

## **Chapter 5: Allosteric Modulation of DNA by Small Molecules**

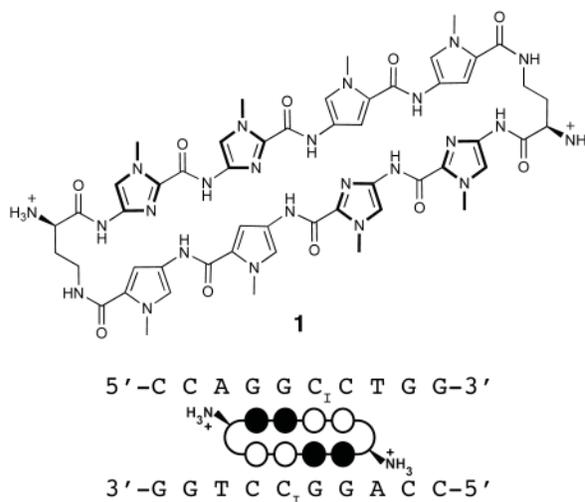
**Abstract**

Many human diseases are caused by dysregulated gene expression. The oversupply of one or more transcription factors may be required for the growth and metastatic behavior of human cancers. Cell permeable small molecules which can be programmed to disrupt transcription factor-DNA interfaces could silence aberrant gene expression pathways. Pyrrole-imidazole polyamides are DNA minor-groove binding molecules that are programmable for a large repertoire of DNA motifs. A high resolution X-ray crystal structure of an 8-ring cyclic Py/Im polyamide bound to the central six base pairs of the sequence d(5'-CCAGGC<sub>1</sub>CTGG-3')<sub>2</sub> reveals a 4 Å widening of the minor groove and compression of the major groove along with a >18° bend in the helix axis toward the major groove. This allosteric perturbation of the DNA helix provides a molecular basis for disruption of transcription factor-DNA interfaces by small molecules, a minimum step in chemical control of gene networks.

## 5.1 Introduction

Py/Im polyamides bind the minor groove of DNA sequence specifically,<sup>1,2</sup> encoded by side-by-side arrangements of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) carboxamide monomers. Im/Py pairs distinguish G•C from C•G base pairs, whereas Py/Py pairs are degenerate for T•A and A•T.<sup>3-6</sup> Antiparallel Py/Im strands are connected by  $\gamma$ -aminobutyric acid linker (GABA) to create a hairpin-shaped oligomer. Hairpin Py/Im polyamides have been programmed to bind a broad library of different DNA sequences.<sup>7</sup> They have been shown to permeate cell membranes,<sup>8-10</sup> access chromatin,<sup>11,12</sup> and disrupt protein-DNA interactions.<sup>2</sup> Disruption of transcription factor-DNA interfaces six bp in size such as HIF-1 $\alpha$ ,<sup>13-15</sup> androgen receptor (AR),<sup>16</sup> and AP-1<sup>17,18</sup> have been exploited for controlling expression of medically relevant genes such as VEGF, PSA, TGF- $\beta$ 1 and LOX-1 in cell culture experiments.<sup>13-18</sup> X-ray crystallography of antiparallel 2:1 binding polyamides in complex with DNA reveal a 1 Å widening of the minor groove.<sup>5,6</sup> This modest structural perturbation to the DNA helix by the side-by-side stacked arrangement of aromatic rings does not explain the large number of transcription factor-DNA interfaces disrupted by minor-groove binding hairpin Py/Im polyamides.<sup>2,5,6,19</sup> It must be that the turn unit in the hairpin oligomer connecting the two antiparallel strands plays a structural role.

Here we report the atomic resolution structure (1.18 Å resolution) of an 8-ring cyclic polyamide in complex with double helical DNA. The cyclic polyamide **1** is comprised of two antiparallel ImImPyPy strands capped by (*R*)- $\alpha$ -amino- $\gamma$  turn units. Polyamide **1**, which codes for the sequence 5'-WGGCCW-3' was co-crystallized with the palindromic DNA oligonucleotide sequence 5'-CCAGGC<sub>1</sub>CTGG-3' 10 base pairs in length (Figure 5.1). We observe significant structural allosteric perturbations of the DNA helix induced upon binding of GABA ( $\gamma$ -aminobutyric acid) turn-linked polyamides in the minor groove. A detailed view of the  $\alpha$ -amino- $\gamma$ -turn conformation and hydration

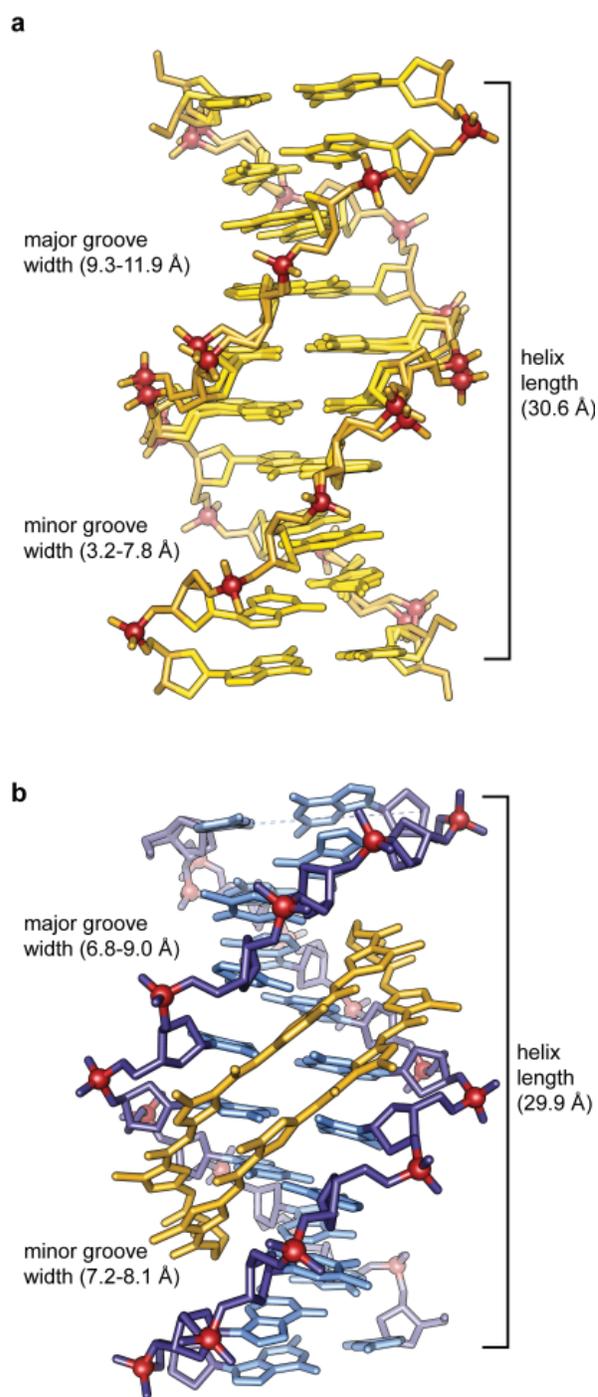


**Figure 5.1** Chemical structure of the cyclic polyamide and DNA sequence. Cyclic polyamide **1** targeting the sequence 5'-WGGCCW-3' shown with ball-and-stick model superimposed onto the DNA oligonucleotide used for crystallization. (Black circles represent imidazoles, open circles represent pyrroles, and ammonium substituted half circles at each end represent the (*R*)- $\alpha$ -amine- $\gamma$ -turn.

