J|A|C|S
ARTICLES
Published on Web 06/20/2006

Programmable Oligomers for Minor Groove DNA Recognition

Raymond M. Doss, Michael A. Marques, Shane Foister, David M. Chenoweth, and Peter B. Dervan*

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

Received March 30, 2006; E-mail: dervan@caltech.edu

Abstract: The four Watson−Crick base pairs of DNA can be distinguished in the minor groove by pairing side-by-side three five-membered aromatic carboxamides, imidazole (**Im**), pyrrole (**Py**), and hydroxypyrrole (**Hp**), four different ways. On the basis of the paradigm of unsymmetrical paired edges of aromatic rings for minor groove recognition, a second generation set of heterocycle pairs, imidazopyridine/pyrrole (**Ip/Py**) and hydroxybenzimidazole/pyrrole (**Hz/Py**), revealed that recognition elements not based on analogues of distamycin could be realized. A new set of end-cap heterocycle dimers, oxazole-hydroxybenzimidazole (**No**−**Hz**) and chlorothiophene-hydroxybenzimidazole (**Ct**−**Hz**), paired with **Py**−**Py** are shown to bind contiguous base pairs of DNA in the minor groove, specifically 5′-GT-3′ and 5′-TT-3′, with high affinity and selectivity. Utilizing this technology, we have developed a new class of oligomers for sequence-specific DNA minor groove recognition no longer based on the *N*-methyl pyrrole carboxamides of distamycin.

Introduction

Aberrant gene expression is the cause of many diseases, and the ability to reprogram transcriptional pathways using cell-permeable small molecules may, one day, have an impact on human medicine.[1] DNA-binding polyamides, which are based on the architecture of the natural products netropsin and distamycin A,[2,3a,b] are capable of distinguishing all four Watson−Crick base pairs in the DNA minor groove and have been the subject of intense study along with many other classes of minor groove binders.[3c−f,4,5] Sequence-specific recognition of the minor groove of DNA by polyamides arises from the pairing of three different antiparallel five-membered heterocyclic amino acids, pyrrole (**Py**), imidazole (**Im**), and hydroxypyrrole (**Hp**).[4,5] The direct read out, or information face, on the inside of the crescent-shaped polyamide may be programmed by the incremental change of atoms on the corners of the ring pairs presented to the DNA minor groove floor. Stabilizing and, importantly, destabilizing interactions with the different edges of the four

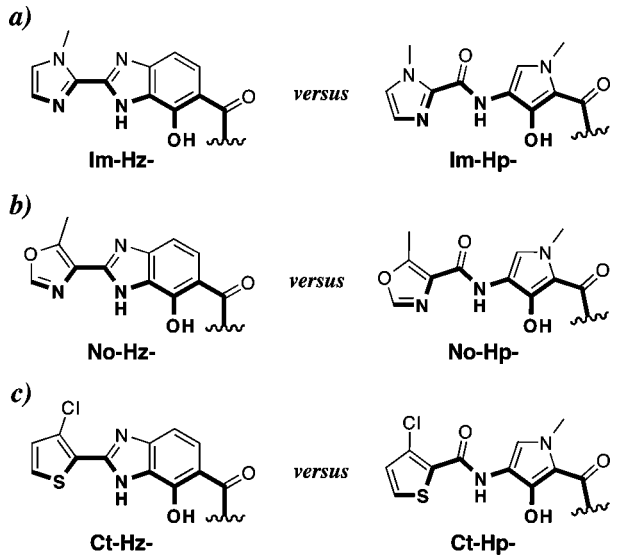


Figure 1. Structures of dimers (a) imidazole−hydroxybenzimidazole (**Im**−**Hz**), (b) oxazole−hydroxybenzimidazole (**No**−**Hz**), and (c) chlorothiophene−hydroxybenzimidazole (**Ct**−**Hz**) dimer caps in comparison with their respective five membered ring systems. Hydrogen-bonding surfaces to the DNA minor-groove floor are bolded.

Watson−Crick bases are modulated by shape complementarity and specific hydrogen bonds.[4−6,7l] For example, the **Im/Py** pair distinguishes G·C from C·G, T·A, and A·T. **Im** presents a lone pair of electrons to the DNA minor groove and can accept a hydrogen bond from the exocyclic amine of guanine.[5] Additionally, the **Hp/Py** pair distinguishes T·A from A·T, G·C, and C·G.[4−6] **Hp** projects an exocyclic OH group toward the minor

