

## Alternative Heterocycles for DNA Recognition: The Benzimidazole/Imidazole Pair

Christoph A. Briehn, Philipp Weyermann, and Peter B. Dervan\*<sup>[a]</sup>

**Abstract:** Boc-protected benzimidazole-pyrrole, benzimidazole-imidazole, and benzimidazole-methoxypyrrole amino acids were synthesized and incorporated into DNA binding polyamides, comprised of *N*-methyl pyrrole and *N*-methyl imidazole amino acids, by means of solid-phase synthesis on an oxime resin. These hairpin polyamides were designed to determine the DNA recognition profile of a side-by-side benzimidazole/imidazole pair for the designated six base pair recognition

sequence. Equilibrium association constants of the polyamide-DNA complexes were determined at two of the six base pair positions of the recognition sequence by quantitative DNase I footprinting titrations on DNA fragments each containing matched and single base pair mismatched binding sites. The re-

**Keywords:** base pairing • binding affinity • DNA recognition • hydrogen bonds • polyamides

sults indicate that the benzimidazole-heterocycle building blocks can replace pyrrole-pyrrole, pyrrole-imidazole, and pyrrole-hydroxypyrrole constructs while retaining relative site specificities and subnanomolar match site affinities. The benzimidazole-containing hairpin polyamides represent a novel class of DNA binding ligands featuring tunable target recognition sequences combined with the favorable properties of the benzimidazole type DNA minor groove binders.

### Introduction

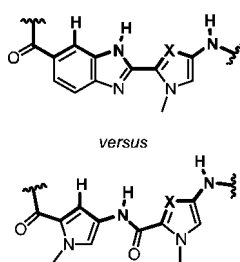
The deciphering of the human genome has spurred efforts to develop various strategies for gene regulation at the transcriptional and translational level. As promising modulators of gene transcription, DNA binding polyamides have proved to be useful tools that recognize predetermined sequences within the minor groove of DNA at subnanomolar concentrations.<sup>[1, 2]</sup> DNA recognition by polyamides depends on a code of side-by-side amino acid pairings that are oriented N–C with respect to the 5′–3′-direction of the DNA helix in the minor groove.<sup>[1]</sup> An antiparallel pairing of *N*-methyl imidazole opposite *N*-methyl pyrrole (Im/Py pair) distinguishes G•C from C•G and both of these from A•T/T•A base pairs (bp). An *N*-methyl pyrrole/*N*-methyl pyrrole (Py/Py) pair binds both A•T and T•A in preference to G•C/C•G. The discrimination of T•A from A•T by means of *N*-methyl hydroxypyrrole/*N*-methyl pyrrole (Hp/Py) pairs completes the four base pair code.<sup>[1]</sup>

In addition to the development of various polyamide binding motifs such as hairpins,<sup>[3]</sup> cycles,<sup>[4]</sup> H-pins,<sup>[5]</sup> and dimers,<sup>[6]</sup> efforts have been devoted to extend the ensemble of five-membered aromatic amino acids that are capable of cooperatively pairing with each other to recognize DNA base pair sequences in the minor groove.<sup>[7]</sup> Hitherto, less attention was focused on the incorporation of six-membered ring heterocycles and fused heteroaromatic rings into polyamide sequences and on the manipulation of the heterocycle tethering amide bonds. In earlier studies of polyamide-type minor groove binders pyridine- and phenol-based monomers (when paired with Py) were employed to target G•C and A•T base pairs, respectively.<sup>[8, 9]</sup> Fused heterocycles such as benzimidazoles, imidazopyridines, and indoles as part of small molecule ligands with extended  $\pi$ -conjugated systems are known to bind to the minor groove of DNA with high affinities<sup>[10]</sup> and some of them were shown to have biological activity.<sup>[11]</sup> Hoechst 33258, which comprises of a bis-benzimidazole, a *N*-methylpiperazine, and a phenol moiety, is one of the most prominent examples of these fused heterocycle derivatives due to its favorable fluorescence characteristics.<sup>[12]</sup> The crescent-shaped dye binds to the minor groove of DNA at A/T tracks as a 1:1 complex.<sup>[13]</sup>

As shown in Scheme 1, the benzimidazole moiety has the same atomic connectivity and hydrogen bonding surface along the recognition site (shown as bold bonds) as the 1-methylpyrrole-aminoacetyl moiety. In addition to van der Waals interactions of the aromatic ligand and the minor

[a] Prof. Dr. P. B. Dervan, Dr. C. A. Briehn, Dr. P. Weyermann  
Division of Chemistry and Chemical Engineering  
California Institute of Technology  
Pasadena, CA 91125 (USA)  
Fax: (+1) 626 683 8753  
E-mail: dervan@caltech.edu

Supporting information [ab initio calculations for four-ring polyamide subunits PyPyPyIm and PyBiPyIm (Figure S1)] for this article is available on the WWW under <http://www.chemeurj.org/> or from the author.



X = CH, N, COH

Scheme 1. Structures of benzimidazole-heterocycle construct and pyrrole-heterocycle dimer as they are incorporated into hairpin polyamides. Hydrogen bonding surfaces that are essential for minor-groove binding are indicated in bold type.

minor groove.<sup>[13]</sup> In contrast, when incorporating the benzimidazole ring into a side-by-side 2:1 ligand/DNA-recognition-site binding mode (e.g. in a hairpin polyamide) the benzimidazole might be expected to recognize A·T or T·A base pairs when paired with Py or Hp and a C·G base pair when paired with Im (Figure 1). Despite its formal similarity with the binding surface of aminoacetyl-pyrrole, the different electronic structure, larger ring size, restricted flexibility, and associated difference in curvature of the benzimidazole may interfere with a cooperative side-by-side pairing and result in an altered recognition behavior. In view of biological studies with polyamide-based minor groove binders, one could imagine that the replacement of amide bonds by fused

groove contacts, the proton at the N(1) position of the benzimidazole can form hydrogen bonds with the lone pairs of the N(3) of purines and the O(2) of pyrimidines like the protons of the amide bonds that link the aromatic rings in a polyamide. In Hoechst 33 258, which prefers a 1:1 ligand/DNA-recognition-site binding mode, the N(1) protons of each benzimidazole interact through bifurcated hydrogen bonds with adenine N(3)- and thymine O(2)-atoms on the floor of the

heterocycles may influence the cell permeation and distribution behavior as well as degradation pathways.

Here we report the synthesis of benzimidazole-containing Boc-protected amino acids, their incorporation into hairpin polyamides by applying the solid-phase synthetic protocols published previously,<sup>[14]</sup> and the evaluation of the DNA recognition properties of the side-by-side benzimidazole/imidazole (Bi/Im) pair within the hairpin polyamide motif. A series of benzimidazole-containing hairpin polyamides with the general sequence ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>γ-ImXBiPy-CONHMe (X = Py: **1b**; X = Hp: **3b**), ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>γ-PyImBiPy-CONHMe (**2b**) and the corresponding parent compounds ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>γ-ImXPyPy-CONHMe (X = Py: **1a**; X = Hp: **3a**) and ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>γ-PyImPyPy-CONHMe (**2a**) were synthesized (Figure 2). A schematic drawing of the benzimidazole monomer-containing hairpin polyamides bound to the matched DNA sequences is illustrated in Figure 1. Four plasmids were designed to probe the recognition behavior of these polyamides at the six base pair recognition sites 5'-TGRACA-3' and 5'-TGGXCA-3' (for **1a,b** and **3a,b**; R, X = A, T, G, C) as well as 5'-TGRCAA-3' and 5'-TGGXAA-3' (for **2a,b**), which differ at a single internal binding position R and X, respectively, allowing the systematical analysis of the recognition profile of the benzimidazole/imidazole pair in a hairpin polyamide at the two individual base pair positions R·S and X·Y (Figure 3). The key building block replacements for dimers PyPy and PyIm, benzimidazole-pyrrole (BiPy) and benzimidazole-imidazole (BiIm), respectively, have recently been reported by Singh<sup>[15]</sup> and Matsuba.<sup>[11]</sup> To our knowledge, benzimidazole moieties incorporated into a side-by-side binding motif such as shown in Figure 1 have not been examined in the context of the pairing rules.

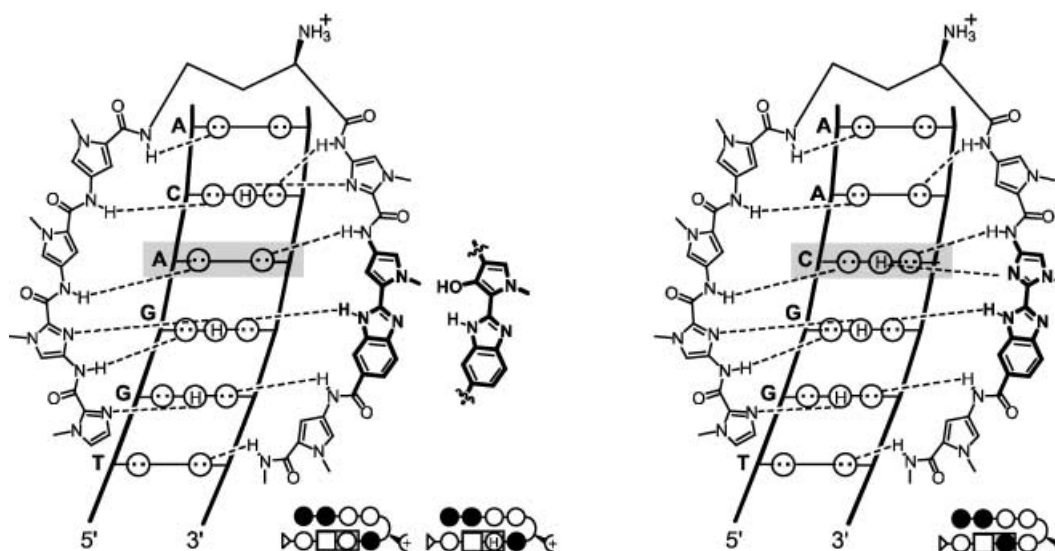


Figure 1. Binding models of the 1:1 polyamide-DNA complexes formed between the benzimidazole-containing polyamides ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>γ-ImXBiPyCONHMe (X = Py: **1b**; X = Hp: **3b**) and ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>γ-PyImBiPyCONHMe (**2b**) and their respective six base pair match sites 5'-TGGACA-3' and 5'-TGGCAA-3', respectively. Circles with two dots represent the lone pairs of N(3) of purines and O(2) of pyrimidines. Circles containing an "H" represent the N(2) hydrogen of guanine. Putative hydrogen bonds are illustrated by dotted lines. The incorporated benzimidazole-pyrrole, benzimidazole-hydroxypyrrole, and benzimidazole-imidazole moieties are shown in bold. Ball-and-stick models of polyamides are shown with shaded and nonshaded circles indicating imidazole and pyrrole carboxamides, respectively. Two touching squares represent benzimidazole-pyrrole, benzimidazole-imidazole, and benzimidazole-hydroxypyrrole constructs: the nonshaded square represents the benzimidazole moiety, the adjacent pyrrole, imidazole rings are depicted as nonshaded and shaded circles inside the second square as described above. Hydroxypyrrole rings are annotated with an "H" in the center of an unshaded circle. (R)-Diaminobutyric acid (DABA) is depicted as a curved line and a plus sign, the methyl amide tail is annotated as a triangle.

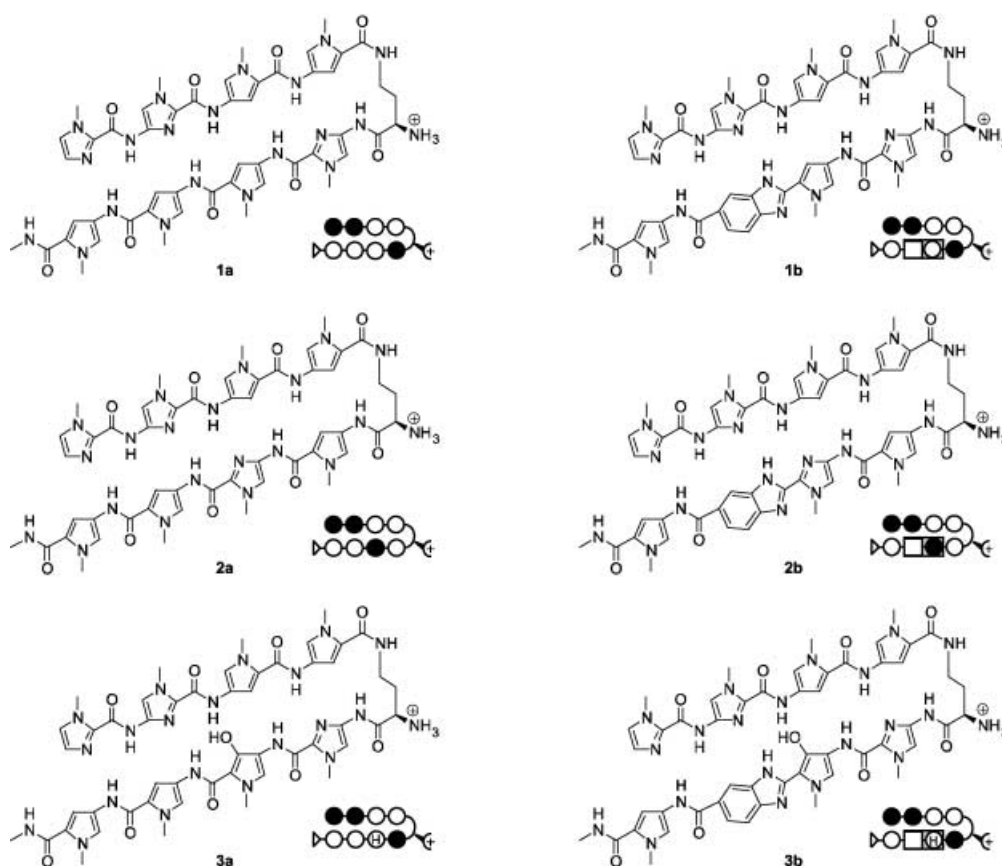


Figure 2. Structures and ball-and-stick models of the parent eight-ring polyamides **1a**, **2a**, and **3a**, and the corresponding benzimidazole-containing polyamides **1b**, **2b**, and **3b**.

