈1-2 kcal/mol enhanced affinity exhibited by the hairpin polyamide for the 5′-TGTTA-3′ site relative to two sites with single base pair changes is enthalpic in origin.

Correlation Between Thermodynamic and Structural Properties. A NMR and molecular modeling study (8) on the antiparallel, side-by-side heterodimeric complex of PyPyPy and ImPyPy bound to the minor groove of the d(CCTTGTTAGGC)·d(GCCTAACAAGG) B-form duplex provides us with a structural context in which to interpret our thermodynamic data. The binding site in the NMR study (8) is the same as that present in duplex 1 of this study, while the heterodimeric ligand is the same as that studied here except for the absence of the hairpin linker domain. The structural picture that emerges from the NMR study (8) is schematically shown in Fig. 1. Note that two classes of ligand-base hydrogen bonds are proposed that may be critical to the sequence specificity exhibited by the dimeric ligand complex. One class of hydrogen bonds involves an imidazole nitrogen and the 2-amino hydrogen of guanine, while a second class involves an amide hydrogen and either the N3 nitrogen of adenine, the O2 oxygen of thymine, or the O2 oxygen of cytosine. The differences we observe in the thermodynamics of ImPyPy-γ-PyPyPy-β-Dp binding to the three duplex targets studied here may be related to the ability of the ligand to form either or both of these types of hydrogen bonds with its duplex target, although we recognize that potential differences in van der Waals contacts and solvation also could be important contributors. According to the NMR-derived structural model (Fig. 1), when the hairpin polyamide is complexed with the high-affinity 5′-TGTTA-3′ site of duplex 1, it should be able to form one imidazole(N)–(amino H2)guanine hydrogen bond, two amide–(O2)thymine hydrogen bonds, one amide–(O2)cytosine hydrogen bond, and four amide–(N3)adenine hydrogen bonds. By contrast, when complexed with the 5′-TATTA-3′ site of duplex 3, the polyamide would be unable to form the imidazole(N)–(amino H2)guanine hydrogen bond (see Fig. 1), while when complexed with the 5′-TGGTA-3′ site of duplex 2, the polyamide would be unable to form one of the amide–(O2)thymine hydrogen bonds noted above. The polyamide may not be able to compensate this latter loss by forming a hydrogen bond between the unbonded amide hydrogen and the N3 atom of guanine due to steric interference from the neighboring 2-amino group (Fig. 1).

The inability of the hairpin polyamide to form the hydrogen bonds noted above when complexed with the 5′-TATTA-3′ and 5′-TGGTA-3′ sites of duplexes 3 and 2, respectively, may give rise to its reduced binding free energy relative to that which it exhibits when complexed with the 5′-TGTTA-3′ site of duplex 1 (Table 4). The enthalpic origin we observe for this reduction in binding free energy may reflect the enthalpic cost of failing to form these hydrogen bonds. We recognize that other factors, such as hydrophobic interactions and differential hydration of the polyamide and the DNA duplexes in their free and complexed states, also may contribute to the observed thermodynamic differences. However, we offer this simple

Table 4. Thermodynamic consequences of single base pair changes on the binding of ImPyPy-γ-PyPyPy-β-Dp to the 5′-TGTTA-3′ match site

Duplex	$\Delta\Delta H_b$, kcal/mol	$\Delta(T\Delta S_b)$, kcal/mol	$\Delta\Delta G_{b\text{-}20}$, kcal/mol
1 (5′-TGTTA-3′)	—	—	—
2 (5′-TGGTA-3′)	+2.1	−0.1	+2.2
3 (5′-TATTA-3′)	+2.3	+1.1	+1.2

$\Delta\Delta H_b$, $\Delta(T\Delta S_b)$, and $\Delta\Delta G_{b\text{-}20}$ were determined by subtracting the values of ΔH_b, $T\Delta S_b$, and $\Delta G_{b\text{-}20}$ for duplex 1 from the corresponding ΔH_b, $T\Delta S_b$, and $\Delta G_{b\text{-}20}$ values for either duplex 2 or duplex 3.

Table 3. Thermodynamic parameters for the binding of ImPyPy-γ-PyPyPy-β-Dp to the three 11-mer DNA duplexes

Duplex	ΔH_b,* kcal/mol	$T\Delta S_b$,† kcal/mol	$\Delta G_{b\text{-}20}$,‡ kcal/mol	K_{20},* M^{-1}
1 (5′-TGTTA-3′)	−6.7 ± 0.6	+2.5 ± 0.4	−9.2 ± 0.1	$7.3 \pm 1.3 \times 10^6$
2 (5′-TGGTA-3′)	−4.6 ± 0.8	+2.4 ± 0.4	−7.0 ± 0.2	$1.6 \pm 0.7 \times 10^5$
3 (5′-TATTA-3′)	−4.4 ± 0.6	+3.6 ± 0.5	−8.0 ± 0.1	$8.6 \pm 0.9 \times 10^5$

Solution conditions are as described in Table 1.

*The indicated errors in ΔH_b and K_{20} are as described in Tables 1 and 2, respectively.

†ΔS_b is the binding entropy, as determined using Eq. **3** and the corresponding values of ΔH_b and $\Delta G_{b\text{-}20}$. The indicated uncertainties reflect the maximum possible errors in ΔS_b that result from the corresponding uncertainties noted above in ΔH_b and $\Delta G_{b\text{-}20}$, as propagated through Eq. **3**.