(1) (a) Darnell, J. E. *Nat. Rev. Cancer* **2002**, *2*, 740−748. (b) Pandolfi, P. P. *Oncogene* **2001**, *20*, 3116−3127. (c) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284−299.
(2) (a) Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X. *J. Am. Chem. Soc.* **1951**, *73*, 341−343. (b) Arcamone, F. N. V.; Penco, S.; Orezzi, P.; Nicolella, V.; Pirelli, A. *Nature* **1964**, *203*, 1064−1065.
(3) (a) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376−1380. (b) Pelton, J. G.; Wemmer, D. E. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723−5727. (c) Wemmer, D. E. *Annu. Rev. Biomol. Struct.* **2000**, *29*, 439−461. (d) Buchmueller, K. L.; Staples, A. M.; Howard, C. M.; Horick, S. M.; Uthe, P. B.; Minh Le, N.; Cox, K. K.; Nguyen, B.; Pacheco, K. A. O.; Wilson, D. W.; Lee, M. *J. Am. Chem. Soc.* **2004**, *127*, 742−750. (e) Baraldi, P. G.; Bovero, A.; Fruttarolo, F.; Preti, D.; Tabrizi, M. A.; Pavani, M. G.; Romagnoli, R. *Med. Res. Rev.* **2004**, *24*, 475−528. (f) Reddy, P. M.; Toporowski, J. W.; Kahane, A. L.; Bruice, T. C. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5531−5536.
(4) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468−471.
(5) (a) Kielkopf, C. L.; White, S.; Szewcyzk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Science* **1998**, *282*, 111−115. (b) Kielkopf, C. L.; Bremer, R. E.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *J. Mol. Biol.* **2000**, *295*, 557−567.
(6) (a) Urbach, A. R.; Szewczyk, J. W.; White, S.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 11621−11629. (b) White, S.; Turner, J. M.; Szewezyk, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 260−261.

10.1021/ja0621795 CCC: $33.50 **© 2006 American Chemical Society**

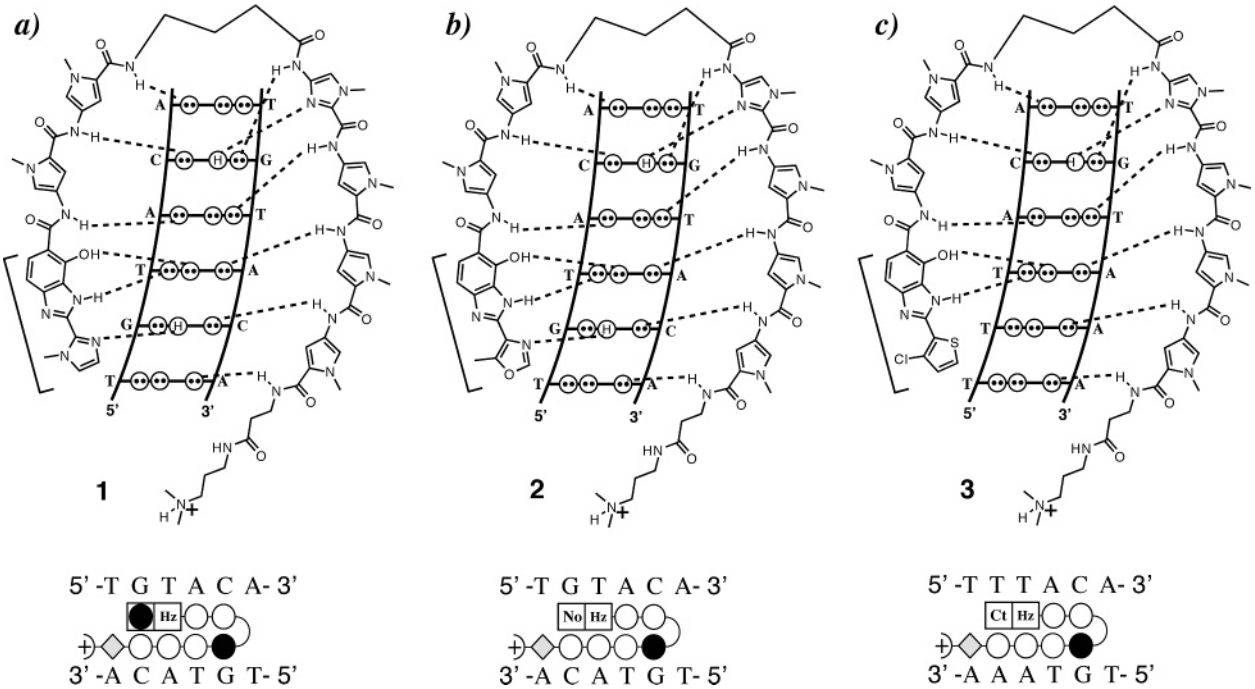


Figure 2. Postulated hydrogen-bonding models for the 1:1 polyamide−DNA complexes with their matched sequence and their ball-and-stick representations. (a) Im−Hz−Py−Py-γ-Im−Py−Py−Py-β-Dp (**1**), (b) No−Hz−Py−Py-γ-Im−Py−Py−Py-β-Dp (**2**), and (c) Ct−Hz−Py−Py-γ-Im−Py−Py−Py-β-Dp (**3**).

groove floor that is sterically accommodated in the cleft of the T·A base pair, preferring to lie over T, not A.[5] These pairing rules have proven useful for programmed recognition of a broad repertoire of DNA sequences; however, the hydroxypyrrole ring system has proven to be unstable over time and in the presence of acid, further prompting our search for new T·A/A·T recognition elements. In addition, sequence-dependent changes in the microstructure of DNA (intrinsic minor groove width, minor groove flexibility, and inherent DNA curvature)[7a−k] combined with structural and conformational changes among polyamides make the targeting of certain sequences less than optimal, leading us to explore whether other novel heterocyclic recognition elements could be discovered for use in DNA groove recognition within the unsymmetrical pairing paradigm.[7l,8−10] Furthermore, from a medicinal chemistry point of view, a broader tool kit of sequence-specific recognition elements for DNA beyond polyamides would be useful as our artificial transcription factor program moves from cell culture[11] to small animal studies.