reveal a network of well-ordered water-mediated interactions between the polyamide and the minor groove floor of DNA.

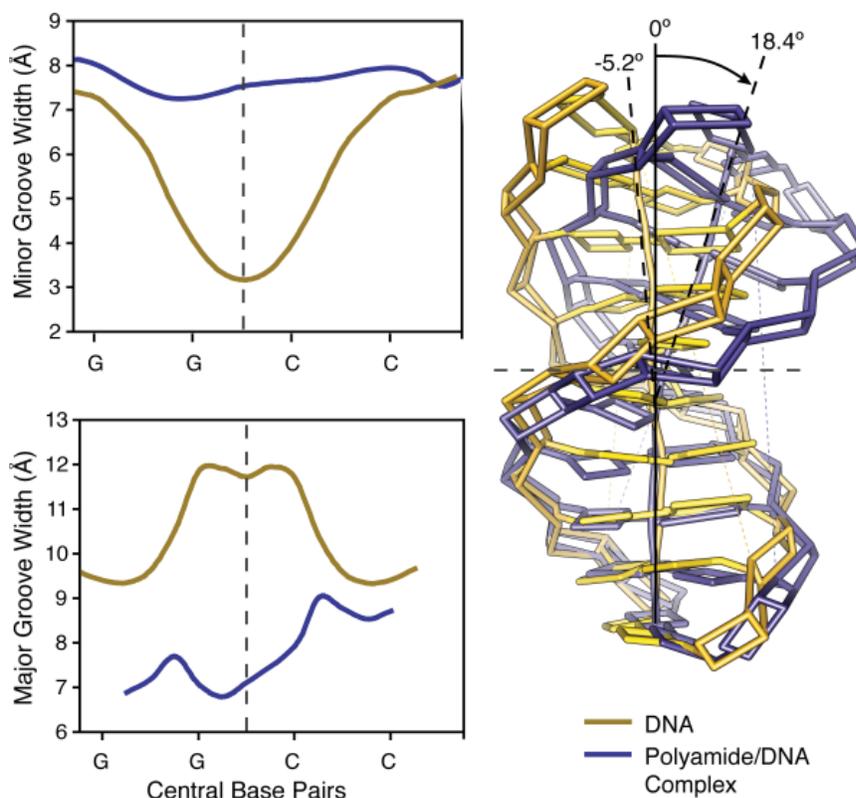
## 5.2 Results and Discussion

The structure was solved by direct methods to 1.18 Å resolution with synchrotron radiation. One cyclic polyamide bound to a single DNA duplex is present in the asymmetric unit of the crystal in the P1 space group. In the DNA complex the aromatic amino acids are bound with an N- to C- orientation of each ImImPyPy strand of the cycle adjacent to the 5' to 3' direction of the DNA. To assess DNA structural perturbations imposed by polyamide binding we compared our polyamide-DNA complex to the free DNA. We solved the X-ray structure of  $d(\text{CCAGGC}_1\text{CTGG})_2$  to a resolution of 0.98 Å for comparison.<sup>20</sup> Interestingly, this structure shows several discrete alternate conformations in 7 of the 10 nucleotides in each strand of the DNA duplex, illustrating the dynamic and conformationally mobile nature of the B-DNA sugar-phosphate backbone. The comparison of polyamide-DNA complex to free DNA is shown in Figure 5.2. Polyamide binding locks out the alternate DNA conformations, rigidifying the sugar-phosphate backbone, and strongly perturbing the overall helix structure. Binding of the polyamide widens the minor groove up to 4 Å while simultaneously compressing the major



**Figure 5.2** Comparison of native DNA to polyamide/DNA complex. a) Native DNA crystal structure at 0.98 Å resolution. b) Comparison to DNA/polyamide co-crystal structure at 1.18 Å resolution. (Both structure solved by direct methods.) showing the bound cyclic polyamide with electron density contoured at the 1.0  $\sigma$  level.

groove by 4 Å. The polyamide bends the DNA strand  $>18^\circ$  toward the major groove as shown in Figure 5.3, and shortening of the overall length of the helix by  $\sim 1$  Å (Figure 5.2).

