Appendix C: Programmable Oligomers for Minor Groove DNA Recognition

The text of this chapter was taken in part from a manuscript coauthored with Raymond M. Doss, Michael M. Marques, Shane Foister, and Peter B. Dervan (Caltech)*

(Doss, R.M., Marques, M.A., Foister, S., Chenoweth, D.M., and Dervan, P. B. *J. Am. Chem. Soc.*, **2006**, *128*, 9074-9079.)

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.

C.1 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 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 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 C.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 C.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 C.2). When positioned at the termini of hairpin polyamides, the No-Hz/Py-Py and Ct-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 pyrroleor imidazole-carboxamide based on the natural product distamycin, a design benchmark for biomimetic chemistry and the field of DNA recognition.

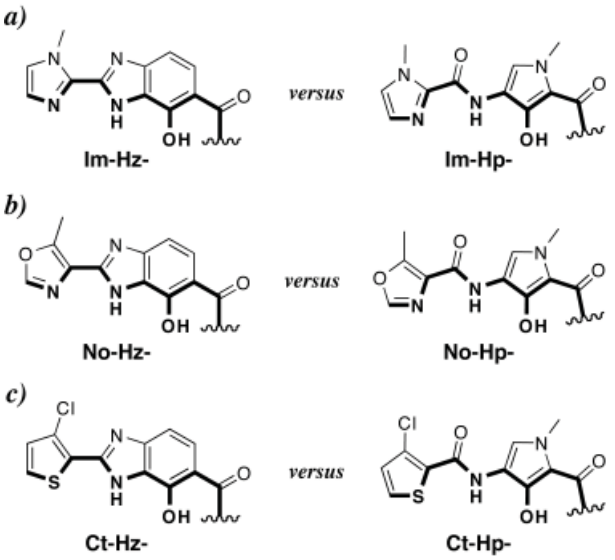


Figure C.1 Structures of dimers. a) imidazole-hydroxybenzimidazole (Im-Hz), b) oxazole-hydroxybenzimidazole (No-Hz), and c) chlorothiophenehydroxybenzimidazole (Ct-Hz) dimer caps in comparison with their respective five membered ring systems. Hydrogen-bonding surfaces to the DNA minor-groove floor are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

C.2 Experimental

C.2.1 Polyamide Synthesis

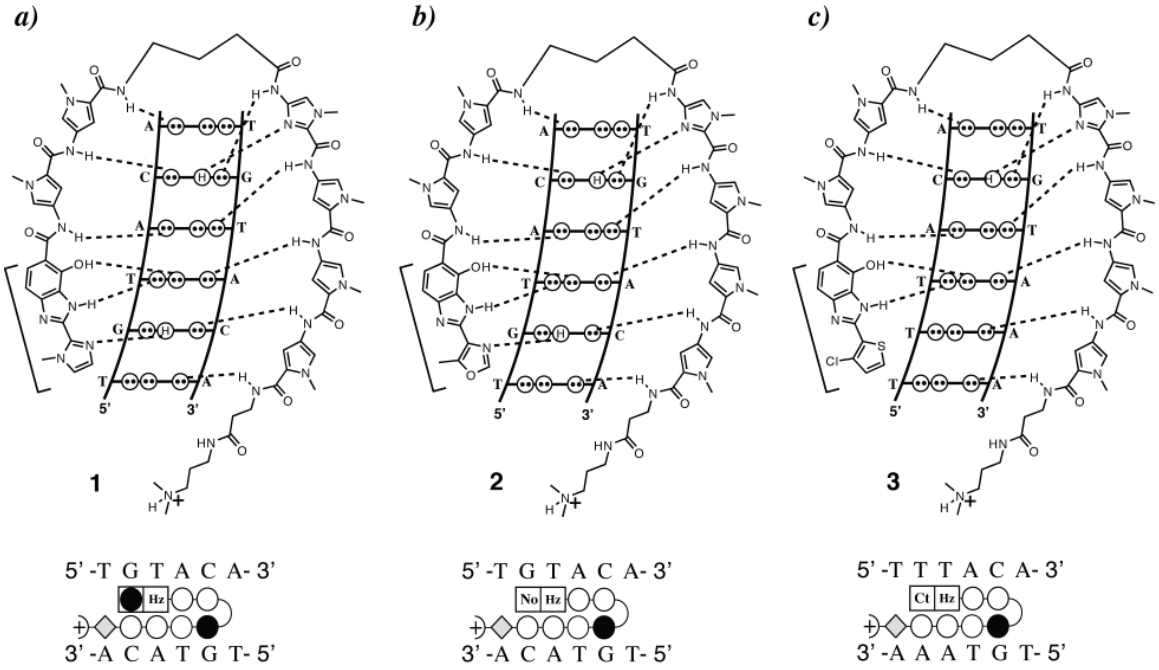


Figure C.2 Postulated hydrogen-bonding models for the 1:1 polyamide-DNA complexes with their matched sequence and their ball-and-stick representations. a) Im-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**1**), b) No-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**2**), and c) Ct-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**3**).

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 C.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-HzOMe-OH (**7**) and Ct-HzOMe-OH (**8**) are detailed in Section C.7 Supplemental 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 Section C.7 Supplemental Information for full experimental details.)

C.3 Results

C.3.1 DNA Affinity and Sequence Specificity of Dimer Caps

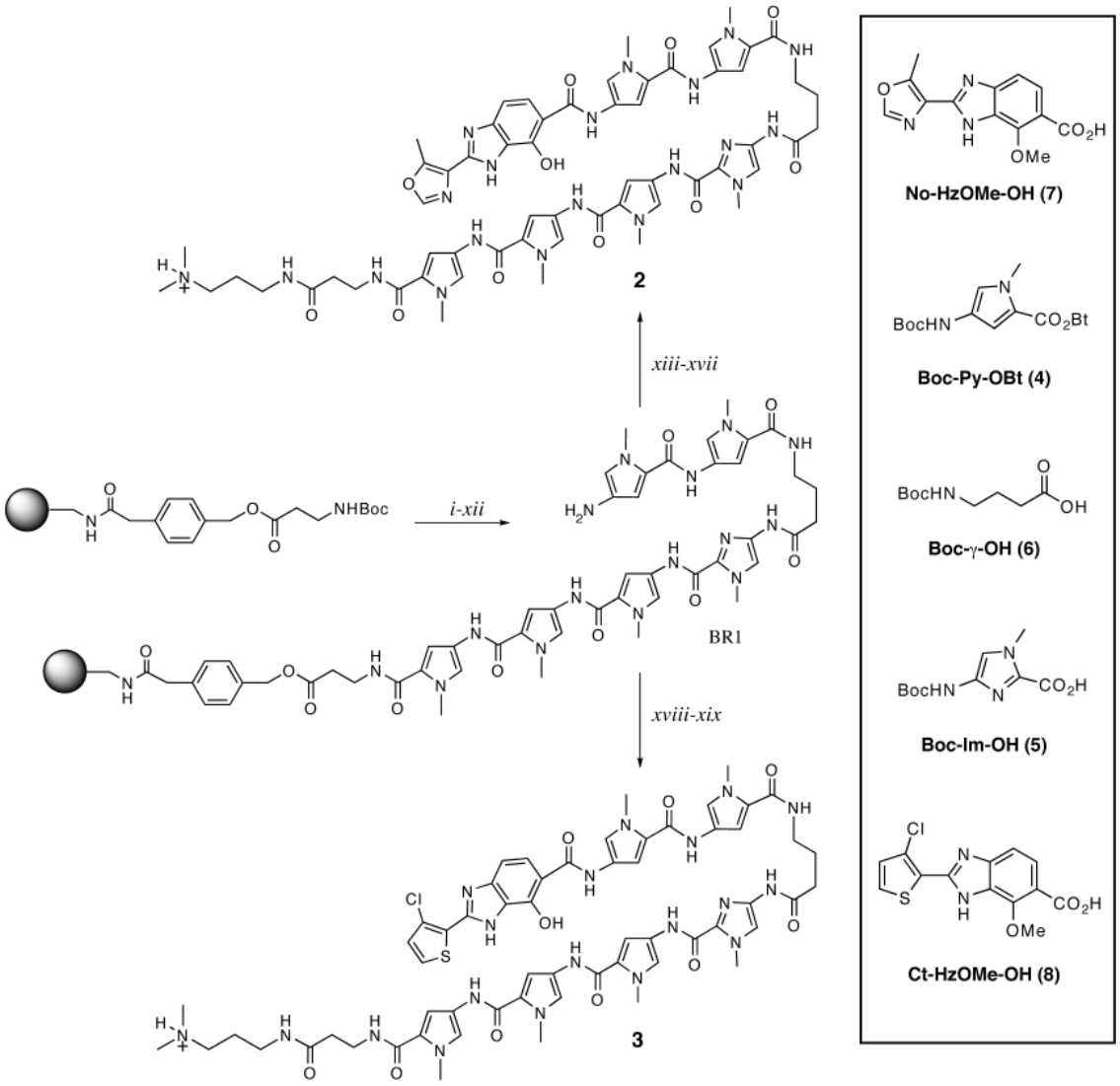
Quantitative DNase I footprinting titrations were carried out for polyamides **1**-**3**. All polyamides were footprinted on the 285-base-pair PCR product of plasmid pCW15. In all cases, the DNA-sequence specificity at the cap position (in bold) was determined by varying a single DNA base pair within 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**XT**ACA-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 x 10^9 M^{-1}) (Figure C.3, Table C.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 x 10^8 and 4.0 x 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 x 10^9 M^{-1} (Figure C.3, Table C.1). With the development of the chlorothiophene and oxazole dimer caps, the range of targetable sequences by polyamides has been expanded (Table C.1).

