Affinity and Specificity of Multiple Hydroxypyrrole/Pyrrole Ring Pairings for Coded Recognition of DNA

Sarah White, James M. Turner, Jason W. Szewczyk, Eldon E. Baird, and Peter B. Dervan*

Division of Chemistry and Chemical Engineering
California Institute of Technology
Pasadena, California, 91125

Received August 25, 1998

Pairings rules have been developed to guide the design of synthetic polyamides for recognition of predetermined sequences in the minor groove of DNA.[1,2] We recently reported that eight-ring hairpin-polyamides containing 3-hydroxypyrrole (Hp), imidazole (Im), and pyrrole (Py) amino acids form four-ring pairings (Im/Py, Py/Im, Hp/Py, and Py/Hp) which distinguish the four Watson−Crick base pairs in the minor groove of DNA.[2] An Im/Py pair distinguishes G·C from C·G and both of these from A·T/T·A base pairs.[1] A Hp/Py specifies T·A from A·T, and both of these from G·C/C·G.[2] The T·A selectivity of the Hp/Py pair likely arises from a combination of differential destabilization of polyamide binding via placement of Hp/Py opposite A·T or T·A, and specific hydrogen bonds between the 3-hydroxy and 4-carboxamido groups of Hp with the O2 of T (Figure 1).[3] A general pairing rule would require the same discrimination to be observed for the recognition of *multiple* T·A base pairs within other sequence contexts, including A·T rich sequences. In the original report, we observed that a single Hp/Py pair replacing a Py/Py pair destabilizes an eight-ring hairpin polyamide by 5-fold for an identical match site.[2] It remains to be determined whether consecutive Hp/Py ring pairings could target binding sites varying in their A·T base pair sequence composition without compromising polyamide affinity and sequence specificity. We report here that three consecutive Hp/Py pairs can be combined within a hairpin template to distinguish core sequence 5′-TTA-3′ from 5′-TAT-3′ in the DNA minor groove.

Application of the "pairings rules" predicts that the ten-ring hairpin polyamide ImHpHpPyPy-γ-ImHpPyPyPy-β-Dp (**1**) will bind to the 7-base pair site 5′-TGTTACA-3′ but will be mismatched with the site 5′-TGTATCA-3′ (Figure 1).[4,5] Quantitative DNase I footprint titration experiments performed on a 361-base pair DNA fragment revealed that polyamide **1** prefers the match site 5′-TGTTACA-3′ ($K_a = (9.4 \pm 1.0) \times 10^8$ M^{-1}) over 5′-TGTATCA-3′ ($K_a \leq 1 \times 10^7$ M^{-1}) by a factor of at least 94-fold (Figure 2). Control experiments on a separate fragment reveal that polyamide **1** binds to the single base mismatch site 5′-TGTT*C*CA-3′ with a $K_a \leq 1 \times 10^7$ M^{-1}, indicating that the Hp/Py pairing is disfavored for placement opposite a G·C base pair (see Table 1.