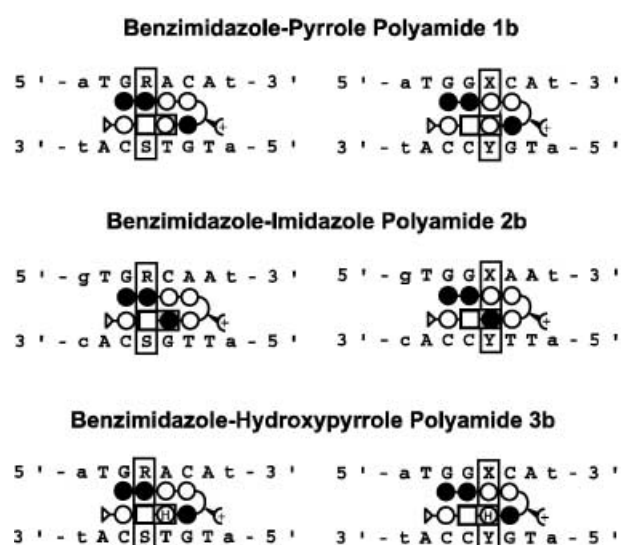


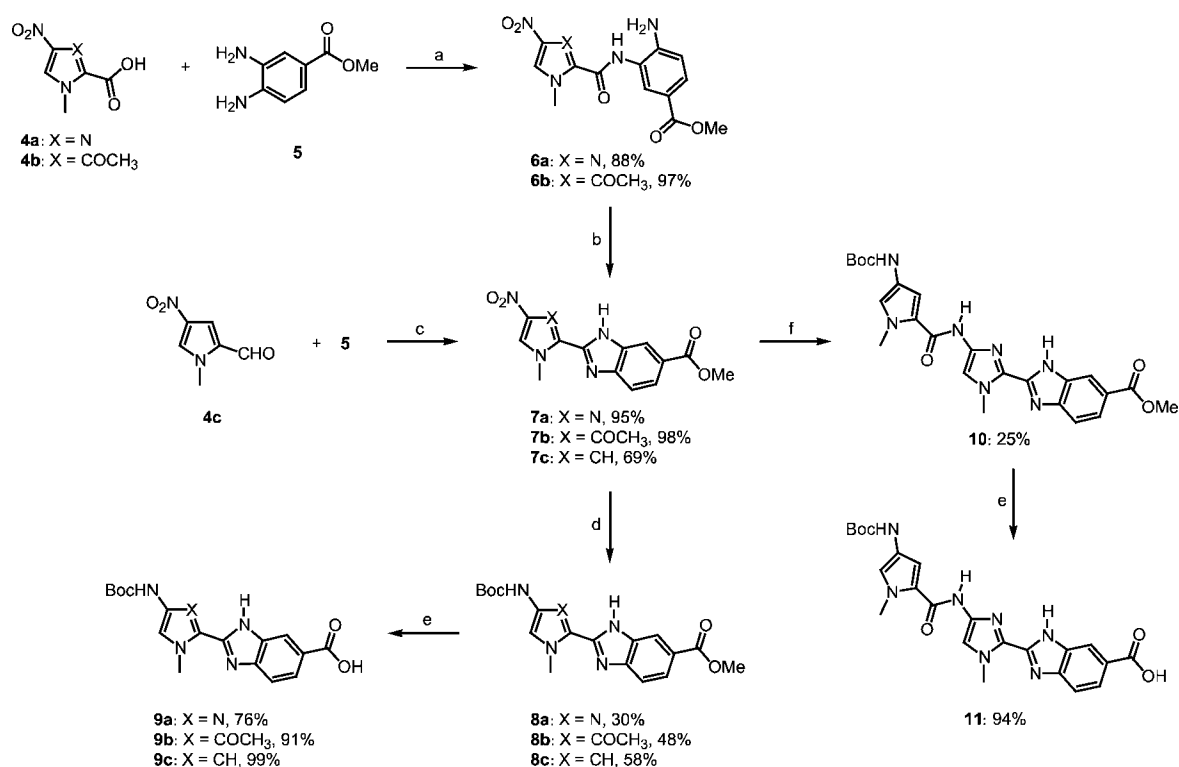
Figure 3. Schematic ball-and-stick models of hairpin polyamides **1b**, **2b**, and **3b** bound to recognition sequences 5'-TGRACA-3', 5'-TGRCAA-3', 5'-TGGXCA-3', and 5'-TGGXAA-3' (R, X = A, T, G or C). Ring pairing-DNA interactions investigated by quantitative DNase I footprint titrations are boxed.

## Results and Discussion

**Synthesis of building blocks:** Preparation of benzimidazole-containing polyamides **1b**, **2b**, and **3b** required the synthesis of Boc-protected benzimidazole-imidazole (Boc-ImBi-OH,

**9a**), benzimidazole-methoxypyrrole (Boc-OpBi-OH, **9b**), benzimidazole-pyrrole (Boc-BiPy-OH, **9c**), and benzimidazole-imidazole-pyrrole (Boc-PyImBi-OH, **11**) amino acids. As shown in Scheme 2, the benzimidazole scaffold was obtained by cyclodehydration of amides **6a,b**, which were in turn prepared from precursors **4a,b** by means of an HBTU-mediated coupling, in acetic acid<sup>[16]</sup> or by condensation of pyrrole-aldehyde **4c**<sup>[17]</sup> with *ortho*-diamine **5** and in situ Fe<sup>III</sup>/Fe<sup>II</sup>-catalyzed oxidative cyclodehydrogenation of the Schiff's base intermediate.<sup>[15]</sup> Subsequent reduction of the nitro group in **7a–c** and Boc protection of the resulting aminoesters furnished the protected aromatic amino acid esters **8a–c** in moderate to good yield. The desired carboxylic acids **9a–c** were finally obtained by hydrolysis of the esters **8a–c** with sodium hydroxide in dioxane. As will be discussed in the following paragraph, it was found advantageous for the solid-phase synthesis to use Boc protected amino acid **11** instead of **9a** as building block to incorporate the BiIm moiety into the hairpin polyamide **2b**. For the synthesis of trimer **11** nitro compound **7a** was reduced and subsequently reacted with Boc-Py-OBt to give ester **10** in moderate yield. Saponification of **10** furnished the protected amino acid **11** in 94% yield.

**Solid-phase synthesis:** Parent polyamides **1a**, **2a**, and the O-protected derivative **12a** were synthesized by solid-phase methods on Kaiser's oxime resin (0.48 mmol g<sup>-1</sup>) in a stepwise manner from Boc-*N*-methyl pyrrole (Boc-Py-OBt, **13a**), Boc-*N*-methyl imidazole (Boc-Im-OH, **14**), Boc-*N*-methyl me-



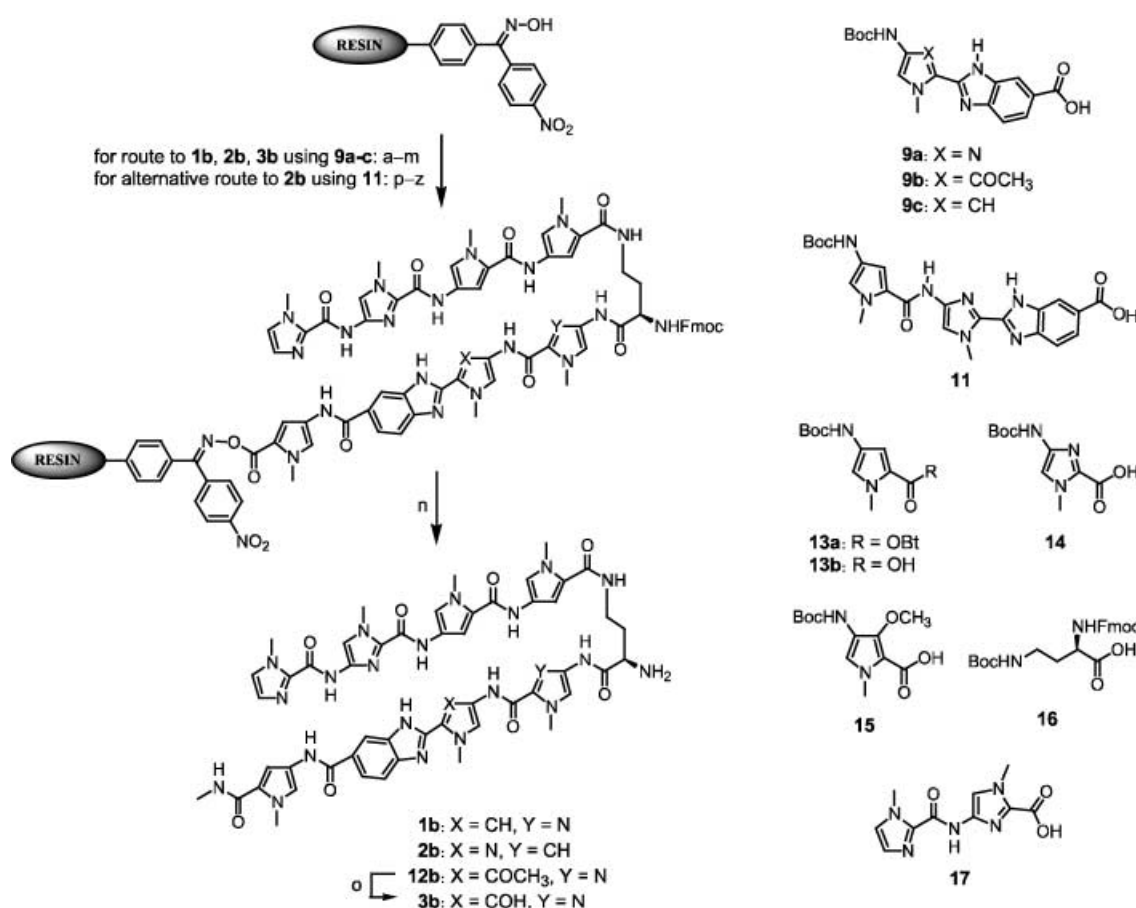
Scheme 2. Synthesis of Boc-protected benzimidazole amino acids **9a–c** and **11**. a) HBTU, DIEA, DMF; b) AcOH; c) FeCl<sub>3</sub>·6H<sub>2</sub>O, DMF; d) for **7a,c**: 1) H<sub>2</sub>, Pd/C, DMF; 2) Boc<sub>2</sub>O, DIEA; for **7b**: 1) SnCl<sub>2</sub>·2H<sub>2</sub>O, DMF; 2) Boc<sub>2</sub>O, DIEA; e) NaOH, dioxane; f) 1) H<sub>2</sub>, Pd/C, DMF; 2) Boc-Py-OBt, DIEA.

thoxyppyrrrole (Boc-Op-OH, **15**) monomers,  $\alpha$ -Fmoc- $\gamma$ -Boc-(*R*)-2,4-diaminobutyric acid ( $\alpha$ -Fmoc- $\gamma$ -Boc-(*R*)-DABA, **16**), and *N*-methyl imidazole dimer (ImIm-OH, **17**) in 16 steps.<sup>[14]</sup> Similarly, benzimidazoles **9a–c** could be successfully used as building blocks for the stepwise polyamide growth furnishing resin-bound benzimidazole-containing polyamides **1b**, **2b**, and the O-protected derivative **12b** in 13 steps (Scheme 3). It is noteworthy that the same reaction conditions proved to be suitable for peptide coupling and amine deprotection of benzimidazole monomers **9a–c** and the standard building blocks **13–17**: Benzimidazoles **9a–c** were coupled in DIEA/NMP after activation with HBTU and (when resin-bound) deprotected with 20% TFA/CH<sub>2</sub>Cl<sub>2</sub> (**9c**) or 50% TFA/CH<sub>2</sub>Cl<sub>2</sub> (**9a, b**). However, due to the moderate efficiency of the Boc-Py-OBt to imidazole amine coupling,<sup>[18]</sup> it was found advantageous in terms of the overall polyamide recovery to employ building block **11** instead of **9a** for the synthesis of polyamide **2b**. To liberate the resin-bound compounds and concomitantly remove the Fmoc group of the DABA turn, a sample of each polyamide-loaded polymer support was treated with methylamine in CH<sub>2</sub>Cl<sub>2</sub>/THF (14 h, 37 °C). The crude polyamides were isolated and purified by preparatory reversed-phase HPLC to yield ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -ImPyPyPy-CONHMe (**1a**), ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -ImPyBiPy-CONHMe (**1b**), ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -PyImPyPy-CONHMe (**2a**), ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -PyImBiPy-CONHMe (**2b**), ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -ImOpPyPy-CONHMe (**12a**), ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -ImOpBiPy-CONHMe (**12b**). *O*-Methyl protected polyamides **12a, b** were *O*-demethylated with sodium thiophenoxide/DMF (85 °C, 2 h)<sup>[19]</sup> and purified by preparatory reversed-phase HPLC to yield ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -ImHpPyPy-CONHMe

(**3a**) and ImImPyPy-(*R*)-H<sub>2</sub>N $\gamma$ -ImHpBiPy-CONHMe (**3b**). The synthetic protocol for the solid-phase synthesis of benzimidazole-containing polyamides **1b**, **2b**, and **3b** is illustrated in Scheme 3.