C.3.2 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 x 10^{10} M^{-1} for the



Scheme C.1 Representative solid-phase synthesis of polyamide **2** and **3** along with a table of the amino acid building blocks used. 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 x 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.

match site 5'-GTAC-3' and affinities of K_a = 3.5 x 10^9 and 9.8 x 10^8 M^{-1} for the mismatch sites 5'-GAAC-3' and 5'-GATC-3', respectively (Figure C.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.

C.4 Discussion

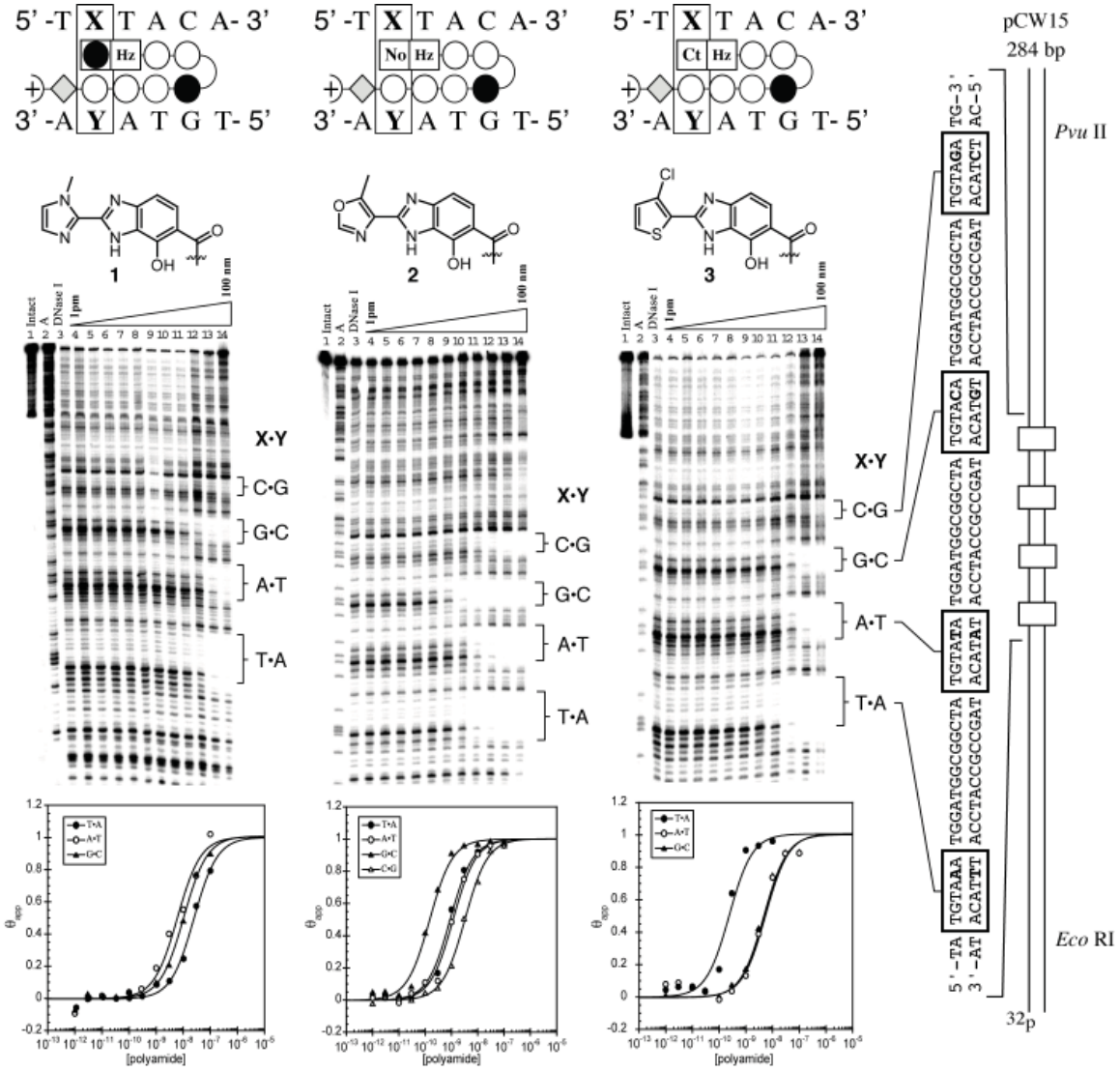
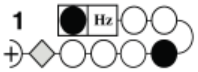
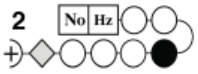


Figure C.3 Quantitative DNase I footprinting experiments in the hairpin motif for polyamides **1**, **2**, and **3**, respectively, on the 278 bp, 5'-end-labeled PCR product of plasmid CW15: (lane 1) intact DNA; (lane 2) A reaction; (lane 3) DNase I standard; (lanes 4-14) 1 pM, 3 pM, 10 pM, 30 pM, 100 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.

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 microstructure of DNA depends on the sequence in question.[17] In addition, structural and conformational changes among polyamides are thought to have an impact

Table C.1 Affinities of X/Py ring pairs proximal to a hydroxybenzimidazole bicycle K_a (M^{-1}).[a,b]

Polyamide	5'-t**A**TACa-3'	5'-t**T**TACa-3'	5'-t**G**TACa-3'	5'-t**C**TACa-3'
1 	$9.7(\pm0.7)\times10^7$	$4.5(\pm0.6)\times10^8$	$1.7(\pm0.4)\times10^8$	$\leq 1.0 \times 10^7$
2 	$8.6(\pm0.3)\times10^8$	$9.5(\pm0.3)\times10^8$	$6.8(\pm0.4)\times10^9$	$2.7(\pm0.5)\times10^8$
3 	$2.1(\pm0.3)\times10^8$	$2.4(\pm0.2)\times10^9$	$2.6(\pm0.4)\times10^8$	$\leq 1.0 \times 10^7$

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 5mM $CaCl_2$ at pH 7.0.