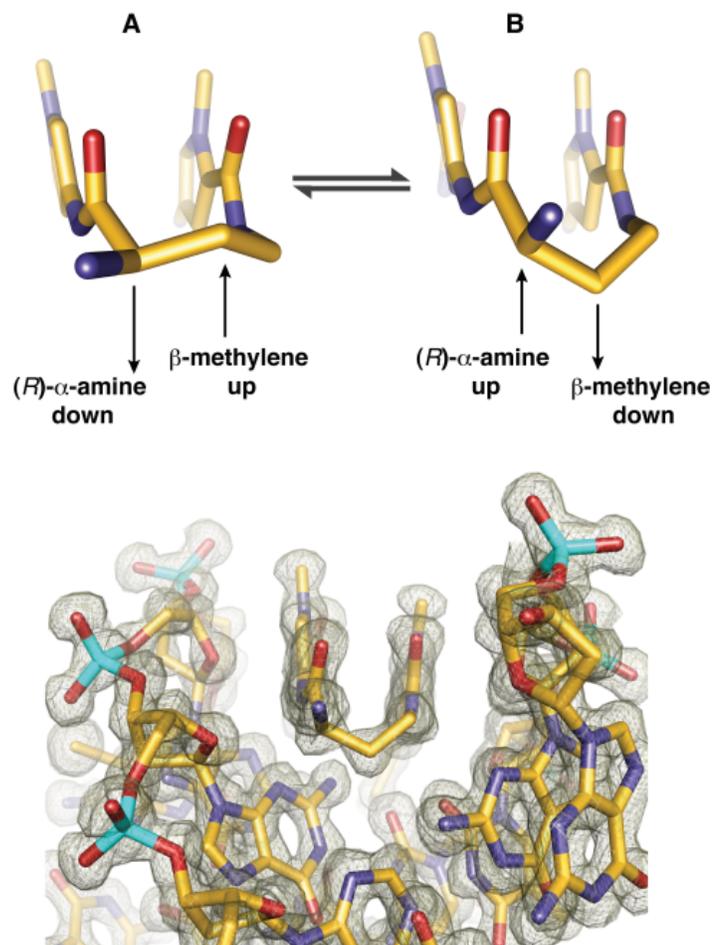


**Figure 5.3** Analysis of native DNA (yellow) compared to polyamide complexed DNA (blue). Chart on the top left shows variation in the minor groove width for native DNA (yellow) and polyamide-complexed DNA (blue) over the central core sequence 5'-GGCC-3'. Chart on the bottom left shows variation in the major groove width for native DNA (yellow) and polyamide complexed DNA (blue) over the central core sequence 5'-GGCC-3'. Overlay of the Curves calculated geometric helix model from each structure showing a DNA bend of  $> 18^\circ$  in the polyamide/DNA complex compared to native DNA.

Py/Im Polyamides linked by a GABA or substituted GABA can adopt either of two possible conformations shown in Figure 5.4. In conformation **A**, the amino group is directed toward the minor-groove wall of DNA with the potential for steric clash with the deoxyribose sugar. In conformation **B** the amine is directed up and out of the minor groove forcing the  $\beta$ -methylene to the floor of the minor groove with the potential for steric interaction with the nucleobases within van der Waals contact distance of the C2 hydrogen of adenine. We observe conformation **B** in our high resolution X-ray structure at both ends of the complex (Figure 5.4). It is possible that there is an intrinsic preference for conformation **A**, which relieves the  $\beta$ -methylene interaction with the floor of the minor groove.

However, for turn substitution at the  $\alpha$ -position interaction with the minor-groove wall becomes the dominant steric interaction, leading to conformational inversion. Figure 5.4 presents a view of the complex looking down the minor groove directly at the polyamide turn linkage. From this view, significant van der Waals interactions can be observed between the outside face of the pyrrole-imidazole strands and the walls of the minor groove, which form a deep binding pocket for the cycle. Approximately 40% of the polyamide surface area is buried leaving only the top of the methyl groups on the heterocycles, the amide carbonyl oxygens, and the chiral  $\alpha$ -ammonium turn solvent exposed. In addition the turn unit introduces conformational constraints preventing slipped or linear binding modes.

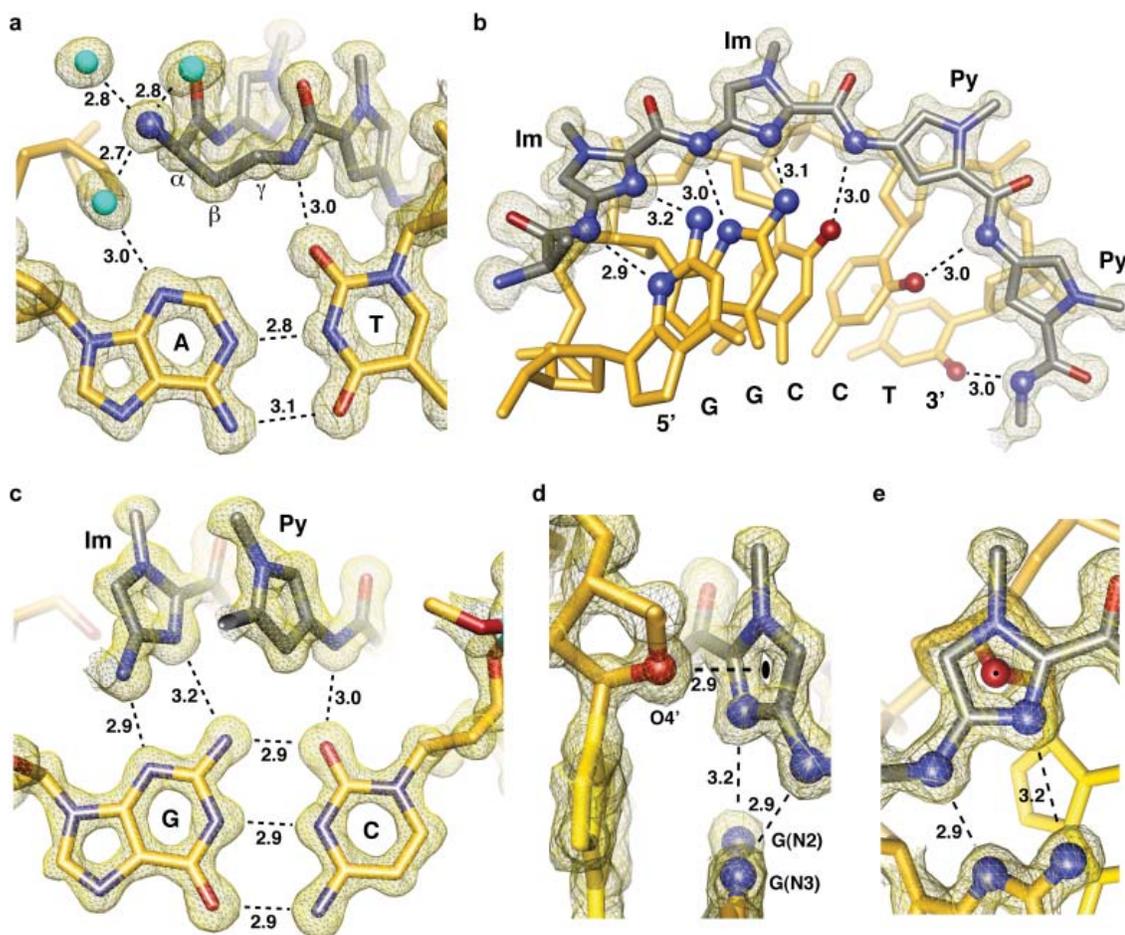
The conformational constraints imposed by the turn linkage result in ring placement that is an intermediate of ring-over-ring and ring-over-amide between adjacent PyPyImIm strands. This alignment allows the ring pairs to remain in phase with the nucleobases as the polyamide adopts an isohelical conformation complementary to the DNA helix. This is highlighted by comparison to the 2:1 structure in which the rings lie over the carboxamide linkages of the adjacent strand.<sup>5</sup> The conformational constraints