‡$\Delta G_{b\text{-}20}$ is the binding free energy at 20°C, as determined using Eq. **2** and the corresponding value of K_{20}. The indicated uncertainties reflect the errors in $\Delta G_{b\text{-}20}$ that result from the corresponding uncertainties noted above in K_{20}, as propagated through Eq. **2**.

hydrogen bonding interpretation as one "explanation" of the experimental data that can serve as a basis for further discussion.

We thank Dr. Jens Völker for his assistance with the acquisition and analysis of the DSC data. This work was supported by National Institutes of Health Grants GM-23509 (K.J.B.), GM-34469 (K.J.B.), CA-47995 (K.J.B.), and GM-27681 (P.B.D.). We are grateful to the Howard Hughes Medical Institute for a Predoctoral Fellowship to E.E.B.

1. Wade, W. S., Mrksich, M. & Dervan, P. B. (1992) *J. Am. Chem. Soc.* **114,** 8783–8794.
2. Mrksich, M., Wade, W. S., Dwyer, T. J., Geierstanger, B. H., Wemmer, D. E. & Dervan, P. B. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7586–7590.
3. Wade, W. S., Mrksich, M. & Dervan, P. B. (1993) *Biochemistry* **32,** 11385–11389.
4. Pelton, J. G. & Wemmer, D. E. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 5723–5727.
5. Pelton, J. G. & Wemmer, D. E. (1990) *J. Am. Chem. Soc.* **112,** 1393–1399.
6. Chen, X., Ramakrishnan, B., Rao, S. T. & Sundaralingam, M. (1994) *Struct. Biol.* **1,** 169–175.
7. Mrksich, M. & Dervan, P. B. (1993) *J. Am. Chem. Soc.* **115,** 2572–2576.
8. Geierstanger, B. H., Jacobsen, J. P., Mrksich, M., Dervan, P. B. & Wemmer, D. E. (1994) *Biochemistry* **33,** 3055–3062.
9. Geierstanger, B. H., Dwyer, T. J., Bathini, Y., Lown, J. W. & Wemmer, D. E. (1993) *J. Am. Chem. Soc.* **115,** 4474–4482.
10. Dwyer, T. J., Geierstanger, B. H., Mrksich, M., Dervan, P. B. & Wemmer, D. E. (1993) *J. Am. Chem. Soc.* **115,** 9900–9906.
11. Mrksich, M. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 3663–3664.
12. Mrksich, M., Parks, M. E. & Dervan, P. B. (1994) *J. Am. Chem. Soc.* **116,** 7983–7988.
13. Parks, M. E., Baird, E. E. & Dervan, P. B. (1996) *J. Am. Chem. Soc.* **118,** 6141–6146.
14. Baird, E. E. & Dervan, P. B. (1996) *J. Am. Chem. Soc.* **118,** 6147–6152.
15. Griswold, B. L., Humoller, F. L. & McIntyre, A. R. (1951) *Anal. Chem.* **23,** 192–194.
16. Marky, L. A. & Breslauer, K. J. (1987) *Biopolymers* **26,** 1601–1620.
17. Breslauer, K. J. (1994) in *Methods in Molecular Biology*, ed. Agrawal, S. (Humana, Totowa, NJ), Vol. 26, pp. 347–372.
18. Mudd, C. P. & Berger, R. L. (1988) *J. Biochem. Biophys. Methods* **17,** 171–192.
19. Remeta, D. P., Mudd, C. P., Berger, R. L. & Breslauer, K. J. (1991) *Biochemistry* **30,** 9799–9809.
20. Robinson, A. L. (1932) *J. Am. Chem. Soc.* **54,** 1311–1318.
21. Gulbransen, E. A. & Robinson, A. L. (1934) *J. Am. Chem. Soc.* **56,** 2637–2641.
22. Crothers, D. M. (1971) *Biopolymers* **10,** 2147–2160.
23. McGhee, J. D. (1976) *Biopolymers* **15,** 1345–1375.
24. Neidle, S. & Abraham, Z. (1984) *CRC Crit. Rev. Biochem.* **17,** 73–121.
25. Snyder, J. G., Hartman, N. G., D'Estantoit, B. L., Kennard, O., Remeta, D. P. & Breslauer, K. J. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 3968–3972.
26. Marky, L. A., Curry, J. & Breslauer, K. J. (1985) in *Molecular Basis of Cancer*, ed. Rein, R. (Liss, New York), Part B, pp. 155–173.
27. Chou, W. Y., Marky, L. A., Zaunczkowski, D. & Breslauer, K. J. (1987) *J. Biomol. Struct. Dyn.* **5,** 345–359.
28. Breslauer, K. J., Freire, E. & Straume, M. (1992) *Methods Enzymol.* **211,** 533–567.
29. Drew, H. R. & Dickerson, R. E. (1981) *J. Mol. Biol.* **151,** 535–556.
30. Kopka, M. L., Yoon, C., Goodsell, D., Pjura, P. & Dickerson, R. E. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1376–1380.
31. Marky, L. A. & Breslauer, K. J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 4359–4363.
32. Chalikian, T. V., Plum, G. E., Sarvazyan, A. P. & Breslauer, K. J. (1994) *Biochemistry* **33,** 8629–8640.