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 C.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'-TGTACA-3' with a 25-fold preference for G > C and an affinity of $K_a = 6.8\ \ 10^9\ M^{-1}$ at its match site (Figure C.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 positions 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 $K_a = 2.4$ x

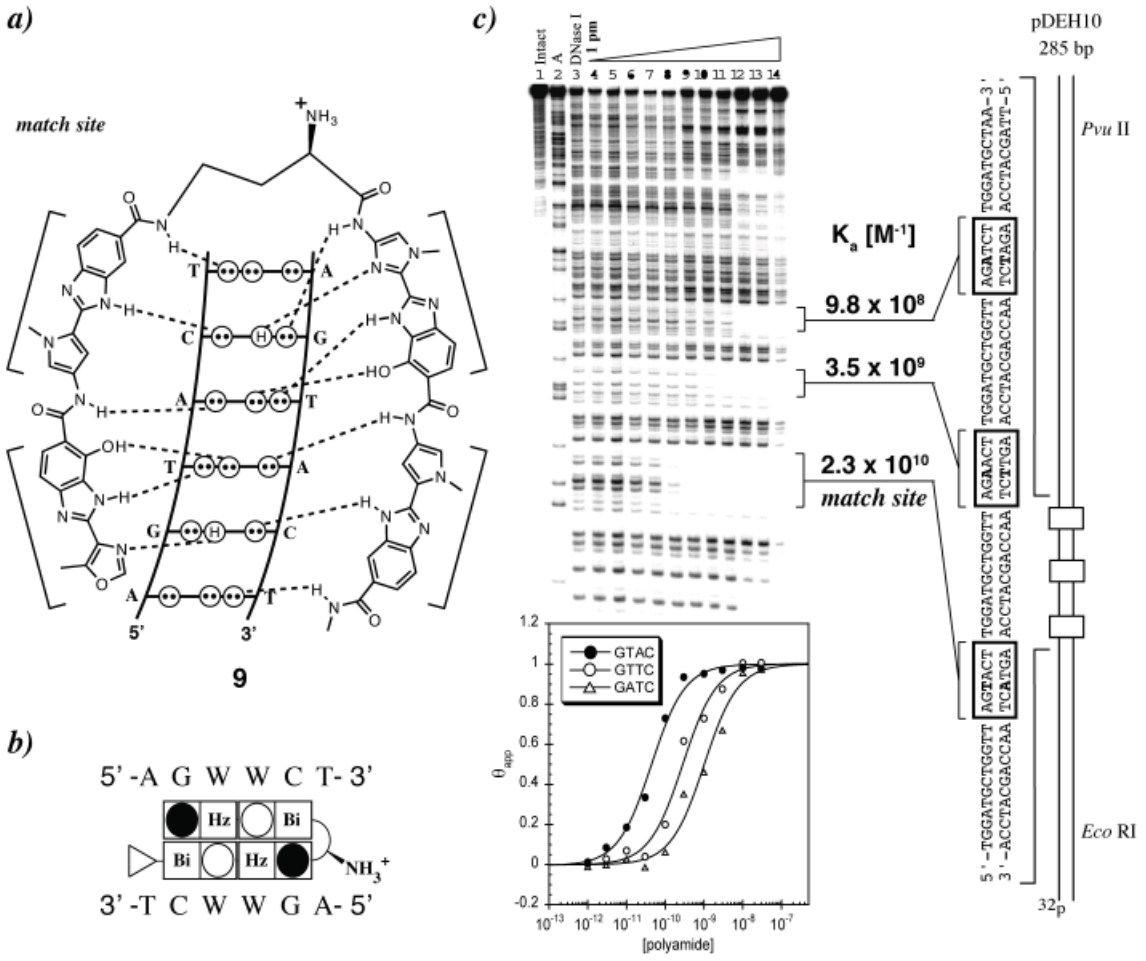


Figure C.4 a) Postulated hydrogen-bonding model and structure of oligomer **9**. b) Ball and stick representation of **9** and the 6-base-pair binding site with variable region (W = A or T) shown. c) Quantitative DNase I footprint titration experiment on the 5'-^{32}P-labeled PCR product shown with an illustration and complete sequence of the 285 bp *Eco*RI/P*vu*II restriction fragment from plasmid pDEH10. Binding affinities are shown next to their respective binding sites and the match site is designated.sites.

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 C.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 C.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.

C.5 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.

C.6 Notes and References

1. Darnell, J. E. Transcription factors as targets for cancer therapy. *Nat. Rev. Cancer* **2002**, *2*, 740-748. (b) Pandolfi, P. P. Transcription therapy for cancer. *Oncogene* **2001**, *20*, 3116-3127. (c) Dervan, P. B; Edelson, B. S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284-299.

2. (a) Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X. Netropsin, A new antibiotic produced by a streptomyces. *J. Am. Chem. Soc.* **1951**, *73*, 341-343. (b) Arcamone, F. N. V.; Penco,

S.; Orezzi, P.; Nicolella, V.; Pirelli, A. Structure and synthesis of Distamycin A. *Nature* **1964**, *203*, 1064-1065.

3. (a) Kopka, M. L.; Yoon, C.; Goodsell, D.; Pjura, P.; Dickerson, R. E. The molecular origin of DNA-drug specificity in netropsin and distamycin. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 1376-1380. (b) Pelton, J. G.; Wemmer, D. E. Structural characterization of a 2:1 distamycin A.d(CGCAAATTGGC) complex by two-dimensional NMR. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 5723-5727. (c) Wemmer, D. E. Designed sequence-specific minor groove ligands. *Annu. Rev. Biophys. 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. 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.* **2004**, *127*, 742-750. (e) Baraldi, P. G.; Bovero, A.; Fruttarolo, F.; Preti, D.; Tabrizi, M. A.; Pavani, M. G.; Romagnoli, R. DNA minor groove binders as potential antitumor and antimicrobial agents. *Med. Res. Rev.* **2004**, *24*, 475-528. (f) Reddy, P. M.; Toporowski, J. W.; Kahane, A. L.; Bruice, T. C. Recognition of a 10 base pair sequence of DNA and stereochemical control of the binding affinity of chiral hairpin polyamide-Hoechst 33258 conjugates. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 5531-5536.

4. White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. Recognition of the four Watson-Crick base pairs in the DNA minor groove by synthetic ligands. *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. A structural basis for recognition of A•T and T•A base pairs in the minor groove of B-DNA. *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. Structural effects of DNA sequence on TA recognition by hydroxypyrrole/pyrrole pairs in the minor groove. *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. Sequence selectivity of 3-hydroxypyrrole/pyrrole ring pairings in the DNA minor groove. *J. Am. Chem. Soc.* **1999**, *121*, 11621-11629. (b) White, S.; Turner, J. M.; Szewezyk, J. W.; Baird, E. E.; Dervan, P. B. Affinity and specificity of multiple hydroxypyrrole/pyrrole ring pairings for coded recognition of DNA. *J. Am. Chem. Soc.* **1999**, *121*, 260-261.