**Figure 5.4** Conformation of the  $\alpha$ -amino substituted GABA turn. (top), Two possible Conformations **A** and **B** are shown with conformation **A** directing the  $\beta$ -methylene up and away from the minor groove floor while orienting the  $\alpha$ -ammonium toward the minor groove wall. Conformation **B** presents the  $\beta$ -methylene down toward the minor-groove floor while orienting the  $\alpha$ -ammonium up and out of the minor-groove, relieving possible steric interaction with the sugar-phosphate backbone (minor-groove wall). (bottom), View looking down the DNA minor-groove, showing the  $(R)$ - $\alpha$ -amine- $\gamma$ -turn conformation observed in the X-ray crystal structure, which matches that of conformation **B**. Electron density map is contoured at the  $1.0 \sigma$  level.

imposed by the turn and inability of the ligand to slip into a possibly more preferred orientation may impact the overall DNA structure by inducing bending and other allosteric distortions accommodated by the plasticity of DNA. In addition, the turn-constrained cycle may have a major entropic driving force leading to substantial pre-organization, increased affinity, and increased specificity by locking out unproductive conformations and alternate binding modes. The van der Waals interactions between rings may also lead to cooperativity in the binding process. In addition, we find a shell of highly ordered water molecules around the  $\alpha$ -ammonium substituent and a water-mediated hydrogen bond from the ammonium to the N3 lone-pair of the adenine under the turn. The hydration pattern around the turn is highly conserved at both ends of the structure and the water mediated hydrogen bonds are within  $\sim 2.7$ – $2.9$  Å from the ammonium to water to the adenine lone-pair (Figure 5.5a).

The amide NH's and imidazole lone-pairs form a continuous series of direct hydrogen bonds to the floor of the DNA minor-groove, while the imidazoles impart specificity for the exocyclic amine of guanine through relief of a steric interaction and a G(N2-hydrogen)-Im (lone pair) hydrogen bond. The amides linking the aromatic rings and the turns contribute hydrogen bonds to the purine N3 and pyrimidine O2 lone pairs. All amides are within hydrogen bonding distance of a single DNA base ( $\sim 3.0$  Å average, Figure 5.9). In all there are 10 direct amide hydrogen bonds (average distance =  $2.97$  Å), 4 direct imidazole hydrogen bonds (2 terminal average distance =  $3.27$  Å and 2 internal average distance =  $3.05$  Å), and 2 (*R*)- $\alpha$ -ammonium turn water-mediated hydrogen bonds (average distance amine to water =  $2.75$  Å and average distance from the water to adenine N3 =  $2.98$  Å) to the floor of the DNA minor groove with at least one interaction for all 12 DNA base-pairs in the 6 bp binding site for a total of 16 hydrogen bond interactions between the cyclic polyamide and the floor of the DNA minor-groove. These 16 hydrogen bonds utilize every heteroatom of the polyamide presented to the floor of the DNA minor-groove, which exactly matches the total number of Watson-Crick hydrogen bonds between all the DNA base pairs in the 6 bp binding site. In addition to these 16 hydrogen bonds, we find unique weak interactions in the form of lone pair- $\pi$  interactions<sup>21,22</sup> between the center of the leading imidazole ring and the lone pair of the adjacent deoxyribose O4' oxygen (Figure 5.5d and 5.5e). Interestingly this interaction is only observed for the terminal imidazole aromatic ring and analysis of qualitative electrostatic potential surfaces substantiates the electropositive nature of the imidazole under these conditions (Figure 5.13).<sup>23</sup>



**Figure 5.5** Direct and water-mediated non-covalent molecular recognition interactions. a) Geometry of the  $\alpha$ -amino turn interacting with the AT base pair through water-mediated hydrogen bonding interactions. Structural basis for the turn preference for AT versus GC is demonstrated by the  $\beta$ -methylene conformational preference, which points down toward the DNA minor-groove floor within van der Waals contact distance of the adenine base. b) Isolated view of one half of the macrocyclic-polyamide showing hydrogen bond distances made to the DNA minor groove floor by the imidazoles and amides of compound **1**. c) Im-Py pair showing the mechanism for GC specificity. d) Interaction of the O4' oxygen of a deoxyribose sugar with the terminal imidazole aromatic ring through a lone pair- $\pi$  interaction. The sugar conformation is C2'-endo at the N-terminal imidazole of the polyamide with the sugar oxygen lone-pair pointing directly to the centroid of the imidazole ring. The distance between the sugar oxygen and the ring centroid is 2.90 Å, which is less than the sum of the van der Waals radii to any atom in the imidazole ring. Electrostatic potential maps calculated at the HF/3-21g\* level of theory show the slightly electropositive nature of the imidazole ring under these conditions (Figure 5.13). e) View of the O4' deoxyribose oxygen atom looking through the imidazole ring showing the ring centroid superimposed on the oxygen atom. All distances are reported in angstroms (Å) and all electron density maps are contoured at the 1.0  $\sigma$  level (Im = imidazole and Py = pyrrole).

### 5.3 Conclusion

The crystal structure presented highlights the molecular recognition of turn-linked polyamides in the minor-groove of DNA and provides insight into the allosteric modulation of

B-form DNA by hairpin oligomers. The DNA structural distortion induced upon polyamide minor-groove binding provides an allosteric model for disrupting DNA:transcription factor interfaces in the promoters of selected genes. The ability of DNA to undergo short and long-range allosteric effects coupled with DNA binding by proteins can have influence over important processes such as modulation of eukaryotic gene networks.<sup>24-27</sup> Allosteric communication along and through the DNA helix forms the basis for cooperative interactions among transcription factor regulatory networks such as the interferon- $\beta$  enhanceosome.<sup>25</sup> The potential for allosteric control over the transcriptional machinery provides a powerful concept for the design of small molecules that can bind to distinct locations on DNA with the possibility of modulating transcription factor activity.<sup>1,2,19,27</sup>

## 5.4 Experimental Section

### 5.4.1 General

Chemicals and solvents were purchased from Sigma-Aldrich and Hampton Research and were used without further purification. Water (18 M $\Omega$ ) was purified using a Millipore MilliQ purification system. Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250 x 4.6 mm, 5  $\mu$ M), a diode array detector, and the mobile phase consisted of a gradient of acetonitrile (MeCN) in 0.1% (v/v) aqueous trifluoroacetic acid (TFA). Preparative HPLC was performed on an Agilent 1200 system equipped with a solvent degasser, diode array detector, and a Phenomenex Gemini column (5  $\mu$ m particle size, C18 110A, 250 x 21.2 mm, 5 micron). A gradient of MeCN in 0.1% (v/v) aqueous trifluoroacetic acid (TFA) was utilized as the mobile phase. UV-Vis measurements were made on a Hewlett-Packard diode array spectrophotometer (model 8452 A) and polyamide concentrations were measured in 0.1% (v/v) aqueous TFA using an extinction coefficient of 69200 M<sup>-1</sup>cm<sup>-1</sup> at  $\lambda_{\text{max}}$  near 310 nm. Matrix-assisted LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) was performed on an Applied Biosystems Voyager DR Pro spectrometer using  $\alpha$ -cyano-4-hydroxycinnamic acid as matrix.