**DNA binding affinities and sequence specificities:** When replacing PyPy, PyIm, and PyHp dimers with BiPy (as in **1b**), BiIm (as in **2b**), and BiHp constructs (as in **3b**), respectively, the question arises whether a) the benzimidazole ring is capable of cooperating with the imidazole ring on the opposite strand of the hairpin polyamide to specifically recognize a C·G base pair according to the binding model and the general pairing rules discussed above, and b) whether the steric and electronic peculiarities of the benzimidazole ring disturb the binding ability of the adjacent heterocycle. In order to address this question for the given polyamide design, the recognition behavior of the benzimidazole-containing polyamides was probed in two of the six base pair positions R·S and X·Y of the recognition sequence (5'-TGRACA-3' and 5'-TGGXCA-3' for **1b** and **3b**; 5'-TGRCAA-3' and 5'-TGGXAA-3' for **2b**; Figure 3) and compared with the DNA-binding ability of the parent compounds **1a**, **2a**, and **3a** at the same recognition sequences.

Evaluation of the DNA recognition properties of the new ligands **1b**, **2b**, and **3b** at the imidazole/benzimidazole position (referred to as R·S position) should reveal the sequence preference of the benzimidazole ring. The binding properties at the neighboring ring position (referred to as X·Y position) were expected to disclose whether the recognition capabilities of the adjacent internal five-membered ring are disturbed by the benzimidazole moiety. To examine the



Scheme 3. Solid-phase synthesis of hairpin polyamides **1b**, **2b**, and **3b** and amino acid building blocks for the polyamide synthesis. a) **13a**, DIEA, NMP; b) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; c) for X = N: **9a**, HBTU, DIEA, NMP; for X = COCH<sub>3</sub>: **9b**, HBTU, DIEA, NMP; for X = CH: **9c**, HBTU, DIEA, NMP; d) for X = CH: 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; for X = N, COCH<sub>3</sub>: 50% TFA/CH<sub>2</sub>Cl<sub>2</sub>; e) for X = COCH<sub>3</sub>, CH: **14**, HBTU, DIEA, NMP; for X = N: **13b**, DCC, DMAP, CH<sub>2</sub>Cl<sub>2</sub>; f) for Y = CH: 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; for Y = N: 50% TFA/CH<sub>2</sub>Cl<sub>2</sub>; g) **16**, HBTU, DIEA, NMP; h) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; i) **13a**, DIEA, NMP; j) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; k) **13a**, DIEA, NMP; l) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; m) **17**, DCC, HOBT, DIEA, NMP; n) 2.0 M CH<sub>3</sub>NH<sub>2</sub>/THF, CH<sub>2</sub>Cl<sub>2</sub>; o) PhSnA, DMF; p) **13a**, DIEA, NMP; q) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; r) **11**, HBTU, DIEA, NMP; s) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; t) **16**, HBTU, DIEA, NMP; u) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; v) **13a**, DIEA, NMP; w) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; x) **13a**, DIEA, NMP; y) 20% TFA/CH<sub>2</sub>Cl<sub>2</sub>; z) **17**, DCC, HOBT, DIEA, NMP.

consequences of the ring replacement for DNA binding in terms of affinity and specificity, the benzimidazole-containing hairpin polyamides **1b**, **2b**, and **3b** were evaluated by quantitative DNase I footprint titrations and the results compared to those from parent compounds **1a**, **2a**, and **3a**, which were investigated under identical conditions. Quantitative DNase I footprint titration experiments (10 mM Tris-HCl, 10 mM KCl, 10 mM MgCl<sub>2</sub>, 5 mM CaCl<sub>2</sub>, pH 7.0, 22 °C, equilibration time: 12 h)<sup>[20]</sup> were performed on <sup>32</sup>P-3'-labeled *EcoRI/PvuII* restriction fragments of plasmids pPWF2 and pCAB1 (to probe the R·S position, Figure 4); and pAU2<sup>[21]</sup> and pCAB2 (to probe the X·Y position, Figure 5). Equilibrium association constants ( $K_a$ ) for polyamides **1a**, **b**, **2a**, **b**, and **3a**, **b** at the binding sites of interest were determined as described previously.<sup>[20]</sup>

On first inspection, analysis of the footprint data at position R·S (Table 1) and at position X·Y (Table 2) extracted from the autoradiograms shown in Figures 6–8 reveals that the benzimidazole-containing polyamides **1b**, **2b**, **3b** and the corresponding parent compounds **1a**, **2a**, **3a** exhibit the same binding site preferences for matched and mismatched sites on the 3'-labeled restriction fragments derived from pPWF2, pCAB1, pAU2, and pCAB2. The expected match sites out of

the sequence families 5'-TGRACA-3' or 5'-TGRCAA-3' (R·S position) and 5'-TGGXCA-3' or 5'-TGGXAA-3' (X·Y position) are bound with comparable equilibrium association constants by the benzimidazole-containing compounds **1b**, **2b**, and **3b** and the parent hairpin polyamides **1a**, **2a**, and **3a**. All compounds bind their match sites with high affinities in the subnanomolar range ( $K_a$  values range from  $4.9 \times 10^{10} \text{ M}^{-1}$  for R·S position of **1b** and X·Y position of **1a** down to  $3.5 \times 10^9 \text{ M}^{-1}$  for X·Y position of **2a**).

A closer inspection of the binding affinities at the two individual base pair positions elucidates the subtle effects of the benzimidazole introduction in the molecular recognition:

**Molecular recognition at the R·S position:** While both benzimidazole-containing polyamide **1b** and parent compound **1a** display comparable match-site affinities ( $K_a = 4.9 \times 10^{10} \text{ M}^{-1}$  and  $K_a = 2.7 \times 10^{10} \text{ M}^{-1}$ , respectively), benzimidazole-containing polyamides **2b** and **3b** exhibit a six to seven-fold higher match site affinity ( $K_a = 3.1 \times 10^{10} \text{ M}^{-1}$  and  $K_a = 2.5 \times 10^{10} \text{ M}^{-1}$ , respectively) than the parent DNA binders **2a** and **3a** ( $K_a = 5.2 \times 10^9 \text{ M}^{-1}$  and  $K_a = 3.8 \times 10^9 \text{ M}^{-1}$ , respectively). A significant difference can be seen for the two sets of compounds when their mismatch-site affinities and the

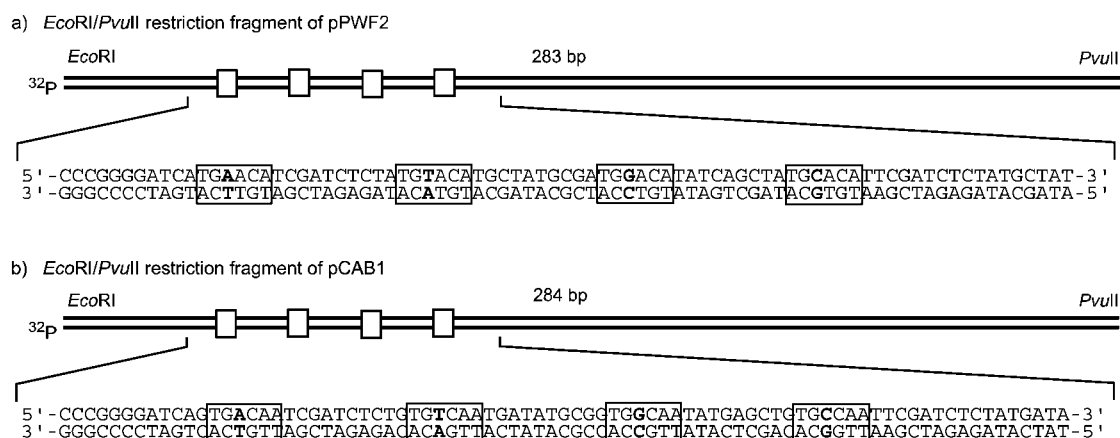


Figure 4. Illustration of the *EcoRI/PvuII* restriction fragments derived from plasmids a) pPWF2 and b) pCAB1. The four designed binding sites that were analyzed in quantitative footprint titrations are indicated with a box surrounding each of the six bp sites. The R·S base pair position within the binding sites is shown in bold type.

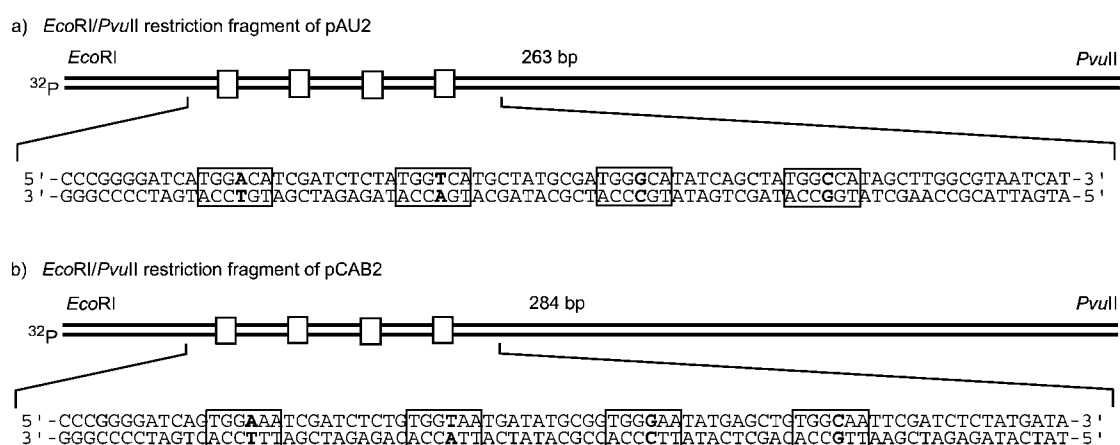


Figure 5. Illustration of the *EcoRI/PvuII* restriction fragments derived from plasmids a) pAU2 and b) pCAB2. The four designed binding sites that were analyzed in quantitative footprint titrations are indicated with a box surrounding each of the six bp sites. The X·Y base pair position within the binding sites is shown in bold type.

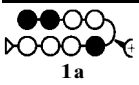
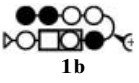
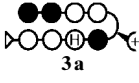
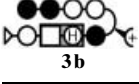
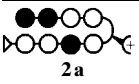
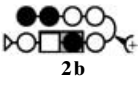
associated single site specificities are compared. The replacement of the PyPy, PyIm, and PyHp constructs by the benzimidazole-Py, -Im, and -Hp analogues is generally accompanied by an increase in the binding affinities for the single base pair mismatch sites of benzimidazole-containing polyamides **1b**, **2b**, and **3b** resulting in overall mildly diminished binding specificities. More precisely, for polyamide **1b** single base pair mismatch specificities are lower by roughly one order of magnitude when compared with the corresponding parent compounds **1a**. By contrast, compound **3b** displays a specificity profile that is comparable to that of the parent compound **3a**. For benzimidazole-containing compound **2b** the match site/mismatch site discrimination is comparable for R=A but reveals at least ten-fold lower specificities for R=T or C.

**Molecular recognition at the X·Y position:** As expected, for the X·Y position of the sequence family 5'-TGGXCA-3' (compounds **1a, b** and **3a, b**) and 5'-TGGXAA-3' (compounds **2a, b**) the response to the heterocycle replacement in terms of match site/single base pair mismatch site discrimination is less pronounced. Binding affinities for match and

mismatch sites and the associated binding specificities reveal the same overall trend and range in the same order of magnitude for both the benzimidazole-containing polyamides **1b**, **3b** and the parent compounds **1a**, **3a**. By contrast, benzimidazole-containing polyamide **2b** exhibits a six-fold higher match site affinity than its parent binder **2a** ( $K_a = 1.9 \times 10^{10} \text{ M}^{-1}$  vs  $K_a = 3.5 \times 10^9 \text{ M}^{-1}$ , respectively). Along with the significantly higher match site affinity of compound **2b** slightly increased binding affinities for single base pair mismatch sites are displayed, which result in mildly diminished specificities when compared with the values of the parent compounds **2a**. Obviously, the introduction of the benzimidazole moiety has only a marginal influence on the binding properties of the adjacent five-membered internal ring.