7. (a) Hays, F. A.; Teegarden, A.; Jones, Z. J. R.; Harms, M.; Raup, D.; Watson, J.; Cavaliere, E.; Shing Ho, P. How sequence defines structure: a crystallographic map of DNA structure and conformation. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7157-7162. (b) Beveridge, D. L.; Barreiro, G.; Byun, K. S.; Case, D. A.; Cheatham 3rd, T. E.; 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. Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. I. Research design and results on d(CpG) steps. *Biophys. J.* **2004**, *87*, 3799–3813. (c) Dixit, S. B.; Beveridge, D. L.; Case, D. A.; Cheatham 3rd, T. E.; Giudice, E.; Lankas, F.; Lavery, R.; Maddocks, J. H.; Osman, R.; Sklenar, H.; Thayer, K. M.; Varnai, P. Molecular dynamics simulations of the 136 unique tetranucleotide sequences of DNA oligonucleotides. II: sequence context effects on the dynamical structures of the 10 unique dinucleotide steps. *Biophys. J.* **2005**, *89*, 3721–3740. (d) Wu, H.; Crothers, D. M. The locus of sequence-directed and protein-induced DNA bending *Nature* **1984**, *308*, 509. (e) Steitz, T. A. Structural studies of protein-nucleic acid interaction: the sources of sequence-specific binding. *Q. Rev. Biophys.* **1990**, *23*, 205. (f) Goodsell, D. S.; Kopka, M. L.; Cascio, D.; Dickerson, R. E. Crystal structure of CATGGCCATG and its implications for A-tract bending models. *Proc. Natl. Sci. U.S.A.* **1993**, *90*, 2930. (g) Paolella, D. N.; Palmer, R.; Schepartz, A. DNA targets for certain bZIP proteins distinguished by an intrinsic

bend. *Science* **1994**, *264*, 1130. (h) Kahn, J. D.; Yun, E.; Crothers, D. M. Detection of localized DNA flexibility. *Nature* **1994**, *368* 163. (h) Geierstanger, B. H.; Wemmer, D. E. Complexes of the minor groove of DNA. *Annu. Rev. Biophys. Biochem. Struct.* **1995**, *24*, 463. (i) Hansen, M. R.; Hurley, L. H. Pluramycins. Old Drugs Having Modern Friends in Structural Biology. *Acc. Chem. Res.* **1996**, *29*, 249. (j) Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. Aliphatic/aromatic amino acid pairings for polyamide recognition in the minor groove of DNA. *J. Am. Chem. Soc.* **1998,** *120*, 6219-6226. (k) Marques, M. A.; Doss, R. M.; Urbach, A. R.; Dervan, P. B. Toward an understanding of the chemical etiology for DNA minor-groove recognition by polyamides. *Helv. Chim. Acta* **2002,** *85*, 4485-4517.

8. Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B. Shape selective secognition of T•A base pairs by hairpin polyamides containing N-terminal 3-methoxy (and 3-chloro) thiophene residues. *Bioorg. Med. Chem.* **2003,** *11*, 4333-4340.

9. Briehn, C. A.; Weyermann, P.; Dervan, P. B. Alternative heterocycles for DNA recognition: the benzimidazole/imidazole pair.*Chem. Eur. J.* **2003,** *9*, 2110-2122.

10. (a) Renneberg, D.; Dervan, P. B. Imidazopyridine/Pyrrole and hydroxybenzimidazole/pyrrole pairs for DNA minor groove recognition. *J. Am. Chem. Soc.* **2003,** *125*, 5707-5716. (b) Marques, M. A.; Doss, R. M.; Foister, S.; Dervan, P. B. Expanding the repertoire of heterocycle ring pairs for programmable minor groove DNA recognition. *J. Am. Chem. Soc.* **2004,** *126*, 10339-10349.

11. For down regulation 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. Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *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.

13. Baird, E. E.; Dervan, P. B. Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **1996,** *118*, 6141-6146.

14. Urbach, A. R.; Dervan, P. B. Toward rules for 1: 1 polyamide: DNA recognition. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** *98*, 4343-4348.

15. Trauger, J. W.; Dervan, P. B. Footprinting methods for analysis of pyrrole-imidazole polyamide/ DNA complexes. *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 x 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 x 10^8 M^{-1} where W = A or T[10b]

17. Urbach, A. R. Love, J. J.; Ross, S. A.; Dervan, P. B. Structure of a beta-alanine-linked polyamide bound to a full helical turn of purine tract DNA in the 1: 1 motif. *J. Mol. Biol.* **2002,** *320*, 55-71.

18. Suzuki, M.; Iwasaki, T.; Miyoshi, M.; Okumura, K.; Matsumoto, K. Synthesis of amino acids and related compounds. 6. New convenient synthesis of alpha-C-acylamino acids and. alpha.-amino ketones. *J. Org. Chem.* **1973**, *38*, 3571-3575.

C.7 Supplemental Information

C.7.1 General

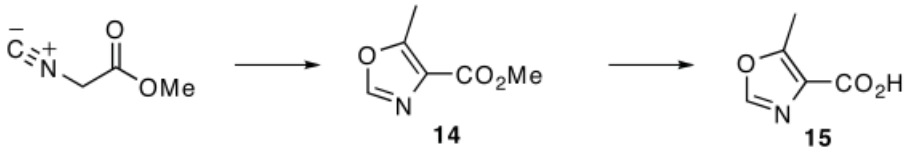
N,*N*-dimethylformamide (DMF), *N*,*N*-diisopropylethylamine (DIEA), thiophenol (PhSH), 3-dimethylamino-1-propylamine (Dp), Triethylamine (TEA), and thiourea were purchased from Aldrich. Boc-β-alanine-(4-carbonylaminomethyl)benzyl-ester-copoly(styrenedivinylbenzene) resin (Boc-β-Pam-resin), dicyclohexylcarbodiimide (DCC), hydroxybenzotriazole (HOBt), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU), 4-dimethylaminopyridine (DMAP), and Boc-β-alanine were purchased from NOVA Biochem. Trifluoroacetic acid (TFA) was purchased from Halocarbon. All other solvents were reagent grade from EMD Chemicals. Oligonucleotide inserts were synthesized by the Biopolymer Synthesis Center at the California Institute of Technology. Precoated silica gel plates 60 F_{254} for TLC and silica gel 60 (40 μm) for flash chromatography were from Merck. Glycogen (20 mg/mL), dNTPs (PCR nucleotide mix), and all enzymes, unless otherwise stated, were purchased from Boehringer-Mannheim. pUC19 was purchased from New England Biolabs, and deoxyadenosine [γ-^{32}P] triphosphate was provided by ICN. Calf thymus DNA (sonicated, deproteinized) and DNaseI (7500 units/mL, FPLC pure) were from Amersham Pharmacia. AmpliTaq DNA polymerase was from Perkin-Elmer and used with the provided buffers. Tris.HCl, DTT, RNase-free water, and 0.5 M EDTA were from United States Biochemical. Calcium chloride, potassium chloride, and magnesium chloride were purchased from Fluka. Tris-borate-EDTA was from GIBCO, and bromophenol blue was from Acros. All reagents were used without further purification. NMR spectra were recorded on a Varian spectrometer at 300 MHz in DMSO-d_6 or $CDCl_3$, with chemical shifts reported in parts per million relative to residual solvent. UV spectra were measured on a Hewlett-Packard model 8452A diode-array spectrophotometer. High-resolution FAB and EI mass spectra were recorded at the Mass Spectroscopy Laboratory at the California Institute of Technology. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) was conducted at the Mass Spectroscopy Laboratory at the California Institute of Technology.

C.7.2 Heterocycle Synthesis

The synthesis of compounds **4** (Boc-Py-OBt), **5** (Boc-Im-OH), **10** (Boc-Py-Bi-OH), **11** (Boc-Im-HzOMe-OH), **19** (Ct-OH), **10**, **22** (Im-HzOMe-OH), and **16** (aryl diamine) have previously been reported.[8,10a]

*C.7.2.1 Methyl 5-methyl-1,3-oxazole-4-carboxylate (No-OMe) (**14**)*

Compound **14** was prepared following an exact preparation reported by Suzuki et al.[18] and purified



Scheme C.2 Synthesis of **15**.