### 5.4.2 Synthesis

Polyamide **1** was synthesized by standard solid-phase synthesis methods<sup>28,29</sup> on oxime resin (Figure 5.6) and purified by reverse-phase high-performance liquid chromatography (Figure 5.7).

### 5.4.3 Cyclo-(-ImImPyPy-(R) <sup>$\alpha$</sup> -BocHN $\gamma$ -ImImPyPy-(R) <sup>$\alpha$</sup> -H<sub>2</sub>N $\gamma$ -) (6)

Oxime resin (**R3**) was generated by manual solid-phase synthesis from Kaiser oxime resin (1 g, 0.48 mmol/g, Novabiochem) using previously described Boc-protected monomers.<sup>28,29</sup> Boc-Py-OBt (**2**) (716 mg, 2 mmol) was dissolved in 2 mL of DMF and added to 1 g of oxime resin followed by 1 mL of DIEA. The reaction was left in a 37 °C shaker for 12 h. The resin was drained, washed with DMF 3x, DCM 3x, and the Boc group was removed upon treatment with 20% TFA/DCM for 30 min. After draining the resin and washing with DCM 3x followed by DMF 3x the second pyrrole residue was coupled in the same fashion as the first, with complete coupling after 2 h at 23 °C. Boc-Im-OH (**3**) (482 mg, 2 mmol) was dissolved in 2 mL of DMF and treated with 1.14 g (2 mmol) of PyBOP and 2 mL of DIEA. This solution was stirred for 5 minutes prior to addition to the resin vessel. Coupling was allowed to proceed for 2 h at 23 °C. The Boc-Im residue was deprotected using a 50% TFA/DCM solution for 30 min at room temperature followed by draining the resin and washing with DCM 3x and DMF 3x. A second imidazole residue was coupled following the exact same procedure as the first. The turn unit,  $\alpha$ -Fmoc- $\gamma$ -Boc-(*R*)-diaminobutyric acid (Fmoc-D-Dab-(Boc)-OH) (**4**) (660 mg, 1.5 mmol) was activated with PyBOP (855 mg, 1.5 mmol) in 2 mL of DMF and 1 mL of DIEA at 23 °C for 15 min prior to addition to the resin. Coupling was allowed to proceed for 2 h at 37 °C. After deprotection with 20% TFA/DCM for 30 min, the next two pyrrole residues were attached in exactly the same manner as previously described using Boc-Py-OBt. The last two imidazoles were added in the same fashion as the previous two. The final turn unit,  $\alpha$ -Boc- $\gamma$ -Fmoc-(*R*)-diaminobutyric acid (Boc-D-Dab-(Fmoc)-OH) (**5**) (660 mg, 1.5 mmol) was activated with PyBOP (855 mg, 1.5 mmol) in 2 mL of DMF and 1 mL of DIEA at 23 °C for 15 min prior to addition to the resin. Coupling was allowed to proceed for 2 h at 37 °C. After Fmoc deprotection with 25% piperidine/DMF for 3x5 min, the resin was washed with DMF 6x and 1:1 DMF/DIEA 3x. Next, the resin was diluted with 10 mL of DMF and stored in a 37 °C shaker on medium speed for 24 h. The resin was filtered off and the DMF concentrated to 1 mL volume, taken up in 9 mL of H<sub>2</sub>O (0.1% TFA), and purified by preparative reverse-phase HPLC to give *cyclo*-(-ImImPyPy-(*R*) <sup>$\alpha$ -BocHN $\gamma$</sup> -ImImPyPy-(*R*) <sup>$\alpha$ -H<sub>2</sub>N $\gamma$</sup> -) (**6**) as a fluffy white solid in 0.1% overall yield (0.480  $\mu$ mol). *Cyclo*-(-ImImPyPy-(*R*) <sup>$\alpha$ -BocHN $\gamma$</sup> -ImImPyPy-(*R*) <sup>$\alpha$ -H<sub>2</sub>N $\gamma$</sup> -) (**6**) MALDI-TOF MS (*m/z*): calc'd for C<sub>57</sub>H<sub>68</sub>N<sub>24</sub>NaO<sub>12</sub> [M+Na]<sup>+</sup> 1303.53, found 1303.36.

#### 5.4.4 *Cyclo*-(-ImImPyPy-(*R*) <sup>$\alpha$ -H<sub>2</sub>N $\gamma$</sup> -ImImPyPy-(*R*) <sup>$\alpha$ -H<sub>2</sub>N $\gamma$</sup> -) (**1**)

A solution of *cyclo*-(-ImImPyPy-(*R*) <sup>$\alpha$ -BocHN $\gamma$</sup> -ImImPyPy-(*R*) <sup>$\alpha$ -H<sub>2</sub>N $\gamma$</sup> -) (**6**) (0.400  $\mu$ mol) in anhydrous TFA/DCM (1:1, 500  $\mu$ L) was stirred at 23 °C for 5 min prior to being taken up in 9.5 mL of H<sub>2</sub>O

(0.1% TFA), and purified by preparative reverse-phase HPLC to give *cyclo*-(-ImImPyPy-(R)<sup>α-H<sub>2</sub>N</sup>γ-ImImPyPy-(R)<sup>α-H<sub>2</sub>N</sup>γ-) **1** as a fluffy white solid in 90% yield (0.360 μmol). *Cyclo*-(-ImImPyPy-(R)<sup>α-H<sub>2</sub>N</sup>γ-ImImPyPy-(R)<sup>α-H<sub>2</sub>N</sup>γ-) **1** MALDI-TOF MS (*m/z*): calc'd for C<sub>52</sub>H<sub>61</sub>N<sub>24</sub>O<sub>10</sub> [M+H]<sup>+</sup> 1181.50, found 1181.55; calc'd for C<sub>52</sub>H<sub>60</sub>N<sub>24</sub>O<sub>10</sub>Na [M+Na]<sup>+</sup> 1203.48, found 1203.37; calc'd for C<sub>52</sub>H<sub>60</sub>N<sub>24</sub>O<sub>10</sub>K [M+K]<sup>+</sup> 1219.46, found 1219.33.