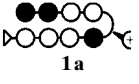
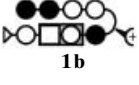
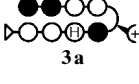
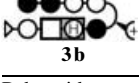
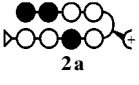
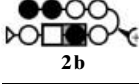
While the relative contributions of various ligand–DNA interactions to the overall recognition profile of a DNA binder are still subject of discussion,<sup>[22, 23]</sup> hydrogen bonding, van der Waals contacts and electrostatic interactions proved to be the major determinants for DNA recognition. Factors that have to be taken into account to rationalize the slightly different DNA recognition properties of the benzimidazole-

Table 1. Equilibrium association constants  $K_a$  [ $M^{-1}$ ] and specificities for polyamides **1a**, **1b**, **2a**, **2b**, **3a**, and **3b** on plasmids pPWF2 and pCAB1 (R·S position).<sup>[a-d]</sup>

Polyamide on pPWF2	5'-aTGAACAt-3'	5'-aTGTACAt-3'	5'-aTGGACAt-3'	5'-aTGCACAt-3'
 <b>1a</b>	$4.4 (\pm 1.2) \times 10^8$ [61]	$\leq 1 \times 10^8$ [ $\geq 270$ ]	<b><math>2.7 (\pm 0.5) \times 10^{10}</math></b>	$\leq 1 \times 10^8$ [ $\geq 270$ ]
 <b>1b</b>	$3.9 (\pm 0.2) \times 10^9$ [13]	$3.3 (\pm 0.3) \times 10^9$ [15]	<b><math>4.9 (\pm 0.1) \times 10^{10}</math></b>	$9.0 (\pm 0.5) \times 10^8$ [54]
 <b>3a</b>	$\leq 3 \times 10^7$ [ $\geq 127$ ]	$\leq 1 \times 10^7$ [ $\geq 380$ ]	<b><math>3.8 (\pm 0.2) \times 10^9</math></b>	$\leq 1 \times 10^7$ [ $\geq 380$ ]
 <b>3b</b>	$3.3 (\pm 0.7) \times 10^8$ [76]	$2.4 (\pm 0.4) \times 10^8$ [104]	<b><math>2.5 (\pm 0.4) \times 10^{10}</math></b>	$1.0 (\pm 0.3) \times 10^8$ [250]
Polyamide on pCAB1	5'-gTGACAAt-3'	5'-gTGTCAAAt-3'	5'-gTGGCAAAt-3'	5'-gTGCCAAAt-3'
 <b>2a</b>	$\leq 1 \times 10^8$ [ $\geq 52$ ]	$\leq 1 \times 10^7$ [ $\geq 520$ ]	<b><math>5.2 (\pm 0.2) \times 10^9</math></b>	$\leq 1 \times 10^7$ [ $\geq 520$ ]
 <b>2b</b>	$7.9 (\pm 1.4) \times 10^8$ [39]	$6.2 (\pm 1.4) \times 10^8$ [50]	<b><math>3.1 (\pm 0.5) \times 10^{10}</math></b>	$4.0 (\pm 0.7) \times 10^8$ [78]

[a] The reported association constants  $K_a$  are the average values obtained from three DNase I footprint titration experiments, with the standard deviation for each data set 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<sub>2</sub>, and 5 mM CaCl<sub>2</sub> with an equilibration time of 12 h. [c] Match site equilibration association constants are shown in bold. [d] Specificities are given in square brackets under the  $K_a$  values and calculated as  $K_a(\text{match})/K_a(\text{mismatch})$ .

Table 2. Equilibrium association constants  $K_a$  [ $M^{-1}$ ] and specificities for polyamides **1a**, **1b**, **2a**, **2b**, **3a**, and **3b** on plasmids pAU2 and pCAB2 (X·Y position).<sup>[a-d]</sup>

Polyamide on pAU2	5'-aTGGACAt-3'	5'-aTGGTCAAt-3'	5'-aTGGGCAt-3'	5'-aTGGCCAt-3'
 <b>1a</b>	<b><math>4.9 (\pm 1.8) \times 10^{10}</math></b>	<b><math>4.9 (\pm 0.2) \times 10^{10}</math></b>	$2.0 (\pm 0.5) \times 10^9$ [25]	$4.5 (\pm 0.8) \times 10^9$ [11]
 <b>1b</b>	<b><math>2.7 (\pm 1.0) \times 10^{10}</math></b>	<b><math>3.6 (\pm 0.9) \times 10^{10}</math></b>	$3.4 (\pm 0.6) \times 10^9$ [11] <sup>[e]</sup>	$5.7 (\pm 0.5) \times 10^9$ [6] <sup>[e]</sup>
 <b>3a</b>	<b><math>8.9 (\pm 1.3) \times 10^9</math></b>	$1.3 (\pm 0.1) \times 10^9$ [7]	$1.8 (\pm 0.4) \times 10^8$ [49]	$1.7 (\pm 0.6) \times 10^8$ [52]
 <b>3b</b>	<b><math>1.8 (\pm 0.7) \times 10^{10}</math></b>	$2.5 (\pm 0.3) \times 10^9$ [7]	$5.8 (\pm 1.8) \times 10^8$ [31]	$7.0 (\pm 0.6) \times 10^8$ [25]
Polyamide on pCAB2	5'-gTGGAAAAt-3'	5'-gTGGTAAAt-3'	5'-gTGGGAAAt-3'	5'-gTGGCAAAt-3'
 <b>2a</b>	$\leq 1 \times 10^7$ [ $\geq 350$ ]	$\leq 1 \times 10^8$ [ $\geq 35$ ]	$\leq 1 \times 10^7$ [ $\geq 350$ ]	<b><math>3.5 (\pm 0.5) \times 10^9</math></b>
 <b>2b</b>	$\geq 1 \times 10^8$ [ $\geq 190$ ]	$1.6 (\pm 0.1) \times 10^9$ [12]	$2.8 (\pm 0.3) \times 10^8$ [68]	<b><math>1.9 (\pm 0.3) \times 10^{10}</math></b>

[a] The reported association constants  $K_a$  are the average values obtained from three DNase I footprint titration experiments, with the standard deviation for each data set 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<sub>2</sub>, and 5 mM CaCl<sub>2</sub> with an equilibration time of 12 h. [c] Match site equilibration association constants are shown in bold type. [d] Specificities are given in square brackets under the  $K_a$  values and calculated as  $K_a(\text{match})/K_a(\text{mismatch})$ . [e] Specificity calculated as  $K_a(5'-aTGGTCAAt-3')/K_a(\text{mismatch})$ .

containing hairpin polyamides when compared to their parent compounds can be categorized into structural and electronic characteristics of the benzimidazole moiety: a) In comparison

with the aminoacetyl-pyrrole system the benzimidazole moiety has with its greater aromatic surface a higher hydrophobicity and polarizability that may alter both van der Waals DNA/ligand interactions and the intra-ligand  $\pi$ -stacking of the two opposite polyamide strands of the hairpin.<sup>[1a, 10b]</sup> b) Gas-phase calculations of PyBiPyIm and PyPyPyIm fragments indicate that the benzimidazole-containing strand PyBiPyIm has a smaller degree of curvature than the corresponding PyPyPyIm strand (see Figure S1, Supporting Information).<sup>[24]</sup> In general, as hairpin polyamides comprised of Py, Im, and Hp are slightly overcurved with respect to the DNA, the decreased curvature of the benzimidazole-containing strand of the hairpin polyamide should improve shape complementarity between the recognition face of the ligand and the floor of the minor groove of DNA and therefore increase binding affinity. However, for the given benzimidazole-containing hairpin polyamides it should be noticed that the larger curvature of the remaining unmodified strand might diminish the cooperative side-by-side ring pairing and as a consequence reduce single base pair mismatch specificity compared to the parent hairpin polyamides. c) The replacement of an amide linkage by a benzimidazole moiety reduces the degree of conformational freedom which is expected to result in an increased preorganisation of the ligand leading to higher binding affinity.

The results from the discussed quantitative DNase I footprint titrations demonstrate that the benzimidazole moiety is able to pair with an Im residue within the context of a hairpin polyamide and that, as a consequence, this pairing is able to distinguish G·C from C·G, A·T, and T·A. Therefore, the recognition behavior of the studied benzimidazole/imidazole

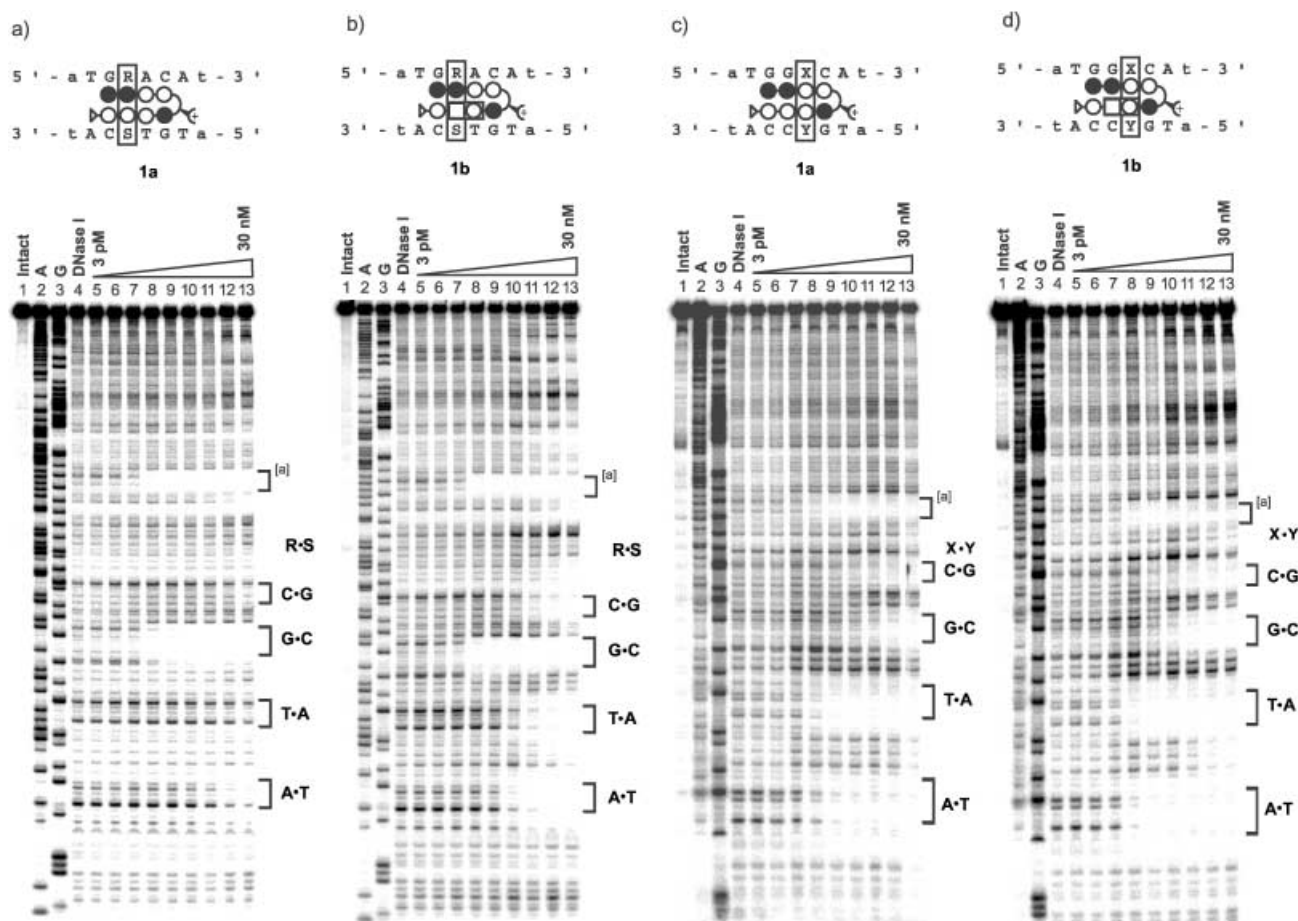


Figure 6. Quantitative DNase I footprint titration experiments on the 3'-<sup>32</sup>P-labeled 283-bp and 263-bp *EcoRI/PvuII* restriction fragments derived from plasmids pWF2 (a and b) and pAU2 (c and d), respectively. a) and c) polyamide **1a**: lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lane 4, DNase I standard; lanes 5–13: 3, 10, 30, 100, 300 pM and 1, 3, 10, 30 nM polyamide, respectively. b) and d) polyamide **1b**: lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lane 4, DNase I standard; lanes 5–13: 3, 10, 30, 100, 300 pM and 1, 3, 10, 30 nM polyamide, respectively. The analyzed 6-bp binding site locations are designated in brackets along the right side of each autoradiogram with their respective unique base pairs indicated. Schematic binding models of **1a** and **1b** with their putative binding sites are shown on the top side of the autoradiograms. Flanking sequences are designated in lower case while the binding site is given in capitals. The boxed R·S and X·Y base pairs indicate the positions that were examined in the experiments. [a] Additional match site for **1a,b**; sequence 5'-aTGGTCAT-3'.