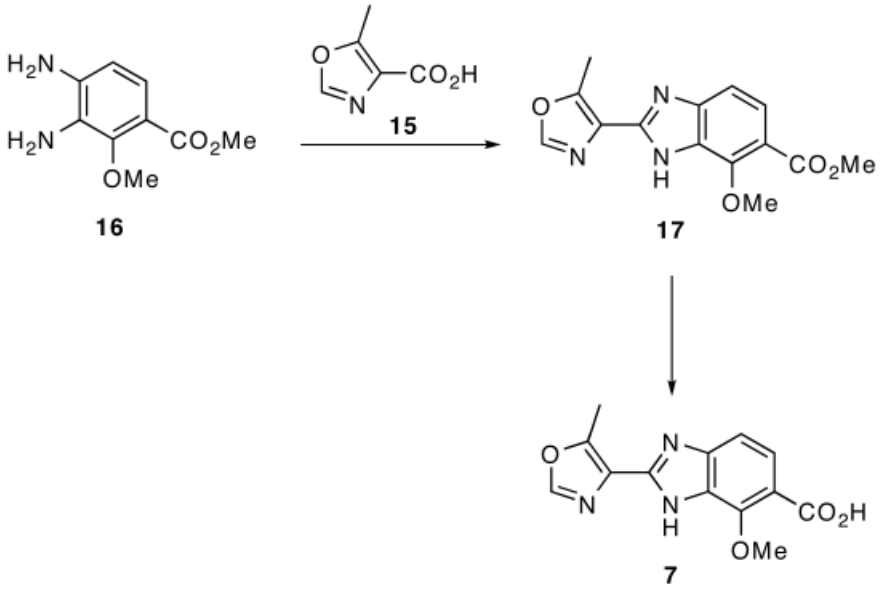
by normal phase column chromatography. According to the Suzuki procedure a mixture of methyl-alpha-isocyanoacetate (3.0 g, 30.2 mmol) and DBU (4.5 g, 30.2 mmol) in dry THF (40 mL) at 10°C, was treated with acetic anhydride (3.06 g, 30 mmol) in dry THF (10 mL) dropwise. The reaction was allowed to warm to room temperature and stirred for 10 h. The solvent was removed by rotoevaporation and water (100 mL) was added. The mixture was extracted with EtOAc (2 x 100 mL). The organic layers were combined, dried over sodium sulfate, filtered, and concentrated in vacuo to provided **14** as a crude amber oil. The oil was subjected to column chromatography using (3:2 Hex/EtOAc) to provide **14** (2.73 g, 64% Yield) as a crystalline white solid. TLC (3:2 Hex/EtOAc) R_f 0.5; ^{1}H NMR (DMSO-d_6) 8.32 (s, 1H), 3.77 (s, 3H), 2.55 (s, 3H); ^{13}C (DMSO-d_6) 161.9, 156.1, 150.3, 126.2, 51.5, 11.5; HR-MS (EI): calculated for $C_6H_7NO_3$: 141.0426; found: 141.0427.

*C.7.2.2 5-Methyl-1,3-oxazole-4-carboxylic acid (No-OH) (**15**)*

A mixture of **14** (1 g, 7.08 mmol), 1N NaOH (10 mL) and MeOH (5 mL) was stirred at 40°C for 4 h. The MeOH was removed by rotoevaporation and the pH adjusted to pH = 2 with 1N HCl. The precipitate was extracted with EtOAc (3 x 10 mL), the organics dried over sodium sulfate and removed by rotoevaporation to provide **15** (738 mg, 82% Yield) as a fibrous white solid. TLC (3:2 EtOAc/Hex +10% AcOH) R_f 0.4; ^{1}H NMR (DMSO-d_6) 8.27 (s, 1H), 2.45 (s, 3H); ^{13}C (DMSO-d_6) 163.0, 155.6, 150.1, 145.6, 11.6; HR-MS (EI): calculated for $C_5H_5NO_3$: 127.0269; found: 127.0268.

*C.7.2.3 Methyl 7-methoxy-2-(5-methyl(1,3-oxazole-4-yl))benzimidazole-6-carboxylate (NoHz(OMe)OMe) (**17**)*

To a solution of **15** (0.3 g, 2.36 mmol) in DMF (4 mL) was added DIEA (915 mg, 1.23 mL, 7.08 mmol) and HBTU (895 mg, 2.36 mmol). The mixture was stirred at room temperature for 1 h, followed by the addition of the aryl diamine **16** (463 mg, 2.36 mmol). The reaction was then heated to 35°C and stirred for an additional 24 h. The reaction was allowed to cool to room temperature and then poured into a separatory funnel containing water (200 mL). The water was then extracted with



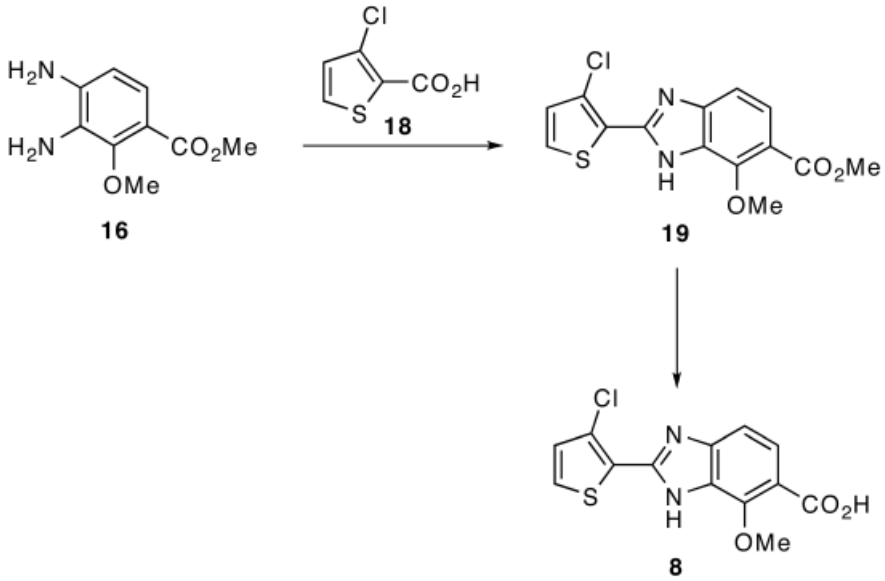
Scheme C.3 Synthesis of **7**.

EtOAc (2 x 100 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated in vacuo to provide a crude solid, which was dissolved in glacial acetic acid (5 mL) and heated to 90°C. The reaction was stirred for 12 h, followed by removal of the solvent by rotoevaporation. The resultant residue was subjected to column chromatography (4:1 EtOAc/Hex) to provide **17** as a thin film. Addition of hexanes to the film, followed by rotoevaporation and drying under high vacuum provided **17** as a white solid (379 mg, 56% Yield). TLC (4:1 EtOAc/Hex) R_f 0.75; ^{1}H NMR (DMSO-d_6) 8.49 (s, 1H), 7.51 (d, 1H, J = 8.7 Hz), 7.16 (d, 1H, J = 8.7 Hz), 4.34 (s, 3H), 3.77 (s, 3H) 2.78 (s, 3H); ^{13}C (DMSO-d_6) 166.7, 151.0, 149.3, 145.5, 139.0, 135.3, 125.6, 124.8, 114.2, 105.3, 60.8, 51.6, 11.3; HR-MS (EI): calculated for $C_{14}H_{13}N_3O_4$: 287.0906; found: 287.0914.