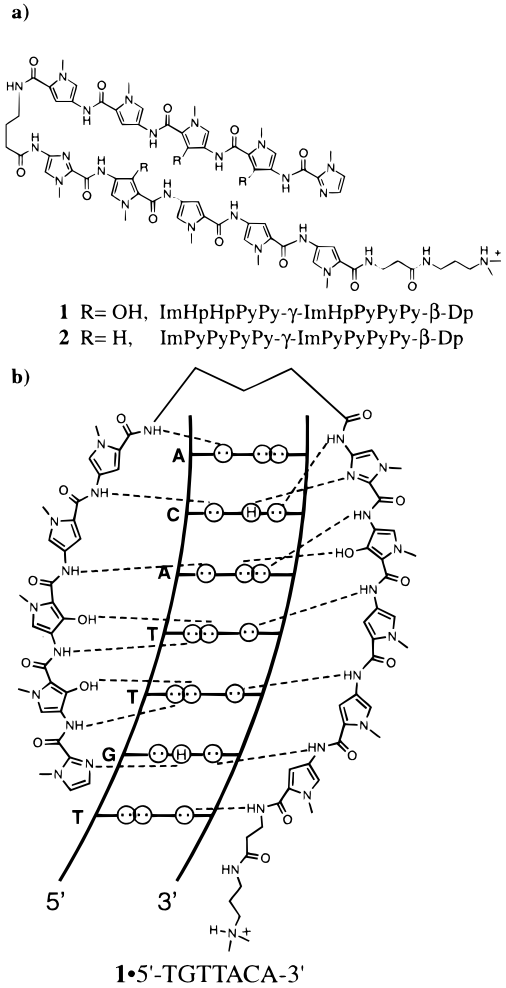


Figure 1. Structures of the 10-ring hairpin polyamides used for this study ($\beta = \beta$-alanine, Dp = dimethylaminopropylamide). (a) Polyamide **1** was synthesized by solid-phase methods using Boc-protected 3-methoxypyrrole, imidazole, and pyrrole aromatic amino acids.[2,7] Polyamide identity and purity was verified by ^{1}H NMR, analytical HPLC, and matrix-assisted laser-desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS-monoisotopic): **1** 1514.7 (1514.7 calcd). (b) Proposed binding model for the complex formed between the DNA and ImHpHpPyPy-γ-ImHpPyPyPy-β-Dp (**1**). A single circle with two dots represents the available lone pair of N3 of purines and the O2 of cytosine. Two touching circles with dots represent the two lone pairs of O2 of thymine. Circles containing an H represent the N2 hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines.

A symmetric pairing of Py/Py is degenerate for placement opposite A·T and T·A base pairs in the minor groove of DNA.[6] We find that control polyamide **2** which contains Py/Py pairs rather than Hp/Py ring pairings binds within a factor of 4 to the sites 5′-TGTTACA-3′ ($K_a = (1.0 \pm 0.4) \times 10^{10}$ M^{-1}) and 5′-TGTATCA-3′ ($K_a = (2.5 \pm 0.9) \times 10^9$ M^{-1}), consistent with the assignment of Py/Py as degenerate for A·T and T·A. In addition, only 18-fold specificity is observed when polyamide **2** is placed opposite the site 5′-TGTTCCA-3′ versus 5′-TGTTACA-3′.[5]

(1) (a) Wade, W. S.; Mrksich, M.; Dervan, P. B. *J. Am. Chem. Soc.* **1992**, *114*, 8783. (b) Mrksich, M.; Wade, W. S.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 7586. (c) White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569. (d) Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nature Struct. Biol.* **1998**, *5*, 104.

(2) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468.

(3) Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, *282*, 111.

(4) (a) Mrksich, M.; Parks, M. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1994**, *116*, 7983. (b) Parks, M. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6147. (c) Pilch, D. S.; Poklar, N. A.; Gelfand, C. A.; Law, S. M.; Breslauer, K. J.; Baird, E. E.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8306−8311. (d) de Claire, R. P. L.; Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7906. (e) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559−561.

(5) Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 7636.

(6) (a) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723. (b) Pelton, J. G.; Wemmer, D. E. *J. Am. Chem. Soc.* **1990**, *112*, 1393. (c) Chen, X.; Ramakrishnan, B.; Rao, S. T.; Sundaralingham, M. *Nature Struct. Biol.* **1994**, *1*, 169. (d) White, S.; Baird, E. E.; Dervan, P. B. *Biochemistry* **1996**, *35*, 12532.

(7) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.

(8) Brenowitz, M.; Senear, D. F.; Shea, M. A.; Ackers, G. K. *Methods Enzymol.* **1986**, *130*, 132.

(9) Iverson, B. L.; Dervan, P. B. *Nucleic Acids Res.* **1987**, *15*, 7823−7830.

10.1021/ja9830593 CCC: $18.00
Published on Web 12/17/1998

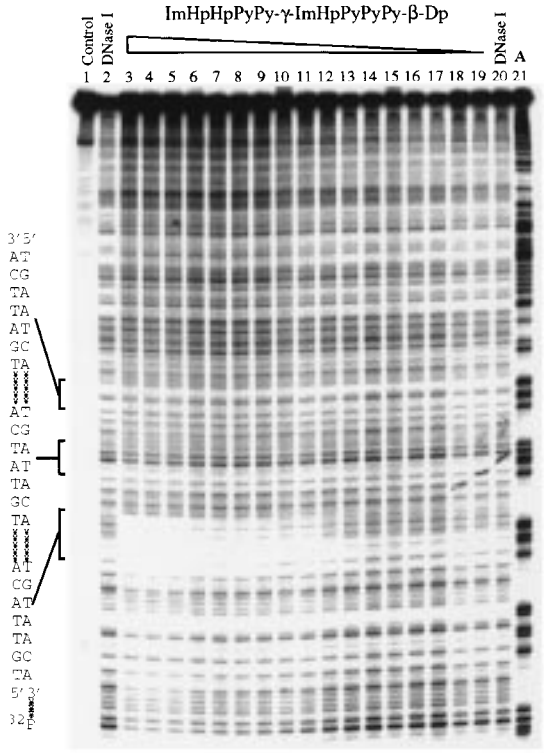
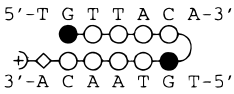
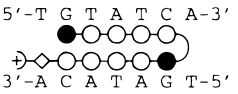
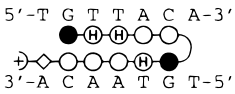
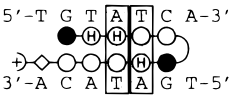