(Bi/Im) pair is consistent with the well-established binding model discussed above. Presumably, a combination of electronic properties, steric bulk, rigidity and the associated alteration of the overall curvature of the benzimidazole moiety translate into a mildly diminished single base pair mismatch specificity in the R·S position, when compared with the parent compounds. However, the benzimidazole moiety minimally disturbs the recognition properties of the adjacent internal five-membered heterocycle resulting in single base pair mismatch specificities at the X·Y position that are similar to those of the parent polyamides. In addition, all benzimidazole-polyamides retain or even exceed the subnanomolar association constants of their corresponding parent compounds at their match sites.

## Conclusion

Benzimidazole moieties were incorporated into hairpin polyamides **1b**, **2b** and **3b** in the form of benzimidazole-imidazole

(**9a**), benzimidazole-methoxypyrrole (**9b**), benzimidazole-pyrrole (**9c**), and benzimidazole-imidazole-pyrrole (**11**) amino acids as a formal replacement for PyIm, PyHp, PyPy dimers and a PyImPy trimer, respectively. The new building blocks were synthesized from diamine **5** and pyrrole carbaldehyde **4c**, imidazole carboxylic acid **4a**, and hydroxypyrrole carboxylic acid **4b**, respectively. The evaluation of the DNA binding properties of the hairpin polyamides **1b**, **2b**, and **3b** by means of quantitative DNase I footprint titrations demonstrated that the benzimidazole ring in an Im/Bi side-by-side pair exhibits the same recognition preferences for a G·C base pair as a Py ring in an Im/Py side-by-side pair. However, the pyrrole-to-benzimidazole ring replacement results in a mildly diminished specificity at the replacement position R·S. The binding properties of the adjacent five-membered heterocycles at position X·Y are only marginally influenced by the steric demand and electronic properties of the benzimidazole ring. The subnanomolar DNA binding affinities for matched binding sites displayed by the parent polyamides are fully retained (for **1b**) or even exceeded (for **2b**, **3b**). The BiIm-



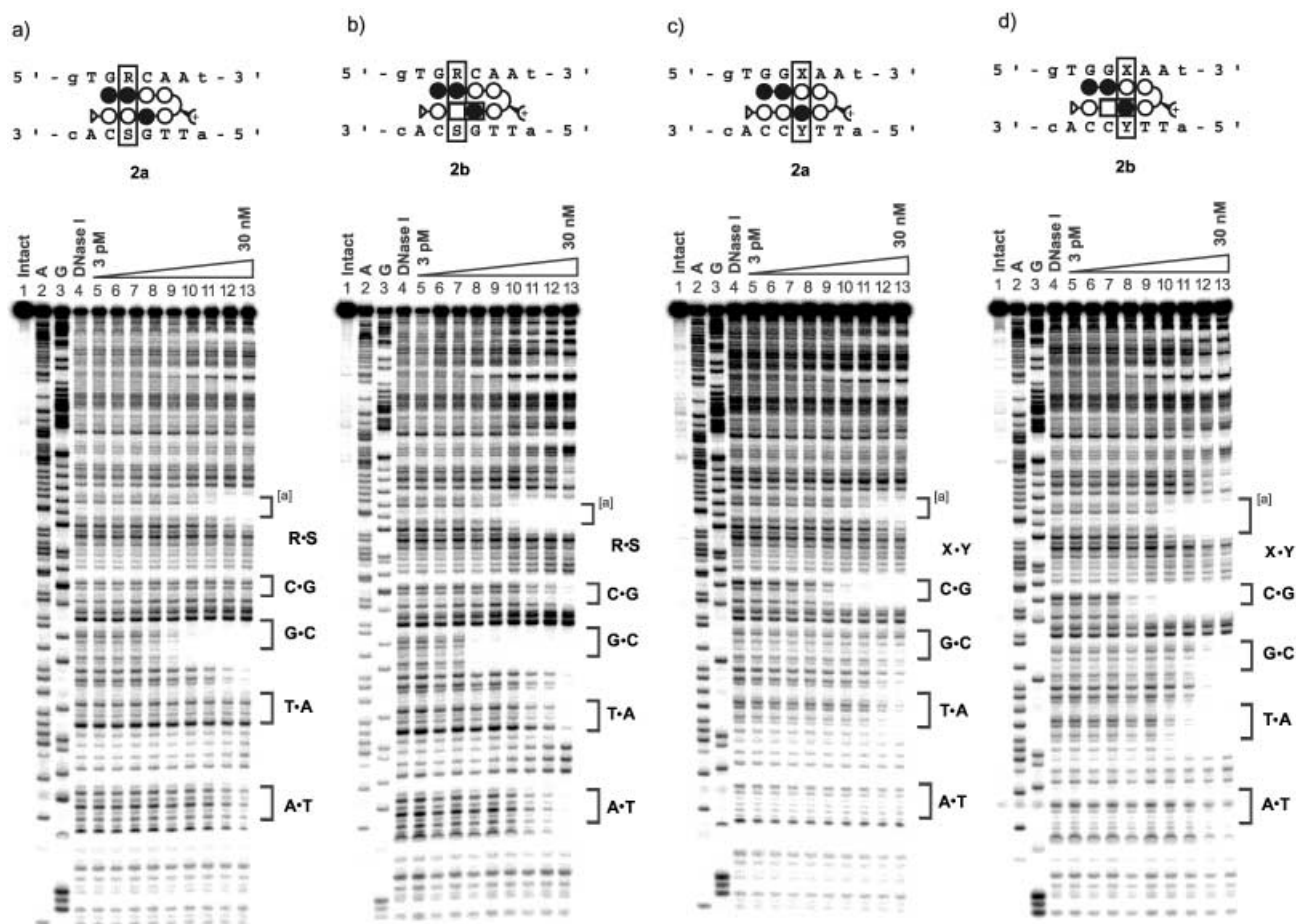


Figure 7. Quantitative DNase I footprint titration experiments on the 3'-<sup>32</sup>P-labeled 284-bp *EcoRI/PvuII* restriction fragments derived from plasmid pCAB1 (a and b) and pCAB2 (c and d). a) and c) polyamide **2a**: lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lane 4, DNase I standard; lanes 5–13: 3, 10, 30, 100, 300 pM and 1, 3, 10, 30 nM polyamide, respectively. b) and d) polyamide **2b**: lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lane 4, DNase I standard; lanes 5–13: 3, 10, 30, 100, 300 pM and 1, 3, 10, 30 nM polyamide, respectively. The analyzed 6-bp binding site locations are designated in brackets along the right side of each autoradiogram with their respective unique base pairs indicated. Schematic binding models of **2a** and **2b** with their putative binding sites are shown on the top side of the autoradiograms. Flanking sequences are designated in lower case while the binding site is given in capitals. The boxed R·S and X·Y base pairs indicate the positions that were examined in the experiments. [a] Three overlapping 1-bp mismatch sites for **2a,b**; sequence 5'-gAAGCTTGGCGTa-3'.

and BiHp-containing hairpin polyamides **2b** and **3b** exhibit a significant increase in binding affinity compared to the parent compounds. Even if the recognition profile of the benzimidazole ring was only studied in a side-by-side pair with Im, it is expected that in analogy Bi/Py and Bi/Hp side-by-side pairs will recognize A·T or T·A base pairs in accordance with the general pairing rules.<sup>[25]</sup> Future studies will address the influence of the benzimidazole moiety on cellular uptake and cellular distribution properties.<sup>[26]</sup>

## Experimental Section

**General:** *N,N'*-Dicyclohexylcarbodiimide (DCC), *N*-hydroxybenzotriazole (HOBt), and 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were purchased from Peptides International. Oxime resin was purchased from Novabiochem (0.48 mmol g<sup>-1</sup>, batch no. A18763). *N,N*-Diisopropylethylamine (DIEA) and *N,N*-dimethylformamide (DMF) were purchased from Applied Biosystems. (*R*)-2-Fmoc-4-Boc-diaminobutyric acid ( $\alpha$ -Fmoc- $\gamma$ -Boc-(*R*)-DABA, **16**) was from Bachem, methyl 3,4-diaminobenzoate (**5**) from Avocado, dichloromethane (DCM) was reagent grade from EM, and trifluoroacetic acid (TFA) was

from Halocarbon. All other chemicals were obtained reagent-grade from Aldrich (unless otherwise stated) and used without further purification. <sup>1</sup>H NMR and <sup>13</sup>C NMR spectra were recorded on a Varian Mercury 300 instrument. Chemical shifts are reported in ppm with reference to the solvent residual signal. UV spectra were measured on a Beckman Coulter DU 7400 diode array spectrophotometer. Matrix-assisted, laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS), electrospray injection (ESI-MS) and high resolution mass spectrometry (FAB, EI) was carried out at the California Institute of Technology. HPLC analysis was performed on a Beckman Gold system using a RAININ C<sub>18</sub>, Microsorb MV, 5  $\mu$ m, 300  $\times$  4.6 mm reversed-phase column in 0.1% (*w/v*) TFA with acetonitrile as eluent and a flow rate of 1.0 mL min<sup>-1</sup>, gradient elution 1.25% acetonitrile min<sup>-1</sup>. Preparatory HPLC was carried out on a Beckman HPLC using a Waters DeltaPak 100  $\times$  25 mm, 100  $\mu$ m C<sub>18</sub> column, 0.1% (*w/v*) TFA, 0.25% acetonitrile min<sup>-1</sup>. 18 M $\Omega$  water was obtained from a Millipore MilliQ water purification system, and all buffers were 0.2  $\mu$ m filtered. DNA oligonucleotides were synthesized by the Biopolymer Synthesis Center at the California Institute of Technology and used without further purification. Plasmids were sequenced by the Sequence/Structure Analysis Facility (SAF) at the California Institute of Technology. dNTPs (PCR nucleotide mix) and all enzymes (unless otherwise stated) were purchased from Roche Diagnostics and used with their supplied buffers. pUC19 was from New England Biolabs. [ $\alpha$ -<sup>32</sup>P]-Deoxyadenosine triphosphate and [ $\alpha$ -<sup>32</sup>P]-thymine triphosphate was from New England Nucleotides. RNase-free water (used for all DNA manipulations) was from