#### 5.4.5 Oligonucleotide Purification and Crystallization

Oligonucleotides were purchased HPLC purified from Trilink Biotechnologies (San Diego, CA). Prior to use, oligonucleotides were de-salted using a Waters Sep-Pak cartridge (5g, C-18 sorbent). The Sep-Pak was pre-washed with acetonitrile (25 mL, 3x) followed by MilliQ water (25 mL, 3x). The oligonucleotide was dissolved in 5 mL of 2.0 M NaCl and loaded directly onto the sorbent followed by a wash with 5 mL of 2.0 M NaCl and 250 mL of MilliQ water. The oligonucleotide was eluted with acetonitrile:water (1:1) and lyophilized to dryness. Single strand DNA was quantitated by UV-Vis spectroscopy and incubated with a 2:1 ratio of DNA to polyamide prior to crystallization. Crystals were obtained after 2-8 weeks from a solution of 0.5 mM duplex DNA, 0.65 mM polyamide, 21% 2-methyl-2,4-pentanediol (MPD), 35 mM calcium acetate, 10 mM Tris pH 7.5 equilibrated in sitting drops against a reservoir of 35% MPD at 4 °C. Crystals were collected in Hampton nylon CryoLoops (10 micron, 0.1 mm) and flash cooled to 100 K prior to data collection (Figure 5.8).

#### 5.4.6 UV-visible analysis

DNA/polyamide complex formation was verified prior to structure solution by UV-Visible spectroscopy. Crystals were collected in Hampton nylon CryoLoops (10 micron, 0.1 mm) and washed with crystallization buffer 3 times prior to dissolution in 50 μL of MilliQ water. UV-Visible spectroscopy of the dissolved crystals confirmed the presence of polyamide and DNA duplex in a 1:1 stoichiometry (Figure 5.8).

#### 5.4.7 Data collection, Structure Determination, and Refinement

Polyamide-DNA crystals grew in space group P1 with unit cell dimensions  $a = 22.50$ ,  $b = 25.14$ ,  $c = 29.09$ ,  $\alpha = 66.53$ ,  $\beta = 79.28$ ,  $\gamma = 79.57$ , and one polyamide-duplex DNA complex in the asymmetric unit. This data set was collected at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 12-2 with a MAR Research imaging plate detector at wavelength 0.97 Å. DNA only crystals grew in space group C2 (C 1 2 1) with unit cell dimensions  $a = 31.827$ ,  $b = 25.636$ ,  $c = 34.173$ ,  $\alpha = 90.00$ ,

$\beta = 116.72$ ,  $\gamma = 90.00$  and one DNA strand in the asymmetric unit. This data set was collected at Stanford Synchrotron Radiation Laboratory (SSRL) beamline 11-1 with a MAR Research imaging plate detector at wavelength 0.999 Å (Table 5.1).

Data was processed with MOSFLM<sup>30</sup> and SCALA<sup>31</sup> from the CCP4 suite of programs.<sup>31</sup> Both crystals were solved by direct methods using the SHELX suite of programs (SHELXD).<sup>32,33</sup> Model building and structure refinement was done with Coot<sup>34</sup> and REFMAC5.<sup>35</sup> The final polyamide-DNA complex was refined to an R factor of 9.8 % and an  $R_{\text{free}}$  of 13.6 %. The final DNA structure was refined to an R factor of 10.9 % and an  $R_{\text{free}}$  of 14.3 %. Anisotropic B factors were refined in the final stages and riding hydrogens included (Table 5.1).

#### 5.4.8 Structure Analysis and Figure Preparation

DNA helical parameters were calculated using the program Curves and 3DNA.<sup>36,37</sup> Molecular electrostatic potential maps were calculated at the HF/3-21g\* level using the Gamess program (Figure 5.13).<sup>38-40</sup> Distance measurements and least squares fitting procedures for ring-centroid measurements were performed using UCSF Chimera<sup>41</sup> and Mercury.<sup>42</sup> Structural figures were prepared using UCSF Chimera.<sup>41</sup>