We recently reported that the benzimidazole ring can be an effective platform for the development of modular paired recognition elements for the minor groove of DNA.[9,10] The benzimidazole 6−5 bicyclic ring structure, though having slightly different curvature from the classic five-membered pyrrole-carboxamides, presents an "inside edge" with a similar atomic readout to the DNA minor groove floor, effectively mimicking **Py**, **Im**, and **Hp**. We demonstrated that the imidazopyridine/pyrrole pair **Ip/Py** distinguishes G·C from C·G and the hydroxybenzimidazole/pyrrole pair **Hz/Py** distinguishes T·A from A·T, providing a solution to the unanticipated hydroxypyrrole instability limitation.[9,10] The question arises whether this second generation solution to DNA recognition can be elaborated further, *deleting incrementally almost all carboxamide linkages in the backbone of the hairpin motif*.[12]

We report here a new set of heterocycle dimer pairs,[12] which represents a step from single base-pair recognition toward a two letter approach to molecular recognition of the minor groove of DNA (Figure 1). We move from single letters to syllables. New heterocycles were designed by combining the T-specific hydroxybenzimidazole (**Hz**) with oxazole (**No**) rings and chlorothiophene (**Ct**) caps[8] to afford the recognition elements **No**−**Hz** and **Ct**−**Hz**, respectively (Figure 1). Quantitative DNase I footprinting titrations were used to determine DNA binding affinities of hairpin oligomers containing the **No**−**Hz** and **Ct**−**Hz** dimers paired with **Py**−**Py** dimer for each of the four Watson−Crick bases (Figure 2). When positioned at the termini of hairpin polyamides, the **No**−**Hz/Py**−**Py** and **Ct**−

(7) (a) Hays, F. A.; Teegarden, A.; Jones, Z. J. R.; Harms, M.; Raup, D.; Watson, J.; Cavaliere, E.; Shing Ho, P. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7157−7162. (b) Beveridge, D. L.; Barreiro, G.; Byun, K. S.; Case, D. A.; Cheatham, T. E., 3rd; Dixit, S. B.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Seibert, E.; Sklenar, H.; Stoll, G.; Thayer, K. M.; Varnai, P.; Young, M. A. *Biophys. J.* **2004**, *87*, 3799−3813. (c) Dixit, S. B.; Beveridge, D. L.; Case, D. A.; Cheatham, T. E., 3rd; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Sklenar, H.; Thayer, K. M.; Varnai, P. *Biophys. J.* **2005**, *89*, 3721−3740. (d) Wu, H.; Crothers, D. M. *Nature* **1984**, *308*, 509. (e) Steitz, T. A. *Annu. Rev. Biophys.* **1990**, *23*, 205. (f) Goodsell, D. S.; Kopka, M. L.; Cascio, D.; Dickerson, R. E. *Proc. Natl. Sci. U.S.A.* **1993**, *90*, 2930. (g) Paolella, D. N.; Palmer, R.; Schepartz, A. *Science* **1994**, *264*, 1130. (h) Kahn, J. D.; Yun, E.; Crothers, D. M. *Nature* **1994**, *368* 163. (i) Geierstanger, B. H.; Wemmer, D. E. *Annu. Rev. Biochem.* **1995**, *24*, 463. (j) Hansen, M. R.; Hurley, L. H. *Acc. Chem. Res.* **1996**, *29*, 249. (k) Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 6219−6226. (l) Marques, M. A.; Doss, R. M.; Urbach, A. R.; Dervan, P. B. *Helv. Chim. Acta* **2002**, *85*, 4485−4517.

(8) Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B. *Bioorg. Med. Chem.* **2003**, *11*, 4333−4340.