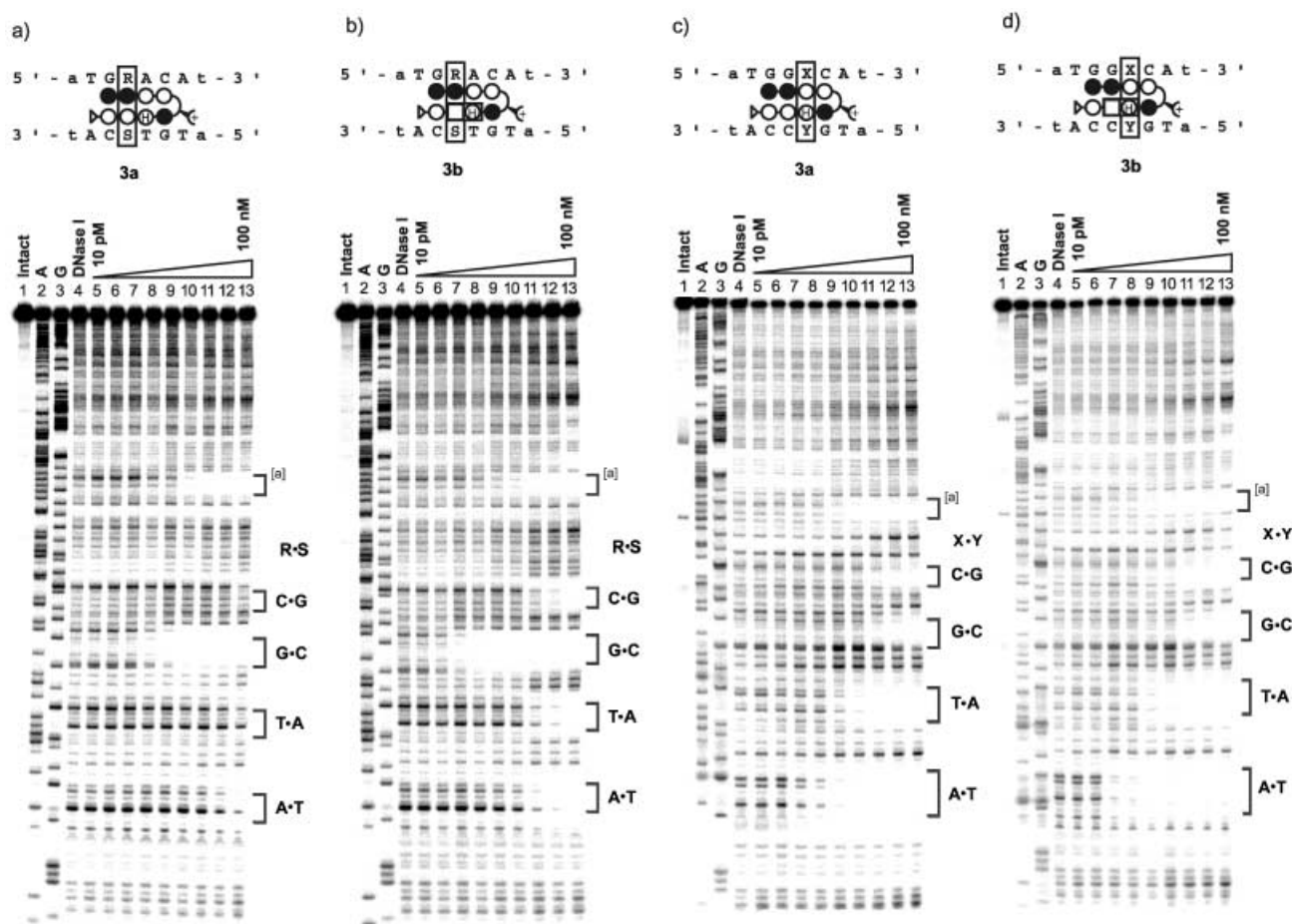


Figure 8. Quantitative DNase I footprint titration experiments on the 3'-<sup>32</sup>P-labeled 283-bp and 263-bp *EcoRI/PvuII* restriction fragments derived from plasmids pPWF2 (a and b) and pAU2 (c and d), respectively. a) and c) polyamide **3a**: lane 1, intact DNA; lane 2, A-specific reaction; lane 3, G-specific reaction; lane 4, DNase I standard; lanes 5–13: 10, 30, 100, 300 pM and 1, 3, 10, 30, 100 nM polyamide, respectively. b) and d) polyamide **3b**: lane 1, intact DNA; lane 2, A specific reaction; lane 3, G specific reaction; lane 4, DNase I standard; lanes 5–13: 10, 30, 100, 300 pM and 1, 3, 10, 30, 100 nM polyamide, respectively. The analyzed 6-bp binding site locations are designated in brackets along the right side of each autoradiogram with their respective unique base pairs indicated. Schematic binding models of **3a** and **3b** with their putative binding sites are shown on the top side of the autoradiograms. Flanking sequences are designated in lower case while the binding site is given in capitals. The boxed R·S and X·Y base pairs indicate the positions that were examined in the experiments. [a] Additional 1-bp mismatch site for **3a,b**; sequence 5'-aTGGTCAT-3'.

US Biochemicals. Ethanol (100%) was from Equistar, isopropanol from Mallinckrodt. Bromophenol blue and xylene cyanol FF were from Acros. DNA manipulations were performed according to standard protocols.<sup>[27]</sup> Autoradiography was performed with a Molecular Dynamics Typhoon Phosphorimager. 1-Methyl-4-nitro-1*H*-imidazole-2-carboxylic acid (**4a**),<sup>[14a]</sup> 1-methyl-4-nitro-1*H*-pyrrole-2-carbaldehyde (**4c**),<sup>[17]</sup> 3-methoxy-1-methyl-4-nitro-1*H*-pyrrole-2-carboxylic acid ethyl ester<sup>[28]</sup> were synthesized according to literature procedures.

#### Building block synthesis

**3-Methoxy-1-methyl-4-nitro-1*H*-pyrrole-2-carboxylic acid (4b)**: 3-Methoxy-1-methyl-4-nitro-1*H*-pyrrole-2-carboxylic ethyl ester (5.50 g, 24.1 mmol) was dissolved in ethanol (100 mL). NaOH (aqueous, 1M, 100 mL) was added and the solution stirred for 13 h at ambient temperature. The yellow solution was carefully acidified with aqueous HCl (1M) to pH 2–3. The formed white precipitate was filtered, washed with water and dried in vacuo to yield the title compound **4b** as a white powder (3.85 g, 80%). <sup>1</sup>H NMR (300 MHz, [D<sub>6</sub>]DMSO): δ = 3.80 (s, 3H), 3.82 (s, 3H), 8.10 (s, 1H); <sup>13</sup>C NMR (75 MHz, [D<sub>6</sub>]DMSO): δ = 38.4, 62.4, 114.6, 126.4, 127.0, 145.1, 160.5; MS (EI): *m/z*: 201 [M+H]<sup>+</sup>; HR-MS (EI): calcd for C<sub>7</sub>H<sub>8</sub>N<sub>2</sub>O<sub>5</sub>: 200.0433; found: 200.0435.

**4-Amino-3-[1-methyl-4-nitro-1*H*-imidazole-2-carbonyl]-amino]-benzoic methyl ester (6a)**: Carboxylic acid **4a** (2.57 g, 15.0 mmol) and methyl 3,4-diamino-benzoate (**5**) (2.49 g, 15.0 mmol) were dissolved in DMF (30 mL).

HBTU (5.72 g, 15.0 mmol) was added to the solution followed by DIEA (3 mL). After stirring for 20 h at ambient temperature the reaction mixture was poured into ice water. The precipitate was filtered, washed with water and dried in vacuo to yield the title compound **6a** as a yellow powder (4.22 g, 88%). <sup>1</sup>H NMR (300 MHz, [D<sub>6</sub>]DMSO): δ = 3.74 (s, 3H), 4.01 (s, 3H), 5.89 (s, 2H), 6.74 (d, *J* = 8.8 Hz, 1H), 7.59 (dd, *J*<sub>1</sub> = 8.8 Hz, *J*<sub>2</sub> = 1.6 Hz, 1H), 7.71 (d, *J* = 1.6 Hz, 1H), 8.64 (s, 1H), 10.04 (s, 1H); <sup>13</sup>C NMR (75 MHz, [D<sub>6</sub>]DMSO): δ = 36.6, 51.4, 114.5, 115.9, 120.3, 126.5, 128.7, 129.0, 137.7, 144.1, 148.6, 156.7, 165.8; MS (ESI): *m/z*: 320 [M+H]<sup>+</sup>; HR-MS (EI): calcd for C<sub>13</sub>H<sub>13</sub>N<sub>5</sub>O<sub>5</sub>: 319.0916; found: 319.0929.

**4-Amino-3-[(3-methoxy-1-methyl-4-nitro-1*H*-pyrrole-2-carbonyl)-amino]-benzoic methyl ester (6b)**: Carboxylic acid **4b** (2.17 g, 10.8 mmol) and methyl 3,4-diamino-benzoate (**5**) (1.81 g, 10.8 mmol) were dissolved in DMF (26 mL). HBTU (4.11 g, 10.8 mmol) was added to the solution followed by DIEA (2.6 mL). After stirring for 20 h at ambient temperature the reaction mixture was poured into ice water. The precipitate was filtered, washed with water and dried in vacuo to yield the title compound **6b** as a yellow powder (3.66 g, 97%). <sup>1</sup>H NMR (300 MHz, [D<sub>6</sub>]DMSO): δ = 3.75 (s, 3H), 3.86 (s, 3H), 3.95 (s, 3H), 5.82 (s, 2H), 6.78 (d, *J* = 8.4 Hz, 1H), 7.57 (dd, *J*<sub>1</sub> = 8.4 Hz, *J*<sub>2</sub> = 2.0 Hz, 1H), 7.93 (d, *J* = 2.0 Hz, 1H), 8.13 (s, 1H), 8.95 (s, 1H); <sup>13</sup>C NMR (75 MHz, [D<sub>6</sub>]DMSO): δ = 38.0, 51.4, 62.9, 114.7, 116.4, 116.5, 121.3, 125.9, 126.4, 127.3, 128.0, 141.4, 147.3, 157.8, 165.9;

MS (ESI):  $m/z$ : 349  $[M+H]^+$ ; HR-MS (EI): calcd for  $C_{15}H_{16}N_4O_6$ : 348.1069; found: 348.1072.

**2-(1-Methyl-4-nitro-1H-imidazol-2-yl)-1H-benzimidazole-5-carboxylic methyl ester (7a):** Amide **6a** (3.99 g, 12.5 mmol) was suspended in acetic acid (50 mL) and heated to 140 °C for 6 h. The solvent was removed in vacuo, diethyl ether was added, the yellow precipitate filtered and dried in vacuo affording the title compound **7a** as an off-white powder (3.57 g, 95 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 3.87 (s, 3H), 4.26 (s, 3H), 7.57–8.34 (m, 3H), 8.71 (s, 1H), 13.78 (s, 1H);  $^{13}C$  NMR (75 MHz,  $[D_6]DMSO$ ):  $\delta$  = 36.8, 52.1, 112.2, 113.8, 119.1, 120.9, 123.8, 124.5, 126.2, 136.3, 144.2, 145.6, 166.3; MS (ESI):  $m/z$ : 302  $[M+H]^+$ ; HR-MS (EI): calcd for  $C_{13}H_{11}N_5O_4$ : 301.0811; found: 301.0805.

**2-(3-Methoxy-1-methyl-4-nitro-1H-pyrrol-2-yl)-3H-benzimidazole-5-carboxylic methyl ester (7b):** Amide **6b** (3.56 g, 10.2 mmol) was suspended in diethyl ether (40 mL) and refluxed at 140 °C for 12 h. Upon cooling to room temperature a yellow solid precipitated that was filtered, suspended in diethyl ether, filtered, and dried in vacuo to yield the title compound **7b** as a yellow solid (3.33, 98 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 3.93 (s, 1.5H), 3.94 (s, 1.5H), 3.97 (s, 3H), 4.15 (s, 3H), 7.72–7.95 (m, 2H), 8.24 (s, 1H), 8.25–8.32 (m, 1H), 12.44 (s, 0.5H), 12.52 (s, 0.5H). At room temperature the presence of two benzimidazole tautomers (1H/3H) could be deduced from the  $^1H$  NMR spectra. The ratio between the two isomers was roughly estimated as 1:1 based on the integration of the protons at  $\delta$  = 3.93/3.94 and 12.44/12.52.  $^{13}C$  NMR (75 MHz,  $[D_6]DMSO$ ):  $\delta$  = 38.5, 52.0, 62.7, 111.8, 113.62, 113.75, 118.2, 120.0, 122.72, 123.18, 123.45, 123.65, 126.38, 126.53, 133.5, 137.4, 140.63, 140.87, 142.2, 144.29, 145.16, 145.98, 166.4; MS (ESI):  $m/z$ : 331  $[M+H]^+$ ; HR-MS (EI): calcd for  $C_{15}H_{14}N_4O_5$ : 330.0964; found: 330.0965.

**2-(1-Methyl-4-nitro-1H-pyrrol-2-yl)-1H-benzimidazole-5-carboxylic methyl ester (7c):** Methyl 3,4-diamino-benzoate (**5**) (2.66 g, 16.0 mmol) was dissolved in DMF (75 mL). A solution of 1-methyl-4-nitro-1H-pyrrol-2-carbaldehyde (**4c**, 2.31 g, 15.0 mmol) in DMF (200 mL) was added and the reaction mixture was heated to 90 °C for 60 min. Iron-trichloride hexahydrate (120 mg) was added and the mixture was heated to 120 °C for 6 h while air was bubbled through the solution. The reaction mixture was cooled to room temperature. The volume of the solvent was reduced in vacuo to ca. 50 mL and the brown solution was cooled to –20 °C for 4 h. The formed precipitate was filtered, washed with  $CH_2Cl_2$  until the washing solution was colorless, and dried in vacuo to yield the title compound **7c** as a yellow powder (3.13 g, 69 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 3.85 (s, 1.5H), 3.86 (s, 1.5H), 4.16 (s, 3H), 7.50–8.31 (m, 5H), 13.15 (s, 0.5H), 13.18 (s, 0.5H). At room temperature the presence of two benzimidazole tautomers (1H/3H) could be deduced from the  $^1H$  NMR spectra. The ratio between the two isomers was roughly estimated as 1:1 based on the integration of the protons at  $\delta$  = 3.85/3.86 and 13.15/13.18.  $^{13}C$  NMR (75 MHz,  $[D_6]DMSO$ ):  $\delta$  = 38.0, 52.0, 106.9, 111.0, 112.5, 118.7, 120.3, 122.9, 123.5, 123.9, 128.0, 134.8, 146.6, 166.4; MS (ESI):  $m/z$ : 301  $[M+H]^+$ ; HR-MS (EI): calcd for  $C_{14}H_{12}N_4O_4$ : 300.0858; found: 300.0847.

**2-(4-tert-Butoxycarbonylamino-1-methyl-1H-imidazol-2-yl)-1H-benzimidazole-5-carboxylic methyl ester (8a):** Nitro-ester **7a** (2.00 g, 6.64 mmol) was dissolved in DMF (200 mL). Pd/C (10 wt. %, 360 mg) was added and the mixture was hydrogenated at 600 psi for 2 h at ambient temperature. The reaction mixture was filtered through Celite to remove the catalyst and the filtrate was treated with  $Boc_2O$  (2.17 g, 9.96 mmol) and DIEA (20 mL) and stirred at 70 °C for 2 d. After evaporation of the solvent the residue was dissolved in diethyl ether (150 mL) and the solution was washed with water and brine. The organic phase was dried with  $MgSO_4$ , the solvent was evaporated, and the crude product purified by column chromatography on silica gel (*n*-hexane/EtOAc 1:1) to give the title compound **8a** as a white solid (0.74 g, 30 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 1.45 (s, 9H), 3.84 (s, 3H), 4.11 (s, 3H), 7.28 (s, 1H), 7.54–8.22 (m, 4H), 9.48 (s, 1H), 12.98, 13.01 (2s, 1H); MS (ESI):  $m/z$ : 372  $[M+H]^+$ ; HR-MS (EI): calcd for  $C_{18}H_{21}N_5O_4$ : 371.1593; found: 371.1584.