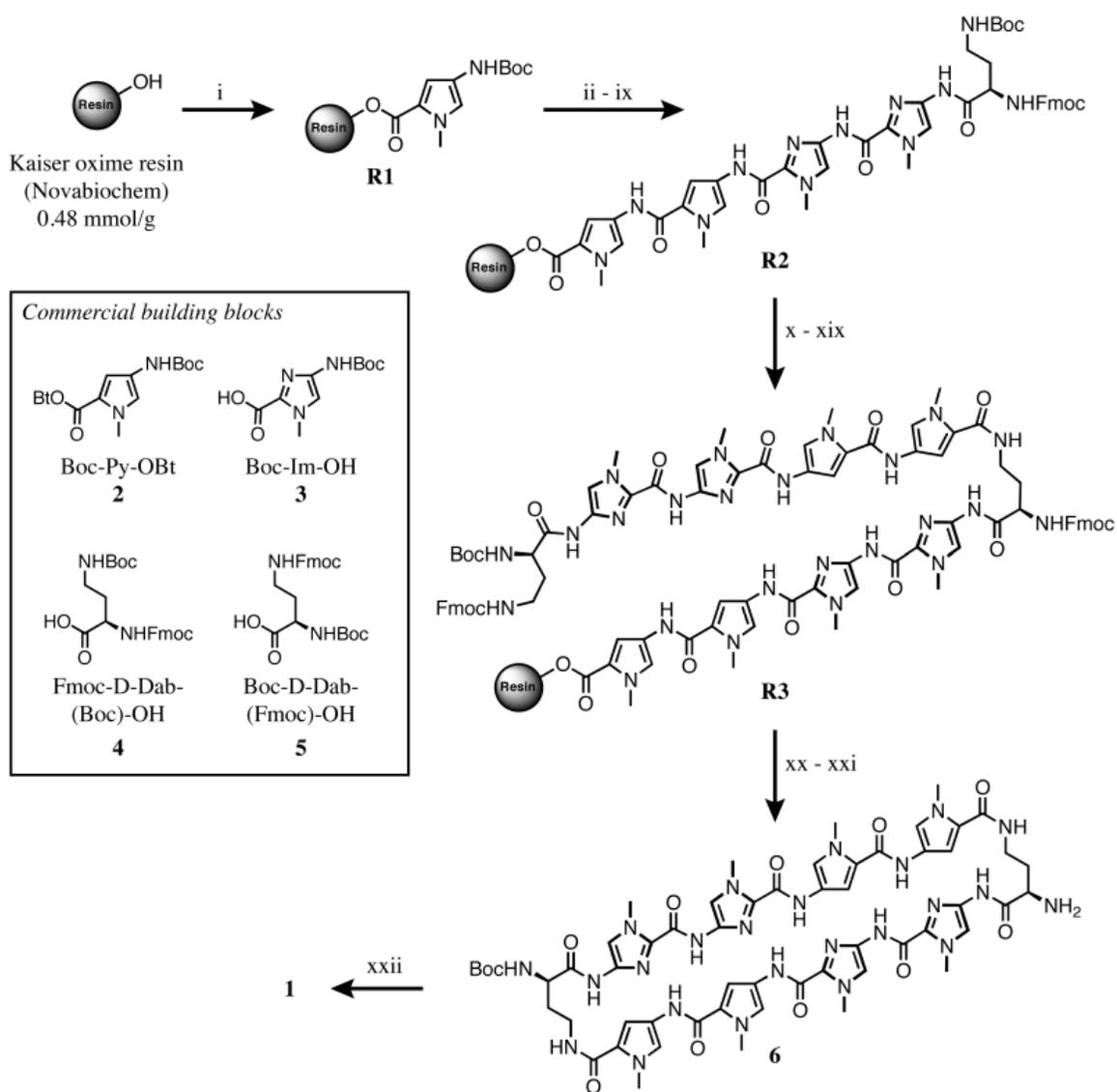
### 5.5 Notes and References

1. Dervan, P. B. Molecular recognition of DNA by small molecules. *Bioorg Med Chem* **9**, 2215-2235 (2001).
2. Dervan, P. B. & Edelson, B. S. Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Current Opinion in Structural Biology* **13**, 284-299 (2003).
3. Trauger, J. W., Baird, E. E. & Dervan, P. B. Recognition of DNA by designed ligands at subnanomolar concentrations. *Nature* **382**, 559-561 (1996).
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* **391**, 468-470 (1998).
5. Kielkopf, C. L., Baird, E. E., Dervan, P. B. & Rees, D. C. Structural basis for G•C recognition in the DNA minor groove. *Nat Struct Biol* **5**, 104-109 (1998).
6. Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., *et al.* A structural basis for recognition of A•T and T•A base pairs in the minor groove of B-DNA. *Science* **282**, 111-115 (1998).
7. Hsu, C. F., Phillips, J. W., Trauger, J. W., Farkas, M. E., *et al.* Completion of a Programmable DNA-Binding Small Molecule Library. *Tetrahedron* **63**, 6146-6151 (2007).
8. Edelson, B. S., Best, T. P., Olenyuk, B., Nickols, N. G., *et al.* Influence of structural variation on nuclear localization of DNA-binding polyamide-fluorophore conjugates. *Nucleic Acids Res* **32**, 2802-2818 (2004).

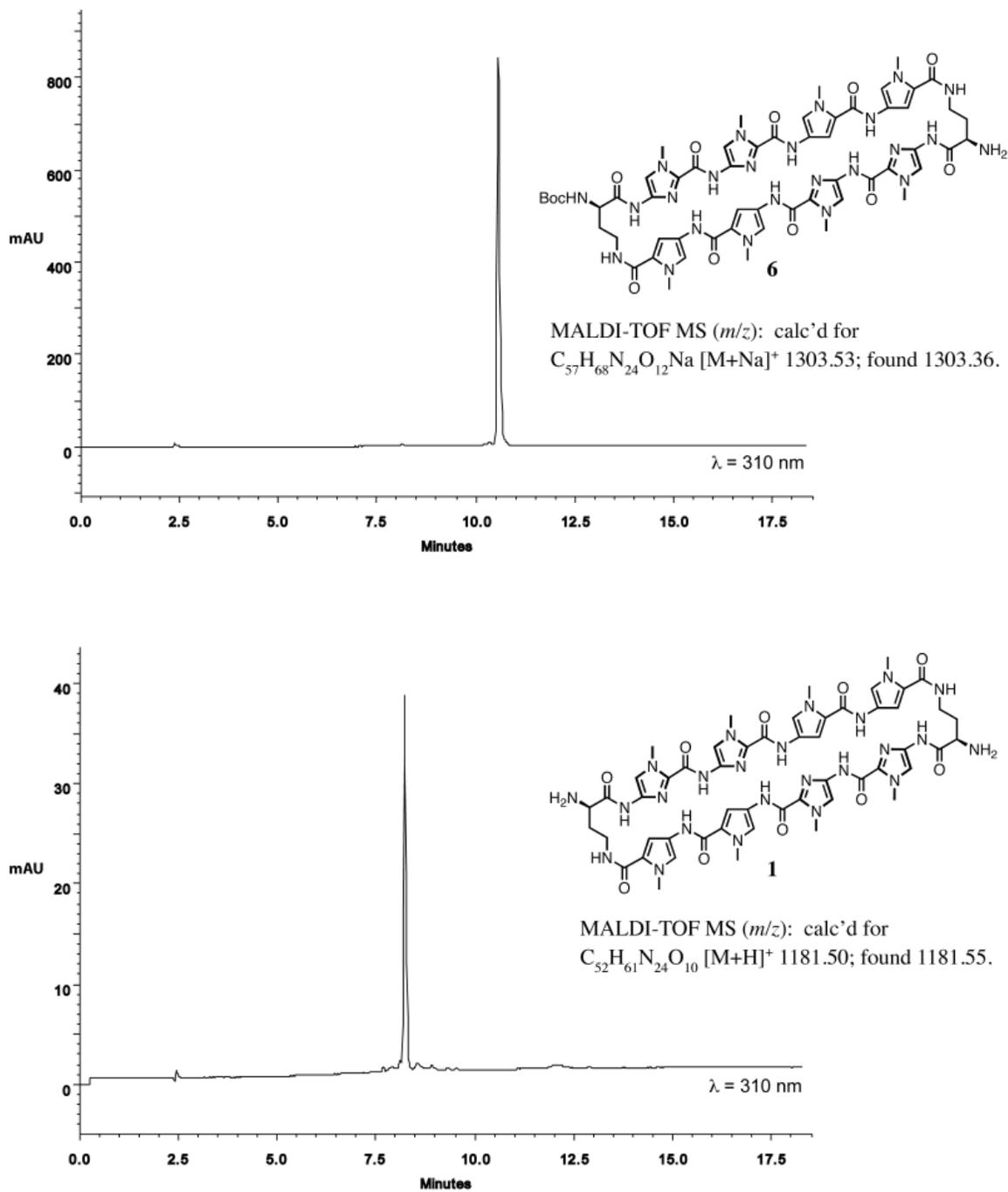
9. Nickols, N. G., Jacobs, C. S., Farkas, M. E. & Dervan, P. B. Improved nuclear localization of DNA-binding polyamides. *Nucleic Acids Res* **35**, 363-370 (2007).
10. Hsu, C. F. & Dervan, P. B. Quantitating the concentration of Py-Im polyamide-fluorescein conjugates in live cells. *Bioorg Med Chem Lett* **18**, 5851-5855 (2008).
11. Gottesfeld, J. M., Melander, C., Suto, R. K., Raviol, H., Luger, K., and Dervan, P. B. Sequence-specific recognition of DNA in the nucleosome by pyrrole-imidazole polyamides. *J. Mol. Biol.* **309**, 615-629 (2001).
12. Suto, R. K., Edayathumangalam, R. S., White, C. L., Melander, C., Gottesfeld, J. M., Dervan, P. B., and Luger, K. Crystal structures of nucleosome core particles in complex with minor groove DNA-binding ligands. *J. Mol. Biol.* **326**, 371-380 (2003).
13. Olenyuk, B. Z., Zhang, G. J., Klco, J. M., Nickols, N. G., *et al.* Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A* **101**, 16768-16773 (2004).
14. Kageyama, Y., Sugiyama, H., Ayame, H., Iwai, A., *et al.* Suppression of VEGF transcription in renal cell carcinoma cells by pyrrole-imidazole hairpin polyamides targeting the hypoxia responsive element. *Acta Oncol* **45**, 317-324 (2006).
15. Nickols, N. G., Jacobs, C. S., Farkas, M. E. & Dervan, P. B. Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem Biol* **2**, 561-571 (2007).
16. Nickols, N. G. & Dervan, P. B. Suppression of androgen receptor-mediated gene expression by a sequence-specific DNA-binding polyamide. *Proc Natl Acad Sci USA* **104**, 10418-10423 (2007).
17. Matsuda, H., Fukuda, N., Ueno, T., Tahira, Y., *et al.* *Journal of the American Society of Nephrology* **17**, 422-432 (2006).
18. Yao, E. H., Fukuda, N., Ueno, T., Matsuda, H., *et al.* Novel gene silencer pyrrole-imidazole polyamide targeting lectin-like oxidized low-density lipoprotein receptor-1 attenuates restenosis of the artery after injury. *Hypertension* **52**, 86-92 (2008).
19. Nguyen-Hackley, D. H., Ramm, E., Taylor, C. M., Joung, J. K., *et al.* Allosteric inhibition of zinc-finger binding in the major groove of DNA by minor-groove binding ligands. *Biochemistry* **43**, 3880-3890 (2004).
20. Heinemann, U., and Alings, C. Crystallographic study of one turn of G/C-rich B-DNA. *J. Mol. Biol.* **210**, 369-381 (1989).
21. Egli, M., and Sarkhel, S. Lone Pair-Aromatic Interactions: To Stabilize or Not to Stabilize *Acc. Chem. Res.* **40**, 197-205 (2007).
22. Gallivan, J. P., and Dougherty, D. A. Can Lone Pairs Bind to a  $\pi$  System? The Water-Hexafluorobenzene Interaction *Org. Lett.* **1**, 103-106 (1999).
23. Mecozzi, S., West, A. P., and Dougherty, D. A. *Proc. Natl. Acad. Sci. USA.* **93**, 10566-10571 (1996).
24. Hogan, M., Dattagupta, N., and Crothers, D. M. Transmission of allosteric effects in DNA *Nature* **278**, 521-524 (1979).
25. Panne, D., Maniatis, T. & Harrison, S. C. An atomic model of the interferon-beta enhanceosome. *Cell* **129**, 1111-1123 (2007).