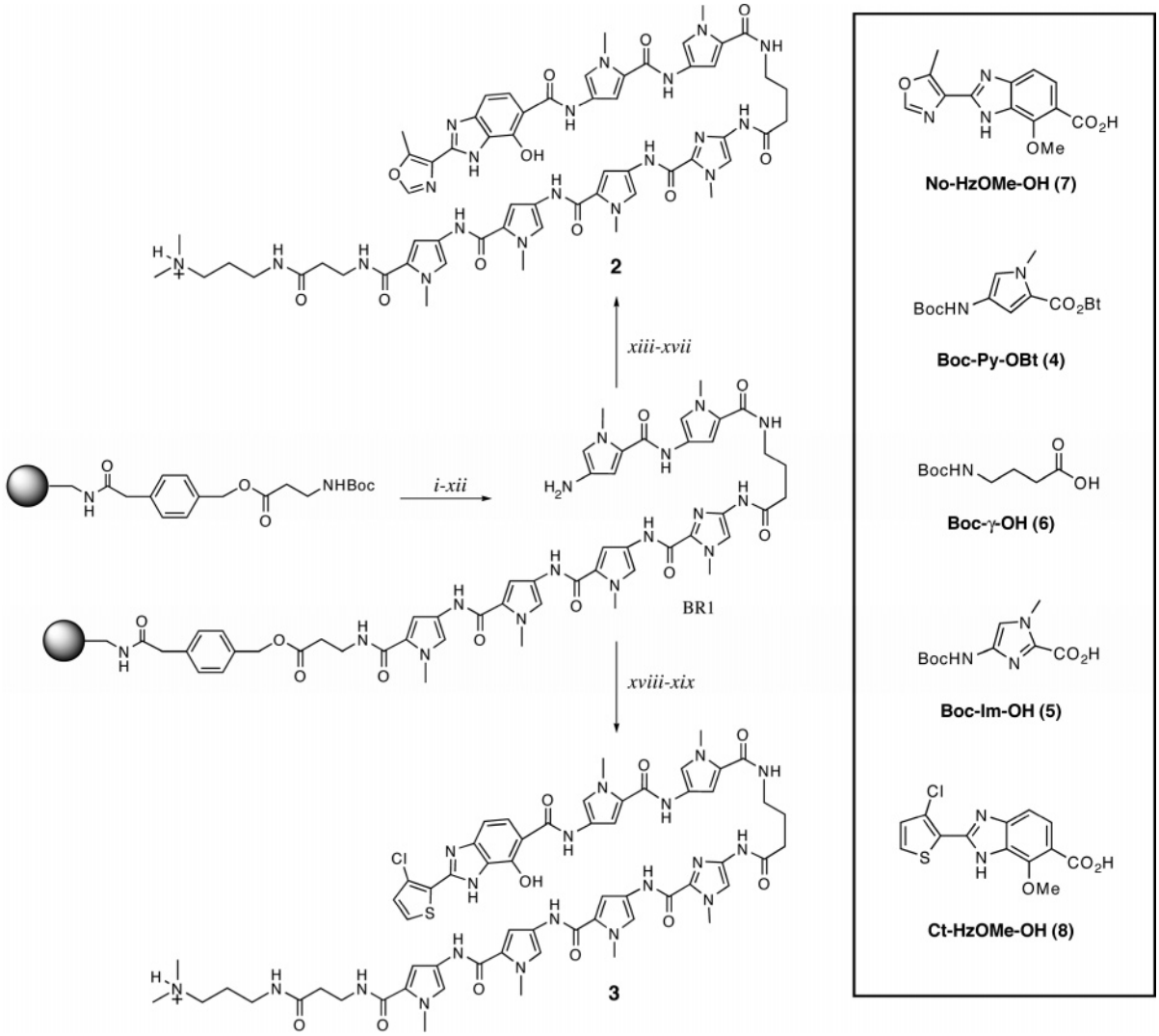
(9) Briehn, C. A.; Weyermann, P.; Dervan, P. B. *Chem.−Eur. J.* **2003**, *9*, 2110−2122.

(10) (a) Renneberg, D.; Dervan, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 5707−5716. (b) Marques, M. A.; Doss, R. M.; Foister, S.; Dervan, P. B. *J. Am. Chem. Soc.* **2004**, *126*, 10339−10349.

(11) For downregulation of an endogenous gene in cell culture by polyamides see: Olenyuk, B. Z.; Zhang, G.; Klco, J. M.; Nickols, N. G.; Kaelin, W. G., Jr.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16768−16773.

(12) Heterocycle pair refers to cofacial stack (noncovalent interaction), whereas dimer refers to two covalently attached heterocycles.

Scheme 1. Representative Solid-Phase Synthesis of Polyamide **2** and **3** along with a Table of the Amino Acid Building Blocks Used[a]



[a] Reaction conditions: (i) 80% TFA/DCM; (ii) Boc-Py-OBt, DIEA, DMF; (iii) Ac_2O, DIEA, DMF; (iv) repeat i−iii × 2; (v) 80% TFA/DCM; (vi) Boc-Im-OH, HBTU, DIEA, DMF; (vii) Ac_2O, DIEA, DMF; (viii) 80% TFA/DCM; (ix) Boc-γ-OH, HBTU, DIEA, DMF; (x) Ac_2O, DIEA, DMF; (xi) repeat i−iii × 2; (xii) 80% TFA/DCM; (xiii) No−HzOMe-OH, HBTU, DIEA, DMF; (xiv) 3-(dimethylamino)-1-propylamine (Dp), 80 °C, 2 h; (xv) preparative HPLC; (xvi) thiophenol, NaH, DMF; (xvii) preparative HPLC; (xviii) Ct-HzOMe-OH, HBTU, DIEA, DMF; (xix) same as steps xiv−xvii.

Hz/Py−Py dimer pairs are found to target 5′-GT-3′ and 5′-TT-3′ sequences, respectively, with high affinity and good specificity. With the development of dimer pairs capable of recognizing a 5′-GT-3′ sequence of DNA, we could address the question whether a hairpin oligomer comprised of four dimer units will bind the site 5′-GTAC-3′, a sequence formally containing all four Watson−Crick base pairs. Such a molecule represents our first programmable oligomer, which demonstrates excellent DNA binding properties without containing a single pyrrole- or imidazole-carboxamide based on the natural product distamycin, a design benchmark for biomimetic chemistry and the field of DNA recognition.