**2-(4-tert-Butoxycarbonylamino-3-methoxy-1-methyl-1H-pyrrol-2-yl)-3H-benzimidazole-5-carboxylic methyl ester (8b):** Nitro-ester **7b** (1.00 g, 3.03 mmol) was dissolved in 9:1 DMF/acetonitrile (60 mL). Tin-dichloride dihydrate (5.00 g, 22.2 mmol) was added to the solution and the reaction mixture was heated to 50 °C for 14 h. After cooling the yellow solution to room temperature  $Boc_2O$  (3.96 g, 18.1 mmol), DMF (5 mL) and DIEA (7 mL) were added and the reaction mixture heated to 50 °C for 4 h. The

solvent was evaporated, and the residue was extracted several times with ethyl acetate. The extracts were combined, washed with water, and dried with  $MgSO_4$ . After evaporation of the solvent the crude product was purified by column chromatography on silica gel ( $CH_2Cl_2$ /ethyl acetate 4:1) to afford the title compound **8b** as a white solid (0.56 g, 48 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 1.44 (s, 9H), 3.79 (s, 3H), 3.85 (s, 3H), 3.96 (s, 3H), 6.96 (s, 1H), 7.56–7.79 (m, 2H), 8.16 (s, 1H), 8.55 (s, 1H), 11.94, 12.01 (2s, 1H); MS (ESI):  $m/z$ : 401  $[M+H]^+$ ; HR-MS (EI): calcd for  $C_{20}H_{24}N_4O_5$ : 400.1746; found: 400.1742.

**2-(4-tert-Butoxycarbonylamino-1-methyl-1H-pyrrol-2-yl)-1H-benzimidazole-5-carboxylic acid methyl ester (8c):** Nitro-ester **7c** (1.50 g, 5.00 mmol) was dissolved in DMF (250 mL). Pd/C (10 wt. %, 300 mg) was added and the mixture was hydrogenated at 600 psi overnight at ambient temperature. The reaction mixture was filtered through Celite to remove the catalyst and the filtrate was treated with  $Boc_2O$  (1.63 g, 7.50 mmol) and DIEA (25 mL) and stirred at 70 °C for 2 d. After evaporation of the solvent the residue was dissolved in diethyl ether (150 mL) and the solution was washed with water and brine. The organic phase was dried with  $MgSO_4$ , the solvent was evaporated, and the crude product purified by column chromatography on silica gel ( $CH_2Cl_2$ /EtOAc 4:1) to give the title compound **8c** as a white solid (1.08 g, 58 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 1.44 (s, 9H), 3.82 (s, 3H), 4.02 (s, 3H), 6.84 (s, 1H), 6.99 (s, 1H), 7.42–8.16 (m, 3H), 9.21 (s, 1H), 12.74, 12.79 (2s, 1H); MS (ESI):  $m/z$ : 371  $[M+H]^+$ ; HR-MS (EI): calcd for  $C_{19}H_{22}N_4O_4$ : 370.1641; found: 370.1645.

**2-(4-tert-Butoxycarbonylamino-1-methyl-1H-imidazol-2-yl)-1H-benzimidazole-5-carboxylic acid (9a):** Methyl ester **8a** (700 mg, 1.88 mmol) was dissolved in dioxane (140 mL). NaOH (aqueous, 1M, 140 mL) was added and the solution stirred for 18 h at ambient temperature. The yellow solution was carefully acidified with aqueous HCl (1M) to pH 3–4. The formed precipitate was filtered, washed with water, and dried in vacuo to provide the title compound **9a** as a white powder (512 mg, 76 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 1.47 (s, 9H), 4.13 (s, 3H), 7.29 (s, 1H), 7.43–8.30 (m, 3H), 9.49 (s, 1H), 12.71 (s, 1H), 12.95 (s, 1H);  $^{13}C$  NMR (75 MHz,  $[D_6]DMSO$ ):  $\delta$  = 28.2, 35.2, 78.9, 111.5, 112.5, 113.5, 118.3, 120.6, 123.3, 124.7, 132.4, 137.9, 145.8, 152.8, 167.6; MS (ESI):  $m/z$ : 358  $[M+H]^+$ ; HR-MS (FAB): calcd for  $C_{17}H_{20}N_5O_4$ : 358.1515; found: 358.1504  $[M+H]^+$ .

**2-(4-tert-Butoxycarbonylamino-3-methoxy-1-methyl-1H-pyrrol-2-yl)-3H-benzimidazole-5-carboxylic acid (9b):** Methyl ester **8b** (400 mg, 1.00 mmol) was dissolved in dioxane (10 mL). NaOH (aqueous, 1M, 10 mL) was added and the solution stirred for 12 h at ambient temperature. The yellow solution was carefully acidified with aqueous HCl (1M) to pH 3–4, and extracted several times with diethyl ether. Evaporation of the solvent afforded the title compound **9b** as a yellow powder (351 mg, 91 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 1.43 (s, 9H), 3.81 (s, 3H), 3.98 (s, 3H), 6.96 (s, 1H), 7.61 (d,  $J$  = 8.4 Hz, 1H), 7.81 (d,  $J$  = 8.4 Hz, 1H), 8.20 (s, 1H), 8.53 (s, 1H), 11.96 (s, 1H), 12.60 (s, 1H);  $^{13}C$  NMR (75 MHz,  $[D_6]DMSO$ ):  $\delta$  = 28.3, 37.3, 61.1, 78.6, 109.6, 111.2, 112.9, 113.3, 117.3, 119.8, 120.3, 123.0, 123.8, 141.4, 147.1, 154.1, 168.0; MS (ESI):  $m/z$ : 387  $[M+H]^+$ ; HR-MS (FAB): calcd for  $C_{19}H_{23}N_4O_5$ : 387.1668; found: 387.1676  $[M+H]^+$ .

**2-(4-tert-Butoxycarbonylamino-1-methyl-1H-pyrrol-2-yl)-1H-benzimidazole-5-carboxylic acid (9c):** Methyl ester **8c** (1.04 g, 2.80 mmol) was dissolved in dioxane (280 mL). NaOH (aqueous, 1M, 280 mL) was added and the solution heated to 80 °C for 2 h. The yellow solution was carefully acidified with aqueous HCl (1M) to pH 3–4. The formed precipitate was filtered, washed with water, and dried in vacuo to provide the title compound **9c** as a white powder (990 mg, 99 %).  $^1H$  NMR (300 MHz,  $[D_6]DMSO$ ):  $\delta$  = 1.46 (s, 9H), 4.04 (s, 3H), 7.04 (s, 1H), 7.24 (s, 1H), 7.73 (d,  $J$  = 8.4 Hz, 1H), 7.96 (dd,  $J_1$  = 8.4 Hz,  $J_2$  = 1.5 Hz, 1H), 8.20 (d,  $J$  = 1.5 Hz, 1H), 9.40 (s, 1H);  $^{13}C$  NMR (75 MHz,  $[D_6]DMSO$ ):  $\delta$  = 28.2, 36.4, 78.7, 107.2, 113.5, 115.0, 120.2, 124.8, 125.6, 126.7, 132.9, 136.2, 144.2, 152.6, 166.7; MS (ESI):  $m/z$ : 357  $[M+H]^+$ ; HR-MS (FAB): calcd for  $C_{18}H_{21}N_4O_4$ : 357.1562; found: 357.1576  $[M+H]^+$ .

**2-[4-(4-tert-Butoxycarbonylamino-1-methyl-1H-pyrrol-2-carbonyl)-amino]-1-methyl-1H-imidazol-2-yl)-1H-benzimidazole-5-carboxylic methyl ester (10):** Pd/C (10 wt. %, 150 mg) was added to a solution of nitro-ester **7a** (1.00 g, 3.32 mmol) in DMF (100 mL) and the mixture was hydrogenated at 450 psi for 2 h at ambient temperature. The reaction mixture was filtered through Celite to remove the catalyst and the filtrate was treated immediately with Boc-Py-Obt (**13a**, 1.42 g, 3.98 mmol) and DIEA (20 mL) and stirred at 60 °C for 18 h. The resulting yellow solution was

poured into ice water and extracted with EtOAc. The organic phase was washed with 10% citric acid, brine, and saturated aqueous sodium bicarbonate and dried with MgSO<sub>4</sub>. Evaporation of the solvent yielded the crude product as a yellow foam which was purified by column chromatography on silica gel (*n*-hexane/EtOAc 1:2) to give the title compound **10** as a beige solid (409 mg, 25%). <sup>1</sup>H NMR (300 MHz, [D<sub>6</sub>]DMSO): δ = 1.45 (s, 9H), 3.82 (s, 3H), 3.87 (s, 3H), 4.16 (s, 3H), 6.88 (s, 1H), 6.99 (s, 1H), 7.60–8.20 (m, 4H), 9.11 (s, 1H), 10.23 (s, 1H), 13.03 (brs, 1H); MS (ESI): *m/z*: 494 [M+H]<sup>+</sup>; HR-MS (FAB): calcd for C<sub>24</sub>H<sub>28</sub>N<sub>7</sub>O<sub>5</sub>: 494.2151; found: 494.2143 [M+H]<sup>+</sup>.

**2-[4-[(4-*tert*-Butoxycarbonylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-1-methyl-1H-imidazol-2-yl]-1H-benzimidazole-5-carboxylic acid (**11**):** Methyl ester **10** (220 mg, 0.45 mmol) was dissolved in dioxane (20 mL). NaOH (aqueous, 1M, 20 mL) was added and the solution stirred for 18 h at ambient temperature. The yellow solution was carefully acidified with aqueous HCl (1M) to pH 3–4. The formed precipitate was filtered, washed with water, and dried in vacuo to provide the title compound **11** as a beige powder (203 mg, 94%). <sup>1</sup>H NMR (300 MHz, [D<sub>6</sub>]DMSO): δ = 1.44 (s, 9H), 3.82 (s, 3H), 4.16 (s, 3H), 6.88 (s, 1H), 6.99 (s, 1H), 7.59–8.18 (m, 4H), 9.12 (s, 1H), 10.23 (s, 1H), 12.95 (brs, 1H); <sup>13</sup>C NMR (75 MHz, [D<sub>6</sub>]DMSO): δ = 28.3, 35.2, 36.2, 78.3, 104.9, 114.01, 114.32, 117.8, 121.99, 122.32, 123.4, 124.71, 124.81, 132.5, 137.4, 143.7, 145.76, 145.84, 152.7, 158.6, 167.6; MS (ESI): *m/z*: 480 [M+H]<sup>+</sup>; HR-MS (ESI): calcd for C<sub>23</sub>H<sub>26</sub>N<sub>7</sub>O<sub>5</sub>: 480.1995; found: 480.2014 [M+H]<sup>+</sup>.

**Solid-phase synthesis:** Polyamides **1a**, **b**, **2a**, **b**, **3a**, **b** were generated by manual solid-phase synthesis on oxime resin according to recently reported protocols<sup>[14]</sup> using Boc-protected amino acid building blocks. The synthesis of polyamide ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-ImPyPyCONHMe (**1a**) has been previously described.<sup>[29]</sup> For the coupling of benzimidazole building blocks Boc protected amino acids **9a–c** and **11** were activated with HBTU (1 equiv in DIEA, NMP) for 15 min and subsequently coupled for 4 h at ambient temperature. Resultant resin-bound benzimidazoles were deprotected using 20% TFA in CH<sub>2</sub>Cl<sub>2</sub> for 25 min (for resin-bound **9c** and **11**) or 50% TFA in CH<sub>2</sub>Cl<sub>2</sub> for 25 min (for resin-bound **9a,b**).