C.7.2.4 7-methoxy-2-(5-methyl(1,3-oxazole-4-yl))benzimidazole-6-carboxylic acid (NoHz(OMe) OH) (7)

A mixture of **17** (200 mg, 0.69 mmol), DMSO (0.8 mL), and 1N NaOH (4 mL) was stirred at 23 °C for 3 h. The solution was diluted with 8 mL of H_2O and the pH was adjusted to pH = 2 with 1N HCl (4.4 mL). The precipitate was filtered and washed with diethyl ether. The crude material was purified by silica gel chromatography (1:1 EtOAc/Hexanes + 2% AcOH to 3:1 EtOAc/Hexanes + 2% AcOH) to provide **7** (156 mg, 82% Yield) as a white solid after drying under high vacuum. TLC (3:1 EtOAc/Hex +2% AcOH) R_f 0.55; ^{1}H NMR (DMSO-d_6) 13.12 (s, 1H), 12.37 (s, 1H), 8.51 (s, 1H), 7.52 (d, 1H, J = 8.5 Hz), 7.15 (d, 1H, J = 8.5 Hz), 4.34 (s, 3H), 2.78 (s, 3H); ^{13}C (DMSO-d_6) 167.8, 151.1, 150.9, 149.3, 145.5, 139.0, 135.5, 125.8, 125.2, 115.2, 105.4, 60.9, 11.4; HR-MS

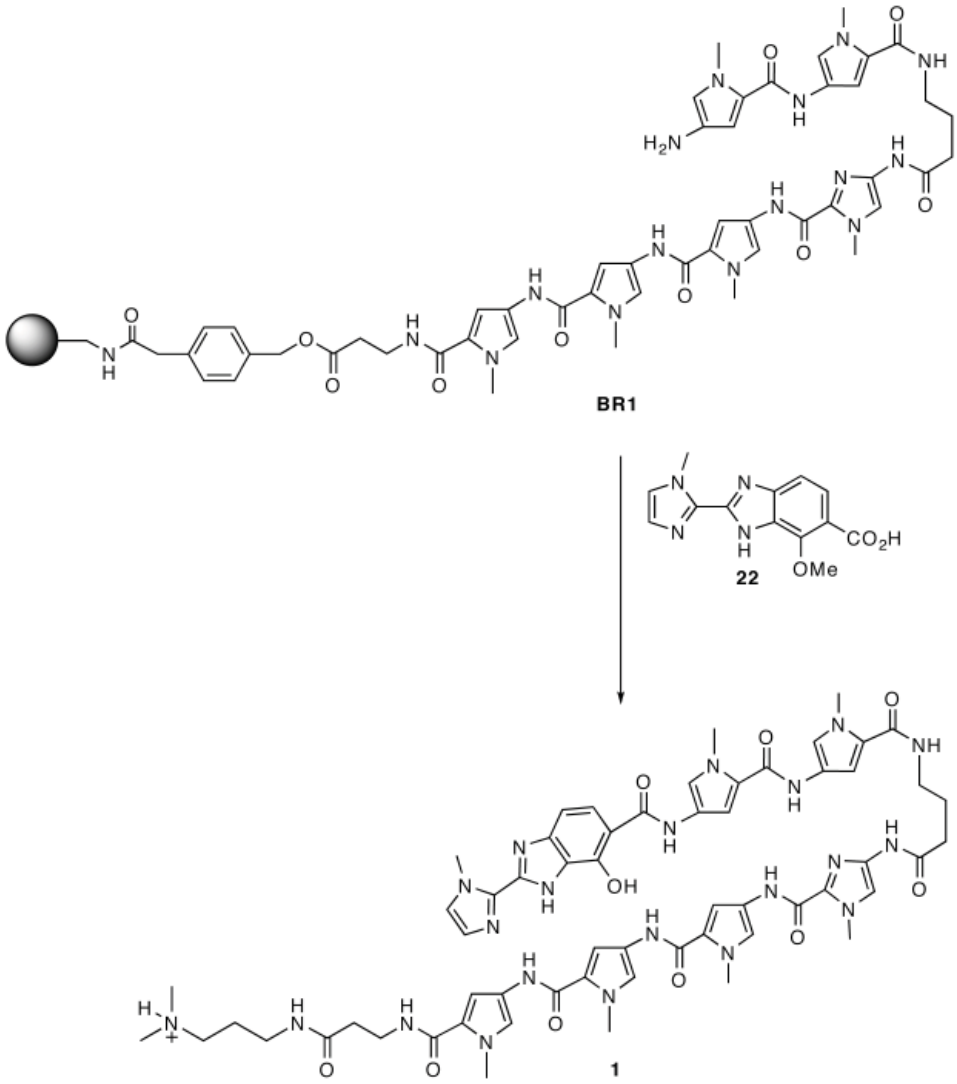
(FAB): calculated for $C_{13}H_{12}N_3O_4$: 274.0828; found: 274.0833 $[M+H]^+$.



Scheme C.4 Synthesis of **8**.

*C.7.2.5 Ct-Hz(OMe)OMe (**19**)*

To a solution of **18** (294 mg, 1.80 mmol) in DMF (5 mL) was added DIEA (0.348 mL, 1.98 mmol) and HBTU (650 mg, 1.72 mmol). The mixture was stirred at room temperature for 1 h, followed by the addition of the aryl diamine **16** (300 mg, 1.80 mmol). The reaction was then heated to 35°C and stirred for an additional 24 h. The reaction was allowed to cool to room temperature and then poured into a Falcon tube containing cold water (40 mL) resulting in a cloudy precipitate. The Falcon tube was centrifuged at 14000 rpm for 10 min, the mother liquor decanted, and the precipitate dried under high vacuum. After drying, the crude solid was dissolved in glacial acetic acid (5 mL) and heated to 90 °C. The reaction was stirred for 12 h, followed by removal of the solvent by rotoevaporation. The resultant residue was subjected to column chromatography (3:2 EtOAc/Hex) to provide **19** as an off-white solid (647 mg, 57% Yield). TLC (3:2 EtOAc/Hex) R_f 0.5; ^{1}H NMR (DMSO-d_6) 12.85 (s, 1H), 7.89 (d, 1H, J = 5.3 Hz), 7.56 (d, 1H, J = 8.5 Hz), 7.27 (d, 1H, J = 8.2), 7.26 (d, 1H, J = 5.3), 4.30 (s, 3H), 3.80 (s, 3H); ^{13}C (DMSO-d_6) 166.6, 151.2, 144.0, 139.8, 134.8, 129.7, 129.4, 126.3, 125.6, 123.2, 114.8, 106.1, 61.2, 51.8; HR-MS (EI): calculated for $C_{14}H_{11}N_2O_3SCl$: 322.0179; found: 322.0171.

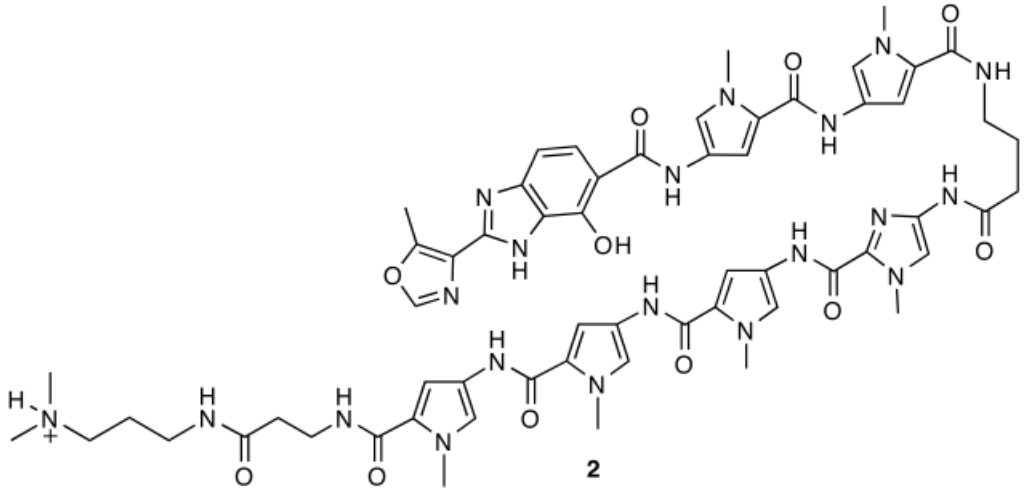


Scheme C.5 Synthesis of **1**.