Experimental Methods

Polyamide Synthesis. Hairpin polyamides were synthesized manually from Boc-β-PAM resin in a stepwise fashion using Boc-protected monomeric and dimeric amino acids according to established solid-phase protocols.[13] Base Resin 1 (**BR1**) (H_2N−Py−Py-γ-Im−Py−Py−Py-β-Pam) was synthesized in gram quantities using the following amino acid building blocks: Boc-**Py**-OBt (**4**), Boc-**Im**-OH (**5**), and Boc-γ-OH (**6**) (Scheme 1). The base resins were then split into smaller batches for coupling to the final dimeric caps. Boc-protected amino acid monomers for Boc-**Py**-OBt (**4**) and Boc-**Im**-OH (**5**) were synthesized according to previously reported procedures.[8,13,14] Dimeric cap synthesis for **No**−**Hz**OMe-OH (**7**) and **Ct**−**Hz**OMe-OH (**8**) are detailed in the Supporting Information. Couplings were achieved using preactivated monomers (Boc-**Py**-OBt) or HBTU activation in a DIEA and DMF mixture. Coupling times ran from 3 to 24 h at 25−40 °C. Deprotection of the growing polyamide was accomplished using 80% TFA/DCM. Polyamides were cleaved from the resin by treatment with neat 3-(dimethylamino)-1-propylamine (Dp) at 80 °C for 2 h and purified by preparatory reverse phase HPLC. Deprotection of the methoxy-protected polyamides was done using a mix of thiophenoxide in DMF at 80 °C to provide the free hydroxy derivatives after a second HPLC purification: **Im**−**Hz**−Py−Py-γ-Im−Py−Py−Py-β-Dp (**1**), **No**−**Hz**−Py−Py-γ-Im−Py−Py−Py-β-Dp (**2**), **Ct**−**Hz**−Py−Py-γ-Im−Py−Py−Py-β-Dp (**3**). (See Supporting Information for full experimental details.)

(13) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141−6146.

(14) Urbach, A. R.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4343−4348.