**General procedure for MeNH<sub>2</sub> cleavage:**<sup>[14b]</sup> A sample of the derivatized resin (100 mg) was suspended in CH<sub>2</sub>Cl<sub>2</sub> (2 mL) to which was added methylamine in THF (2 mL, 2.0M). The mixture was agitated at 37 °C for 14 h. The resin was filtered, rinsed with CH<sub>2</sub>Cl<sub>2</sub> and DMF, the eluant concentrated in vacuo, and the residue purified by prep. reversed-phase HPLC to yield the desired polyamides.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-ImPyBiPyCONHMe (**1b**):** The title compound **1b** was recovered upon lyophilization of the appropriate fractions as a yellow powder. MALDI-TOF-MS: calcd for C<sub>51</sub>H<sub>53</sub>N<sub>21</sub>O<sub>8</sub>: 1090.1; found: 1090.6.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-PyImPyPyCONHMe (**2a**):** The title compound **2a** was recovered upon lyophilization of the appropriate fractions as a white powder. MALDI-TOF-MS: calcd for C<sub>49</sub>H<sub>56</sub>N<sub>22</sub>O<sub>9</sub>: 1097.1; found: 1097.6.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-PyImBiPyCONHMe (**2b**):** The title compound **2b** was recovered upon lyophilization of the appropriate fractions as a yellow powder. MALDI-TOF-MS: calcd for C<sub>50</sub>H<sub>54</sub>N<sub>22</sub>O<sub>8</sub>: 1091.8; found: 1091.1.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-ImOpPyPyCONHMe (**12a**):** The title compound **12a** was recovered upon lyophilization of the appropriate fractions as a beige powder. MALDI-TOF-MS: calcd for C<sub>51</sub>H<sub>59</sub>N<sub>21</sub>O<sub>10</sub>: 1126.2; found: 1126.7.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-ImOpBiPyCONHMe (**12b**):** The title compound **12b** was recovered upon lyophilization of the appropriate fractions as a yellow powder. MALDI-TOF-MS: calcd for C<sub>52</sub>H<sub>57</sub>N<sub>21</sub>O<sub>9</sub>: 1120.2; found: 1120.7.

**O-Demethylation of polyamides 12a,b:** According to the previously reported protocol<sup>[19]</sup> polyamides were treated with sodium thiophenoxide in DMF at 85 °C for 2 h and subsequently purified by prep. reversed-phase HPLC to yield the desired demethylated polyamides.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-ImHpPyPyCONHMe (**3a**):** The title compound **3a** was recovered upon lyophilization of the appropriate fractions as a beige powder. MALDI-TOF-MS: calcd for C<sub>50</sub>H<sub>57</sub>N<sub>21</sub>O<sub>10</sub>: 1112.1; found: 1113.1.

**ImImPyPy-(R)<sup>H<sub>2</sub>N</sup>-ImHpBiPyCONHMe (**3b**):** The title compound **3b** was recovered upon lyophilization of the appropriate fractions as a yellow powder. MALDI-TOF-MS: calcd for C<sub>51</sub>H<sub>55</sub>N<sub>21</sub>O<sub>9</sub>: 1106.1; found: 1106.5.

**Construction of plasmid DNA:** The restriction fragment pAU2 was generated as previously described.<sup>[21]</sup> Similarly, plasmids pPWF2, pCAB1 and pCAB2 were constructed by ligation of the following hybridized inserts into the *Bam*HI/*Hin*DIII polycloning site of pUC19.

Plasmid **pPWF2:** 5'-GATCATGAACATCGATCTCTATGTACATGC-TATGCGATGGACATATCAGCTATGCACATTCGATCTCTATGC-TATGCGA-3'.

Plasmid **pCAB1:** 5'-GATCAGTGACAATCGATCTCTGTGTCAATGATATGCGGTGGCAATATGAGCTGTGCCAATTCGATCTCTATGATATACGA-3'.

Plasmid **pCAB2:** 5'-GATCAGTGGAAATCGATCTCTGTGGTAATGATATGCGGTGGGAATATGAGCTGTGCCAATTCGATCTCTATGATATACGA-3'.

The inserts were obtained by annealing complementary synthetic oligonucleotides and were then ligated to the large *Bam*HI/*Hin*DIII restriction fragment of pUC19 using T4 DNA ligase. The ligated plasmids were then used to transform *E. coli* XL-1 Blue Supercompetent cells. Colonies were selected for  $\alpha$ -complementation on 25 mL Luria-Bertani agar plates containing 50 mg mL<sup>-1</sup> ampicillin and treated with XGAL and IPTG solutions and grown overnight at 37 °C. Well defined white colonies were transferred into 100 mL Luria-Bertani medium containing 50 mg mL<sup>-1</sup> ampicillin. Cells were harvested after overnight growth at 37 °C. Large scale plasmid purification was performed using Qiagen purification kits. The presence of the desired insert was determined by dideoxy sequencing. DNA concentration was determined at 260 nm using the relation 1 OD unit = 50  $\mu$ g mL<sup>-1</sup> duplex DNA.

**Preparation of 3'-end-labeled DNA restriction fragments:** Plasmids pPWF2, pCAB1, pAU2,<sup>[21]</sup> and pCAB2 were linearized with *Eco*RI and *Pvu*II and then radiolabelled by 3'-fill in using [ $\alpha$ -<sup>32</sup>P]-dATP, [ $\alpha$ -<sup>32</sup>P]-TTP, and the Klenow fragment of DNA polymerase II at 37 °C for 25 min. The product mixture was purified on a 7% non-denaturing preparatory polyacrylamide gel (5% cross-linkage) and the desired fragment was isolated after visualization by autoradiography. The DNA was precipitated with isopropanol (1.5 volumes). The pellet was washed with 75% ethanol, lyophilized to dryness, and then resuspended in RNase-free H<sub>2</sub>O. Chemical sequencing reactions were performed according to published protocols.<sup>[30]</sup>

**Quantitative DNase I footprint titrations:** Concentrations of benzimidazole-containing polyamide and parent polyamide stock solutions were determined by UV absorption at 331 nm ( $\epsilon = 41\,500$  cm<sup>-1</sup>) and 310 nm ( $\epsilon = 69\,500$  cm<sup>-1</sup>), respectively. All reactions were carried out in a volume of 400  $\mu$ L according to published procedures.<sup>[20]</sup> Quantitation by storage phosphor autoradiography and determination of equilibrium association constants were as previously described.<sup>[20]</sup>

## Acknowledgement

We are grateful to the National Institute of Health (GM-27681) for research support, to the Alexander von Humboldt-Foundation for a Feodor-Lynen grant to C.A.B. and to the Swiss National Science Foundation for a postdoctoral fellowship to P.W. We also wish to thank Michael A. Marques for conducting the ab initio calculations.

- [1] a) C. L. Kielkopf, S. White, J. W. Szewczyk, J. M. Turner, E. E. Baird, P. B. Dervan, D. C. Rees, *Science* **1998**, *282*, 111–115; b) P. B. Dervan, R. W. Bürl, *Curr. Opin. Chem. Biol.* **1999**, *3*, 688–693; c) D. E. Wemmer, *Biopolymers* **1999**, *52*, 197–211; d) P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.
- [2] a) J. M. Gottesfeld, L. Neely, J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1997**, *387*, 202–205; b) L. A. Dickinson, R. J. Gulizia, J. W. Trauger, E. E. Baird, D. E. Mosier, J. M. Gottesfeld, P. B. Dervan, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12890–12895; c) S. Janssen, T. Durussel, U. K. Laemmli, *Mol. Cell* **2000**, *6*, 999–1011; d) A. Z. Ansari, A. K. Mapp, D. H. Nguyen, P. B. Dervan, M. Ptashne, *Chem. Biol.* **2001**, *8*, 583–592; e) J. J. Coull, G. He, C. Melander, V. C. Rucker, P. B. Dervan, D. M. Margolis, *J. Virology* **2002**, *76*, 12349–12354.
- [3] J. W. Trauger, E. E. Baird, P. B. Dervan, *Nature* **1996**, *382*, 559–561.
- [4] D. M. Herman, J. M. Turner, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 1121–1129.
- [5] W. A. Greenberg, E. E. Baird, P. B. Dervan, *Chem. Eur. J.* **1998**, *4*, 796–805.

- [6] a) P. Weyermann, P. B. Dervan, *J. Am. Chem. Soc.* **2002**, *124*, 6872–6878; b) I. Kers, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 3339–3349.
- [7] a) Z. Y. J. Zhan, P. B. Dervan, *Bioorg. Med. Chem.* **2000**, *8*, 2467–2474; b) D. H. Nguyen, J. W. Szewczyk, E. E. Baird, P. B. Dervan, *Bioorg. Med. Chem.* **2001**, *9*, 7–17; c) C. C. O'Hare, D. Mack, M. Tandon, S. K. Sharma, J. W. Lown, M. L. Kopka, R. E. Dickerson, J. A. Hartley, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 72–77; d) M. A. Marques, R. M. Doss, A. R. Urbach, P. B. Dervan, *Helv. Chim. Acta* **2002**, *85*, 4485–4517.
- [8] a) W. S. Wade, M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **1992**, *114*, 8783–8794; b) M. Mrksich, P. B. Dervan, *J. Am. Chem. Soc.* **1993**, *115*, 9892–9899.
- [9] U. Ellervik, C. C. C. Wang, P. B. Dervan, *J. Am. Chem. Soc.* **2000**, *122*, 9354–9360.
- [10] a) R. L. Lombardy, F. A. Tanius, K. Ramachandran, R. R. Tidwell, W. D. Wilson, *J. Med. Chem.* **1996**, *39*, 1452–1462; b) T. G. Minehan, K. Gottwald, P. B. Dervan, *Helv. Chim. Acta* **2000**, *83*, 2197–2213; c) L. Wang, C. Carrasco, A. Kumar, C. E. Stephens, C. Bailly, D. W. Boykin, W. D. Wilson, *Biochemistry* **2001**, *40*, 2511–2521; d) W. Zhang, Y. Dai, U. Schmitz, T. W. Bruice, *FEBS Lett.* **2001**, *509*, 85–89; e) Y.-H. Ji, D. Bur, W. Haesler, V. R. Schmitt, A. Dorn, C. Bailly, M. J. Waring, R. Hochstrasser, W. Leupin, *Bioorg. Med. Chem.* **2001**, *9*, 2905–2919; f) A. L. Satz, T. C. Bruice, *J. Am. Chem. Soc.* **2001**, *123*, 2469–2477; g) C. Behrens, N. Harrit, P. E. Nielsen, *Bioconjugate Chem.* **2001**, *12*, 1021–1027.
- [11] Y. Matsuba, H. Edatsugi, I. Mita, A. Matsunaga, O. Nakanishi, *Cancer Chemother. Pharmacol.* **2000**, *46*, 1–9.
- [12] T. Stokke, H. B. Steen, *J. Histochem. Cytochem.* **1985**, *33*, 333–338.
- [13] a) P. E. Pjura, K. Grzeskowiak, R. E. Dickerson, *J. Mol. Biol.* **1987**, *197*, 257–271; b) M. Teng, N. Usman, C. A. Frederick, A. Wang, *Nucleic Acids Res.* **1988**, *16*, 2671–2690; c) S. Kumar, B. Yadagiri, J. Zimmermann, R. T. Pon, J. W. Lown, *J. Biomol. Struct. Dyn.* **1990**, *8*, 331–357.
- [14] a) E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1996**, *118*, 6141–6146; b) J. M. Belitsky, D. H. Nguyen, N. R. Wurtz, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 2767–2774.
- [15] M. P. Singh, S. Sasmal, W. Lu, M. N. Chatterjee, *Synthesis* **2000**, 1380–1390.
- [16] J. B. Wright, *Chem. Rev.* **1951**, *48*, 397–541.
- [17] Y. Yamamoto, T. Kimachi, Y. Kanaoka, S. Kato, K. Bessho, T. Matsumoto, T. Kusakabe, Y. Sugiura, *Tetrahedron Lett.* **1996**, *37*, 7801–7804.
- [18] Moderate to low reaction conversions are often observed for couplings of Boc-Py-OBt (**13a**) to resin-bound imidazole amines.<sup>[14a]</sup> To increase the coupling efficiency the use of PyIm dimers proved to be advantageous. Alternatively, Py to Im couplings can be performed by using Boc-Py-OH (**13b**) in the presence of DCC and DMAP.<sup>[14a]</sup>
- [19] C. Melander, D. M. Herman, P. B. Dervan, *Chem. Eur. J.* **2000**, *6*, 4487–4497.
- [20] J. W. Trauger, P. B. Dervan, *Methods Enzymol.* **2001**, *340*, 450–466.
- [21] A. R. Urbach, J. W. Szewczyk, S. White, J. M. Turner, E. E. Baird, P. B. Dervan, *J. Am. Chem. Soc.* **1999**, *121*, 11621–11629.
- [22] R. Kakkar, R. Garg, Suruchi, *J. Mol. Struct.* **2002**, *584*, 37–44; and references therein.
- [23] D. E. Wemmer, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 439–461.
- [24] Gas-phase ab initio HF/6-31G\* energy minimization; *Spartan Essential Software*, Wavefunction Inc., **2001**.
- [25] D. Renneberg, P. B. Dervan, unpublished observations.
- [26] J. M. Belitsky, S. J. Leslie, P. S. Arora, T. A. Beerman, P. B. Dervan, *Bioorg. Med. Chem.* **2002**, *10*, 3313–3318.
- [27] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, **1989**.
- [28] D. P. Phillion, M. Singh (Monsanto Technology LLC), WO 02/04417 A1, **2002** [*Chem. Abstr.* **2002**, *136*, 102283].
- [29] N. R. Wurtz, J. L. Pomerantz, D. Baltimore, P. B. Dervan, *Biochemistry* **2002**, *41*, 7604–7609.
- [30] a) B. L. Iverson, P. B. Dervan, *Nucleic Acids Res.* **1987**, *15*, 7823–7830; b) A. M. Maxam, W. S. Gilbert, *Methods Enzymol.* **1980**, *65*, 499–560.

Received: December 19, 2002 [F4689]