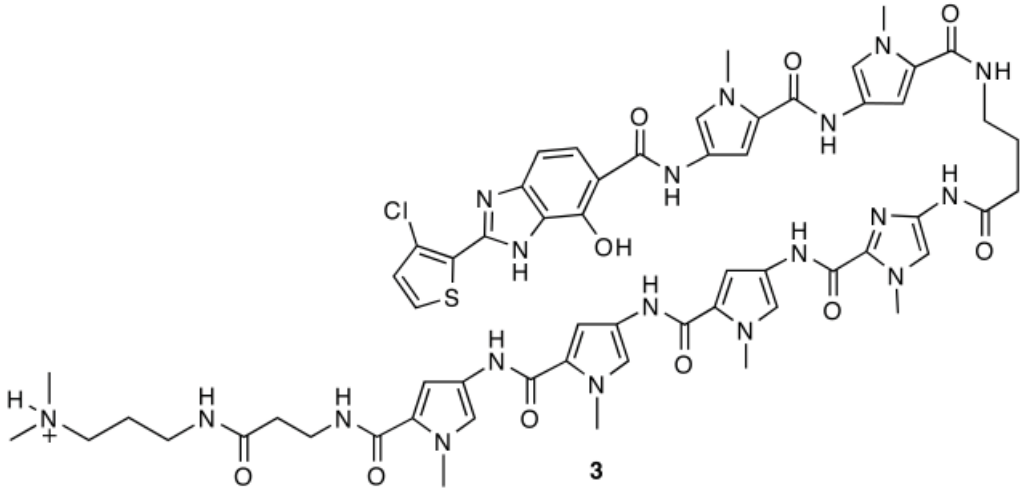
C.7.2.6 Ct-Hz(OMe)OH (8)

A mixture of **19** (1.8 g, 5.58 mmol), 1N NaOH (36 mL) and MeOH (27 mL) was stirred at 45°C for 24 h. The MeOH was removed by rotoevaporation and the pH adjusted to pH = 5.5 with 1N HCl, resulting in a precipitate. The mixture was spun down in a centrafuge (5 min at 14,000 rpm). The tube was decanted leaving a white solid that was dried under high vacuum to provide **8** (2.6 g, Quant. Yield) as a light yellow solid. TLC (3:2 EtOAc/Hex + 10% AcOH) R_f 0.5; ^{1}H NMR (DMSO-d_6) 13.00 (s, 1H), 7.89 (d, 1H, J = 5.3 Hz), 7.57 (d, 1H, J = 8.5 Hz), 7.30 (d, 1H, J = 8.5), 7.25 (d, 1H, J = 5.3 Hz), 4.28 (s, 3H); ^{13}C (DMSO-d_6) 167.7, 151.1, 143.9, 139.7, 134.9, 129.6, 129.3, 126.4, 125.7, 123.2, 115.9, 106.1, 61.2; HR-MS (EI): calculated for $C_{13}H_9N_2O_3SCl$: 308.0022; found: 308.0033.

C.7.3 Polyamide Synthesis



Scheme C.6 Compound **2**.



Scheme C.7 Compound **3**.

Polyamides were synthesized from Boc-β-alanine-Pam resin (50 mg, 0.59 mmol/g) and purified by preparatory HPLC according to published manual solid phase protocols.[4] The synthesis of batch resin BR1 (H_2N-Py-Py-γ-Im-Py-Py-Py-β-Pam) has previously been reported.[8]

*C.7.3.1 Im-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**1**)*

22 (Im-HzOMe-OH) (25 mg, 88.5 μmol) was incorporated by activation with HBTU (32 mg, 84 μmol), DIEA (23 mg, 31 μl, 177 μmol) and DMF (250 μl). The mixture was allowed to stand for

15 min at room temperature and then added to the reaction vessel containing base resin BR1 H_2N-Py-Py-γ-Im-Py-Py-Py-β-Pam. Coupling was allowed to proceed for 12 h at room temperature. The resin-bound polyamide was then washed with DCM and treated as described in the deprotection protocol below to provide Im-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**1**) (0.9 mg, 2.4 % recovery) as a fine white powder under lyophilization of the appropriate fractions. MALDI-TOF-MS: calculated for $C_{59}H_{70}N_{21}O_{10}$: 1232.56; found 1232.55 $[M+H]^+$.

*C.7.3.2 No-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**2**)*

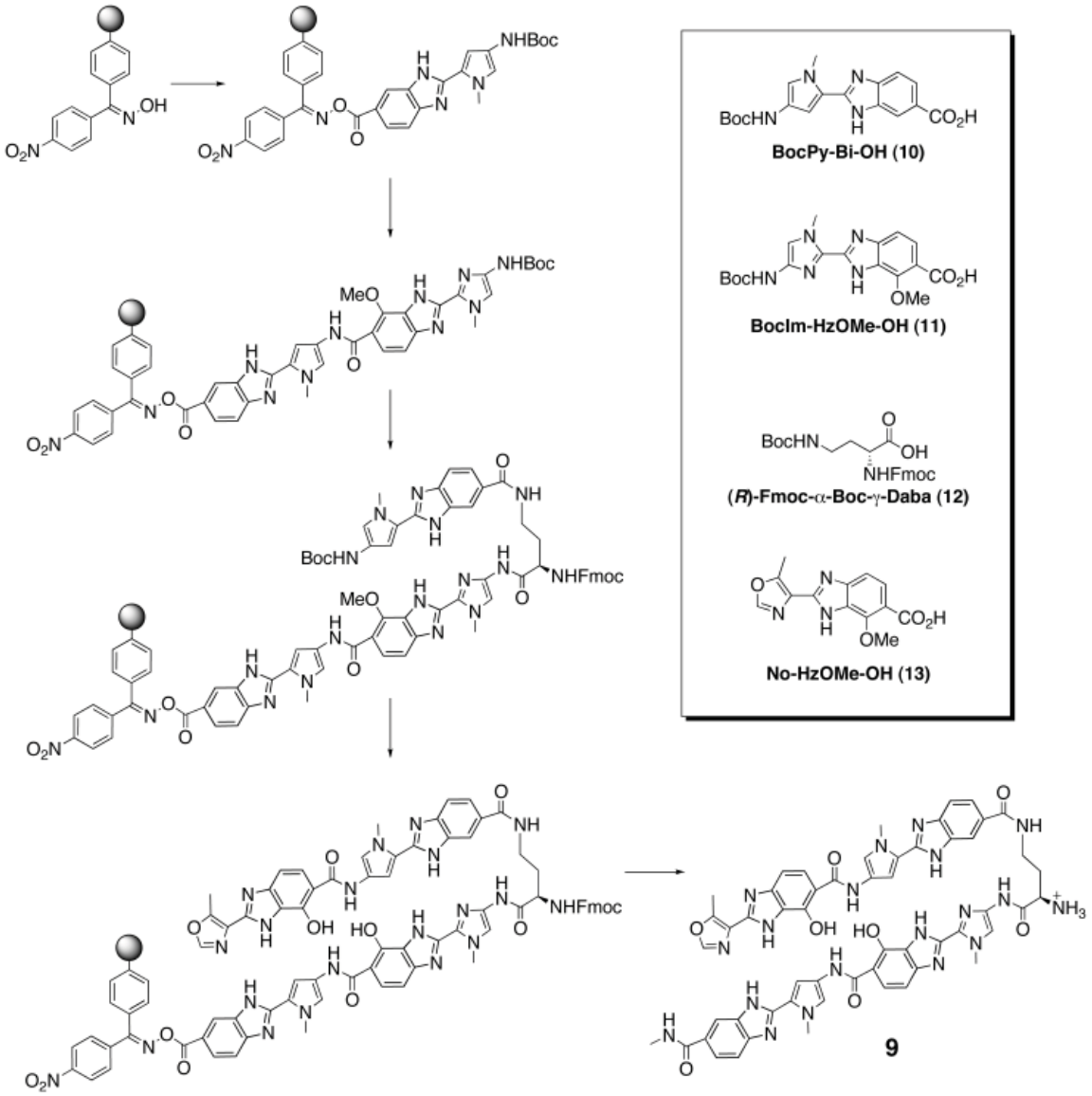
7 (No-HzOMe-OH) was incorporated as described above for Im-Hz-OH (polyamide **1**) to provide No-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**2**) (1.5 mg, 4.1 % recovery) as a fine white powder under lyophilization of the appropriate fractions. MALDI-TOF-MS: calculated for $C_{59}H_{69}N_{20}O_{11}$: 1233.55; found 1233.54 $[M+H]^+$.