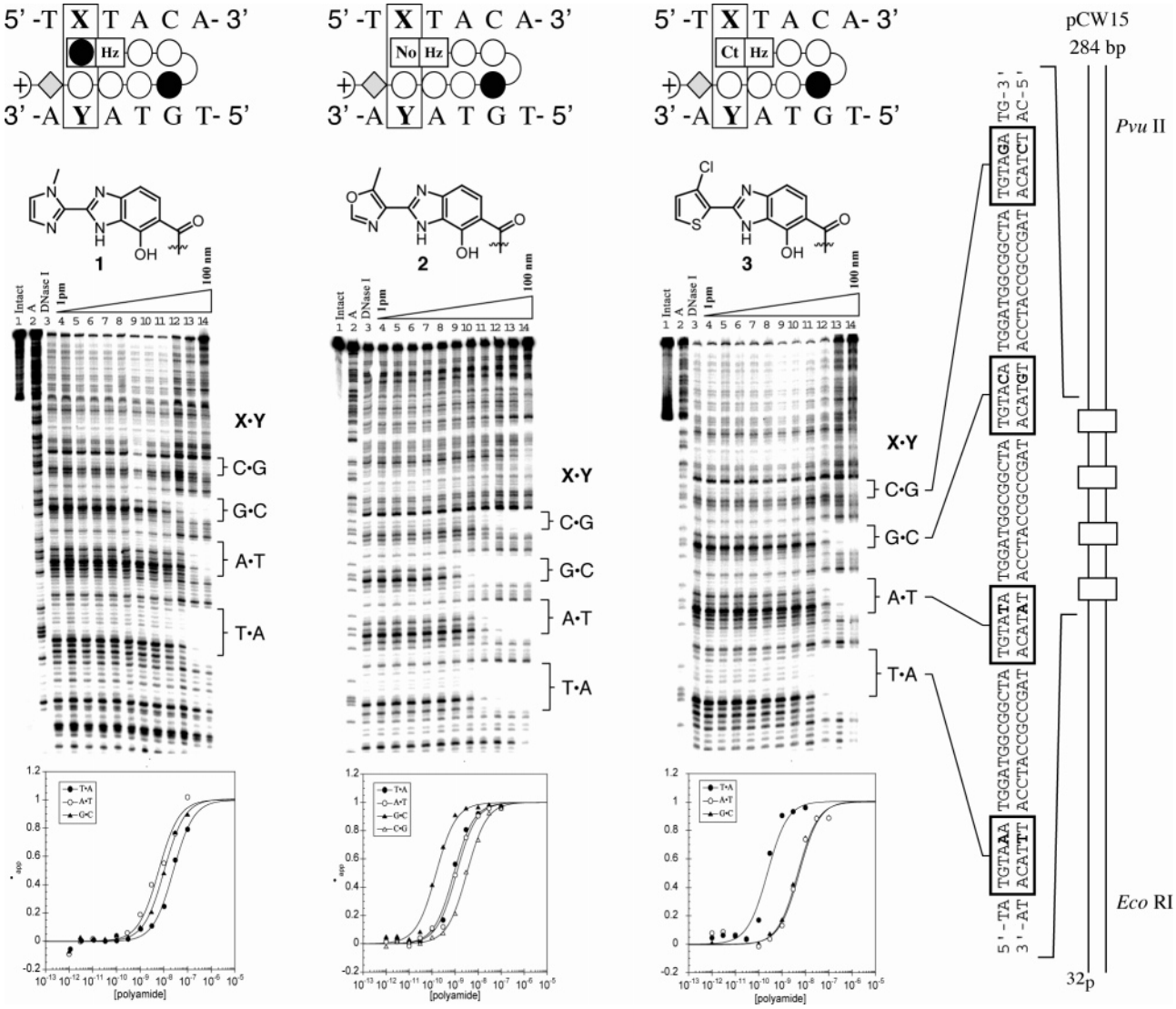


Figure 3. Quantitative DNase I footprinting experiments in the hairpin motif for polyamides 1, 2, and 3, respectively, on the 278 bp, 5′-end-labeled PCR product of plasmid CW15: (lane 1) intact DNA; (lane 2) A reaction; (lane 3) DNase I standard; (lanes 4−14) 1 pM, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, 100 nM polyamide, respectively. Each footprinting gel is accompanied by the following: (top) Chemical structure of the pairing of interest; (bottom) binding isotherms for the four designed sites. θ_{norm} values were obtained according to published methods.[15] A binding model for the hairpin motif is shown centered at the top as a ball-and-stick model with the polyamide bound to its target DNA sequence. Imidazoles and pyrroles are shown as filled and nonfilled circles, respectively; β-alanine is shown as a diamond; the γ-aminobutyric acid turn residue is shown as a semicircle connecting the two subunits.

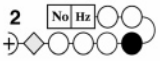
Results

DNA Affinity and Sequence Specificity of Dimer Caps. Quantitative DNase-I footprinting titrations were carried out for polyamides **1**−**3**. All polyamides were footprinted on the 285-base-pair PCR product of plasmid pCW15. In all cases, the DNA-sequence specificity at the cap position (in bold) was determined by varying a single DNA base pair within the sequence, 5′−T**X**TACA-3′, to all four Watson−Crick base pairs (**X** = A, T, G, C) and comparing the relative affinities of the resulting complexes. The variable base-pair position was designed to be adjacent to the **Hz** ring, which has been shown to specify for T when paired across from **Py**, so as to be able to determine the binding properties of each compound to the following two base-pair sequences: AT, TT, GT, and CT.

The sequence specificity of the **Im**−**Hz** and **Ct**−**Hz** dimers for 5′-T**X**TACA-3′ were evaluated in polyamides **1** and **3**, respectively. As expected, polyamide **3** bound its designed match site 5′-T**TT**ACA-3′ ($K_a = 2.4 \times 10^9$ M^{-1}) (Figure 3, Table 1) with both the **Ct** and the **Hz** halves of the dimer preferring to rest over the less bulky T in the asymmetric cleft of a T·A base pair. Placing the **Ct** ring adjacent to the **Hz** resulted in a 10-fold specificity for T > A using the **Ct**−**Hz** system. Polyamide **1**, which contains the **Im**−**Hz** dimer, did not bind its designed match site 5′-T**GT**ACA-3′ with any appreciable level of specificity exhibiting affinities of $K_a = 1.6 \times 10^8$ and 4.0×10^8 M^{-1} for the **GT** and **AT** sites, respectively.

Oxazole cap (polyamide **2**) was incorporated into the dimer cap system, and the affinity for its designed match site, 5′-T**GT**ACA-3′, was examined. Polyamide **2** successfully targeted its designed match site with an appreciable level of specificity (25-fold) and a match site affinity of $K_a = 6.8 \times 10^9$ M^{-1} (Figure 3, Table 1). With the development of the chlorothiophene and oxazole dimer caps, the range of targetable sequences by polyamides has been expanded (Table 1).

Table 1. Affinities of X/Py Ring Pairs Proximal to a Hydroxybenzimidazole Bicycle K_a [M^{-1}][a,b]

Polyamide	5'-tATACa-3'	5'-tTTACa-3'	5'-tGTACa-3'	5'-tCTACa-3'
1 	9.7(±0.7)x10^7	4.5(±0.6)x10^8	1.7(±0.4)x10^8	≤ 1.0 x 10^7
2 	8.6(±0.3)x10^8	9.5(±0.3)x10^8	**6.8(±0.4)x10^9**	2.7(±0.5)x10^8
3 	2.1(±0.3)x10^8	**2.4(±0.2)x10^9**	2.6(±0.4)x10^8	≤ 1.0 x 10^7

[a] Values reported are the mean values from at least three DNase I footprinting titration experiments, with the standard deviation given in parentheses. [b] Assays were performed at 22 °C in a buffer of 10 mM Tris·HCl, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0.