Figure 2. Storage phosphor autoradiogram of an 8% denaturing polyacrylamide gel used to separate the fragments generated by DNase I digestion in a quantitative footprint titration experiment[8] with polyamide **1**: lane 1, intact DNA; lanes 2 and 20, DNase I digestion products in the absence of polyamide; lanes 3−19, DNase I digestion products in the presence of 100, 65,40, 25, 15, 10, 6.5, 4.0, 2.5, 1.5, 1.0, 0.5, 0.2, 0.1, 0.05, 0.02, 0.01 nM polyamide; lane 21, adenine-specific[9] chemical sequencing. All reactions were done in a total volume of 400 mL. All reactions contain 3′-^{32}P-end-labeled *Eco*RI/ *Pvu*II restriction fragment from plasmid pDEH11 (20kcpm), 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ (pH 7.0, 22 °C). Solutions were allowed to equilibrate for 4−12 h at 22 °C before initiation of footprinting reactions. Footprinting reactions, separation of cleavage products, and data analysis were carried out as previously described.[6d] The plasmid pDEH11 was constructed by ligation of an insert, 5′-d(CTAGGCGAGCGCTCATTGTTACAGCGAGCGCTCATTGTATCAGCGAGCGCTCATTGATTCAGCGAGCGCTCATGCA)-3′ and 5′-d(TGAGCGCTCGCTGAATCAATGAGCGCTCGCTGATACAATGAGCGCTCGCTGTAACAATGAGCGCTCGC)-3′ into pUC19 previously cleaved with *Xba*I and *Pst*I. The plasmid was digested with *Eco*RI, labeled at the 3′ end, and digested with *Pvu*II. The 361-base pair fragment was isolated by nondenaturing gel electrophoresis.

High-resolution X-ray structure analysis reveals that Hp/Py polyamides bind undistorted B-form DNA; however, a localized 0.5 Å melting of the T·A Watson−Crick base pair is observed that is potentially responsible for the energetic destabilization of Hp/Py relative to the Py/Py pair.[3] Remarkably, three Hp/Py pairs destabilize polyamide **1** for 5′-TGTTACA-3′ by only 11-fold relative to polyamide **2**. Therefore, multiple Hp/Py pair substitutions within this ten-ring polyamide do not appear to have an additive effect on binding affinity.[2]

Table 1. Equilibrium Association Constants (K_a)[a,b,c]

5'-TTA-3'[d]	5'-TAT-3'	Specificity[e]
		4-fold
$1.0\ (\pm 0.4) \times 10^{10}\ M^{-1}$	$2.5\ (\pm 0.9) \times 10^{9}\ M^{-1}$	
		≥ 94-fold
$9.4\ (\pm 1.0) \times 10^{8}\ M^{-1}$	$\leq 1 \times 10^{7}\ M^{-1}$	

[a] Reported mean values were obtained from three DNase I footprint titration experiments. The standard deviation for each value is indicated in parentheses. [b] The assays were carried out at 22 °C at pH 7.0 in the presence of 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$. [c] Binding models for polyamides **1** and **2** with 5′-TGTTACA-3′ and 5′-TGTATCA-3′. Filled and unfilled circles represent Im and Py rings, respectively, circles containing an H represent Hp, the curved line connecting the polyamide subunits represents γ, the diamond represents β-alanine, and the + represents the positively charged dimethylaminopropylamide tail group. [d] Central 3-base pairs of the 7-base pair binding sites. [e] Specificity is calculated as K_a (5′-TGTTACA-3′)/K_a (5′-TGTATCA-3′).

Hydroxypyrrole−imidazole−pyrrole polyamides complete the minor-groove recognition code using the four-ring pairings (Im/Py, Py/Im, Hp/Py, and Py/Hp) to complement the four Watson−Crick base pairs. The observation of substantially increased A·T/T·A specificity without a severe energetic penalty by incorporation of multiple Hp/Py pairings is a minimal first step toward full integration of multiple Hp/Py ring pairings for Py/Py pairs. Because the sequence dependent microstructure of DNA is still not well-understood, the full scope and limitations of the Hp/Im/Py pairing rules with regard to energetics and specificity remains to be elucidated and will be reported in due course.

Acknowledgment. We are grateful to the National Institutes of Health (GM-27681) for research support and National Research Service Awards to S.W. and J.W.S., J. Edward Richter for an undergraduate fellowship to J.M.T., and the HHMI for a predoctoral fellowship to E.E.B.

JA9830593