*C.7.3.3 Ct-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**3**)*

8 (Ct-HzOMe-OH) was incorporated as described above for Im-Hz-OH (polyamide **1**) to provide Ct-Hz-Py-Py-γ-Im-Py-Py-Py-β-Dp (**3**) (1.1 mg, 2.9 % recovery) as a fine white powder under lyophilization of the appropriate fractions. MALDI-TOF-MS: calculated for $C_{59}H_{67}ClN_{19}O_{10}S$: 1269.47; found 1269.47 $[M+H]^+$.

C.7.4 Deprotection of the O-Methyl-Protected Polyamides

O-Methyl protected polyamides were cleaved from resin, purified, deprotected and subject to further purification using the following general procedure. Upon completion of solid phase synthesis, 3-dimethylamino-1-propylamine (Dp) (500 μL) was added to the synthesis vessel containing the resin (50 mg). The mixture was allowed to stand for 2 h at 85°C with occasional agitation. The resin was then filtered and the solution diluted to 8 mL using 0.1% TFA. The sample was purified by reversed phase HPLC and lyophilized to provide polyamides containing the *O*-methyl protected hydroxybenzimidazole unit (-HzOMe-) as a dry solid. The polyamides were then dissolved in DMF (200 μl) and added to a suspension of sodium hydride (40 mg, 60% oil dispersion) and thiophenol (200 μl) in DMF (400 μl) that was pre-heated for 5 min at 85°C. The mixture was heated for 2 h at 85°C. The mixture was then cooled to 0 °C and 20% TFA (7.0 mL) was added. The aqueous layer was washed three times with diethyl ether (8 mL) and then diluted to a total volume of 9.5 mL using 0.1% TFA. The mixture was then purified by reverse-phase HPLC to give the deprotected



Scheme C.8 Synthesis of **9**-**13**.

Hz-containing polyamides **1, 2,** and **3**.

*C.7.5 Oligomer **9** Synthesis*

Oligomer **9** was prepared using Kaiser oxime resin (0.65 meqiv/gram) from Nova Biochem. Oxime resin (0.1 grams), was added to a solid phase synthesis vessel. The resin was then washed with DCM (15 mL) followed by DMF (15 mL). In a separate vessel was added Boc-PyBi-OH (10) (60 mg, 0.168 mmol), HBTU (61 mg, 0.159 mmol), DIEA (43 mg, 60 μL, 0.336 mmol), and DMF (400 μL). The mixture was vortexed and allowed to activate for 20 minutes at room temperature. The

activated mixture was then added to the resin. The reaction vessel was shaken at room temperature overnight. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). The resin was then washed with a solution of 25% TFA in DCM (20 mL), followed by shaking with 25% TFA in DCM at room temperature for 25 min to deprotect the Boc-protected alkyl amine, followed by shaking at room temperature for 25 min to deprotect the Boc-protected amine of pyrrole. In a separate vessel, a mixture of Boc-ImHz-OH (11) was activated using HBTU, DIEA, and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated Boc-ImHz-OBt dimer was then added to the vessel containing (Resin-BiPy-NH_2). The mixture was shaken at room temperature for 2.5 h. The reaction vessel was then drained and washed with DMF (15 mL) and DCM (15 mL). A solution of 50% TFA in DCM was then washed over the resin (20 mL), followed by shaking at room temperature for 25 minutes. In a separate vessel, a mixture of α-Fmoc-γ-Boc-(*R*)-2,4-diaminobutyric acid, ((*R*)-Fmoc-α-Boc-γ-Daba, and (**12**) was activated using HBTU, DIEA, and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated (*R*)-Fmoc-α-Boc-γ-Daba-OBt monomer was then added to the vessel containing (Resin-BiPy-HzIm-NH_2). The mixture was shaken for 2 h at room temperature. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). The resin was then washed with a solution of 25% TFA in DCM (20 mL), followed by shaking with 25% TFA in DCM at room temperature for 25 min to deprotect the Boc-protected alkyl amine. In a separate vessel, a mixture of Boc-PyBi-OH (**10**) was activated using HBTU, DIEA and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated Boc-PyBi-OBt dimer was then added to the vessel containing (Resin-BiPy-HzIm-γDaba-NH_2). The mixture was shaken at room temperature for 2.5 h. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). The resin was then washed with a solution of 25% TFA in DCM (20 mL), followed by shaking with 25% TFA in DCM at room temperature for 25 min to deprotect the Boc-protected alkyl amine, followed by shaking at room temperature with 25% TFA in DCM for 25 min to deprotect the Boc-protected amine of pyrrole. In a separate vessel, a mixture of NoHzOMe-OH (**13**) was activated using HBTU, DIEA and DMF as described above for (**10**). The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated NoHzOMe-OBt dimer was then added to the vessel containing (Resin-BiPy-HzIm-γDaba-BiPyNH_2). The mixture was shaken at room temperature for

2.5 h. The reaction vessel was then drained and washed with DMF (15 mL) and DCM (15 mL) to provide (Resin-BiPy-HzIm-γDaba-BiPy-HzNo).

C.7.6 O-Methyl Deprotection

To the synthesis vessel containing (Resin-BiPy-HzIm-γDaba-BiPy-HzNo) was added DCM (800 μL) and BCl_3 (400 μL), (2M in heptanes). The mixture was shaken at room temperature for 2.5 h. The vessel was then drained and washed with DCM (15 mL). Deprotection was determined to be complete by analytical HPLC

C.7.7 Cleavage From Resin

Following O-Methyl deprotection, DCM (200 μL) and $MeNH_2$ (1 mL) (2M in THF) was added to the synthesis vessel. The mixture was then shaken at 35°C for 12 h. The filtrate was collected from the synthesis vessel and the organics removed in vacuo. The remaining residue was dissolved in 0.1% TFA and purified using preparatory reverse phase HPLC. Lyophilization of the appropriate fractions provided oligomer **9** as a fine powder solid. MALDI-TOF-MS: calcd for $C_{55}H_{49}N_{19}O_8$: 1103.40; found 1103.41 $[M+H]^+$.