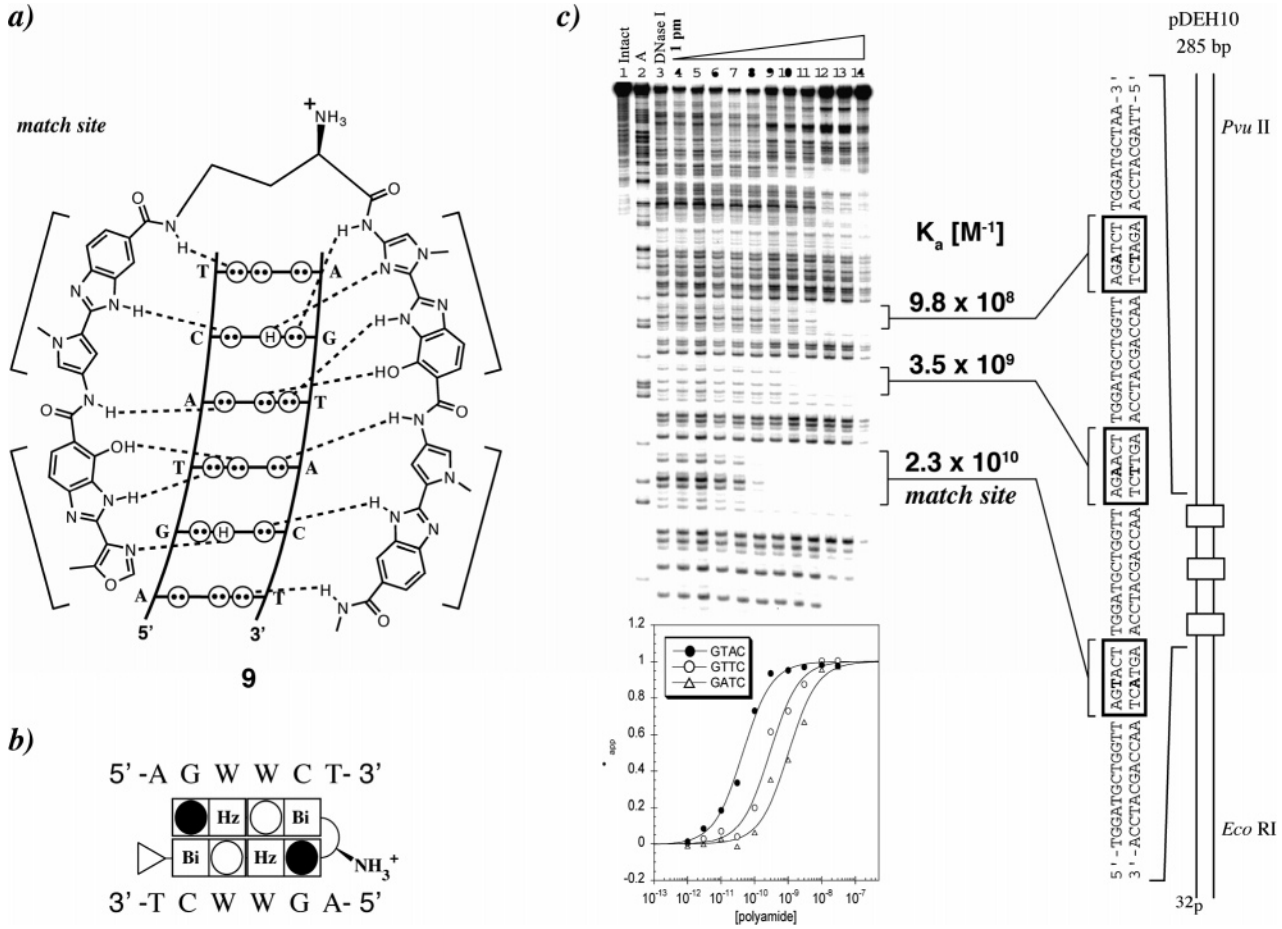


Figure 4. (a) Postulated hydrogen-bonding model and structure of oligomer **9**. (b) Ball and stick representation of **9** and the 6-base-pair binding site with variable region (W = A or T) shown. (c) Quantitative DNase I footprint titration experiment on the 5′-^{32}P-labeled PCR product shown with an illustration and complete sequence of the 285 bp *Eco*RI/*Pvu*II restriction fragment from plasmid pDEH10. Binding affinities are shown next to their respective binding sites and the match site is designated.

Design of a Programmable Oligomer for 5′-GTAC-3′. The synthesis of oligomer **9** containing four dimer units was achieved via the stepwise addition of Boc-amino acid dimers in the same manner as previously described polyamide syntheses. This "third generation" oligomer's binding properties were assessed in the same context[16] as previously reported for first and second generation hairpin polyamides targeting the sequence, 5′-GTAC-3′, containing the four Watson−Crick base pairs.[4,10] Footprinting of the oligomer on the previously characterized plasmid DEH10 showed a binding affinity of $K_a = 2.3 \times 10^{10}$ M^{-1} for the match site 5′-GTAC-3′ and affinities of $K_a = 3.5 \times 10^9$ and 9.8×10^8 M^{-1} for the mismatch sites 5′-G*A*AC-3′ and 5′-G*AT*C-3′, respectively (Figure 4). Such a result demonstrates that a compound consisting exclusively of 6−5 fused ring systems and minimal carboxamide linkages is able to maintain good levels of specificity and excellent binding affinity.

Discussion. Recent advances in hairpin polyamide designs have traditionally focused on developing new modes of single base-pair recognition by heterocyclic ring pairings. Previous studies, however, have highlighted the fact that the microstruc-

(15) Trauger, J. W.; Dervan, P. B. *J. Methods Enzymol.* **2001**, *340*, 450−466.
(16) Im−Hp−Py−Py-γ-Im−Hp−Py−Py-β-Dp targets 5′-WGTACW-3′ with $K_a = 7.0 \times 10^8$ M^{-1} and Im−Hz−Py−Py-γ-Im−Hz−Py−Py-β-Dp targets 5′-WGTACW-3′ with $K_a = 4.6 \times 10^8$ M^{-1} where W = A or T (see ref 10b).

ture of DNA depends on the sequence in question.[17] In addition, structural and conformational changes among polyamides are thought to have an impact on DNA affinity and sequence specificity. Thus, we have taken a more global view of molecular recognition, where our efforts have expanded from designing modules that distinguish the four DNA base-pairs (i.e., pairing rules) to designing those that target short, discrete DNA sequences.

Upon incorporation and evaluation of the G specific **Im** ring into the **Im**−**Hz** dimer cap (polyamide **1**) we were surprised to find that it failed to demonstrate any preference for its designed site in addition to displaying a significantly decreased affinity. The shortcomings of the **Im**−**Hz** dimer prompted a search for a ring system that was capable of specifying for G > C within the **X**−**Hz** context. The oxazole (**No**) cap (Figure 1) was considered because of its structural resemblance to **Im**—both rings present a nitrogen atom capable of hydrogen bonding to the minor groove. When the **No**−**Hz** dimer was incorporated into polyamide **2**, it was found to be specific for its designed sequence of 5′-T**GT**ACA-3′ with a 25-fold preference for G > C and an affinity of $K_a = 6.8 \times 10^9\ M^{-1}$ at its match site (Figure 3). The **No**−**Hz** dimer presents the same functionality to the minor groove as the **Im**−**Hz** dimer but with an enhanced ability to target a 5′-GT-3′ site, which could be due to a combination oxazole lone-pair basicity and differential solvation/desolvation effects between oxazole and imidazole.

The **Ct**−**Hz** dimer cap represents our first effort to target a short sequence of DNA using sequence-inspired recognition elements. Studies have shown that **Hz** exhibits specificity for T > A at the N-1 position—relative to the polyamide N-terminus—and that **Ct** polyamides exhibited specificity for T > A at the cap position with excellent polyamide affinities. We hoped that a hybrid dimer would impart specificity for the TT sequence while maintaining a biologically relevant affinity. Polyamide **3** bound its designed match site 5′-T**TT**ACA-3′ with an affinity of $K_a = 2.4 \times 10^9\ M^{-1}$ and a specificity of 10-fold for T·A over A·T. This result is attributed to the fact that both the sulfur and hydroxyl groups prefer to lie over the less-bulky thymine base in a T·A base pair and that the −OH of the **Hz** ring is able to form an energetically favorable hydrogen bond with the O(2) carbonyl of thymine.[10] Combined, these attributes make this dimer the preferred solution for targeting consecutive thymine residues.

As a first step in the design of programmable oligomers devoid of directly linked pyrrole- or imidazole-carboxamides, we incorporated the new **No**−**Hz** dimer into a hairpin structure, oligomer **9**, consisting only of 6−5 fused ring systems (Figure 2). The oligomer was designed to target the site 5′-GTAC-3′ and is a third generation molecule from our previously reported hairpin polyamides, which were shown to code for the four Watson−Crick base pairs in a sequence specific manner.[16] To evaluate the impact of removing four carboxamide linkages and moving to a system consisting of only 6−5 fused recognition elements, binding properties were evaluated using quantitative DNase I footprinting titrations. Oligomer **9** was found to bind its match site with an impressive affinity of $K_a = 2.3 \times 10^{10}\ M^{-1}$ while discriminating against its mismatch sites of 5′-GAAC-3′ and 5′-GATC-3′ with specificities (K_a match/K_a mismatch) of ∼7- and ∼ 23-fold, respectively (Figure 4). The four 6−5 fused rings of oligomer **9** present an "inside edge" with complimentary shape to the minor groove floor as the carboxamide linkages of traditional Py−Im−Hp polyamides. The complementary bumps and holes fit together between the oligomer and the DNA surface, which is the key to specificity. The dimer recognition elements in oligomer **9** are linked by a single carbon−carbon bond, resulting in fewer degrees of rotational freedom, which may result in a reduced entropic penalty for minor-groove binding. Undoubtedly, much of the favorable energetics for complexation with DNA for all these molecules is a result of differential solvation. The oligomer's large hydrophobic surface may result in increased favorable van der Waals interactions with the walls of the minor groove.

Conclusion

Hairpin polyamides containing the **No**−**Hz** and **Ct**−**Hz** dimer caps at the polyamide N-terminus are able to target 5′-GT-3′ and 5′-TT-3′ sequences with good affinity and specificity and represent new recognition elements for the minor groove of DNA. The **No**−**Hz** and **Ct**−**Hz** dimer caps represent attempts to broaden heterocycle designs beyond single base pair interactions. In addition, the development of the **No**−**Hz** cap has allowed for the design of a DNA binding molecule, which in a formal sense is no longer a polyamide, hence the term programmable oligomer. We are encouraged by the fact that this oligomer demonstrates excellent affinity for DNA while exhibiting good levels of specificity. We hope to apply these new heterocycles to the targeting of biologically relevant sequences in the context of integrating artificial transcription factors with living biological systems.

Acknowledgment. We thank The National Institutes of Health for grant support, Caltech for a James Irvine Fellowship to R.M.D., the Parsons Foundation Fellowship to M.A.M, and the NSF for a fellowship to S.F.

Supporting Information Available: Complete experimental details for the synthesis of building blocks (**7** and **8**) and polyamides (**1**−**3** and **9**) along with compound characterization data. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0621795

(17) Urbach, A. R.; Love, J. J.; Ross, S. A.; Dervan, P. B. *J. Mol. Biol.* **2002**, *320*, 55−71.