26. Lavelle, C. DNA torsional stress propagates through chromatin fiber and participates in transcriptional regulation. *Nat. Struct. Mol. Biol.* **15**, 146-154 (2008).
27. Moretti, R., Donato, L. J., Brezinski, M. L., Stafford, R. L., *et al.* Targeted chemical wedges reveal the role of allosteric DNA modulation in protein-DNA assembly. *ACS Chem Biol* **3**, 220-229 (2008).
28. Baird, E. E., Dervan, P. B. Solid phase synthesis of polyamides containing imidazole and pyrrole amino acids. *J. Am. Chem. Soc.* **118**, 6141-6146 (1996).
29. Belitsky, J. M., Nguyen, D. H., Wurtz, N. R., Dervan, P. B. Solid-phase synthesis of DNA binding polyamides on oxime resin. *Bioorg. Med. Chem.* **10**, 2767-2774 (2002).
30. Leslie, A. G. W. Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter on Protein Crystallography* **26**, (1992).
31. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763 (1994).
32. Sheldrick, G.M. A short history of SHELX. *Acta Crystallogr. A* **64**, 112-122 (2008).
33. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D* **58**, 1772-1779 (2002).
34. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* **60**, 2126-2132 (2004).
35. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr* **53**, 240-255 (1997).
36. Lavery, R., Sklenar, H. Curves 5.2: Helical analysis of irregular nucleic acids. *Biochimie Theorique, CNRS URA*, **77** (1997).
37. Lu, X. J., Olson, W. K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids. Res.*, **31**, 5108-5121 (2003).
38. Schmidt, M. W., Baldrige, K. K., Boatz, J. A., Elbert, S. T., Gordon, M. S., Jensen, J. H., *et al.* General atomic and molecular electronic structure system. *Journal of Computational Chemistry*, **14**, 1347-1363 (1993).
39. Binkley, J. S., Pople, J. A., & Hehre, W. J. Self-Consistent molecular orbital methods. 21. Small split-valence basis sets for first-row elements. *Journal of the American Chemical Society*, **102**, 939-947 (1980).
40. Pietro, W. J., Francl, M. M., Hehre, W. J., DeFrees, D. J., Pople, J. A., & Binkley, J. S. Self-Consistent molecular orbital methods. 24. Supplemented small split-valence basis sets for second-row elements. *Journal of the American Chemical Society*, **104** 5039-5048 (1982).
41. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., *et al.* UCSF chimera-a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, **25**, 1605-12 (2004).
42. Macrae, C.F., Edgington, P.R., McCabe, P., Pidcock, E., Shields, G.P., Taylor, R., Towler, M., van de Streek, J. Mercury: visualization and analysis of crystal structures. *J. Appl. Cryst.* **39**, 453-457, (2006).

## 5.6 Spectra, Data Statistics, and Supplemental Information



**Figure 5.6** Solid-phase synthesis of cyclic polyamide **1** using Kaiser oxime resin and the commercially available building blocks presented above. Reagents and conditions: (i) Boc-Py-OBt **2**, DIEA, DMF; (ii) 20% TFA/DCM; (iii) Boc-Py-OBt **2**, DIEA, DMF; (iv) 20% TFA/DCM; (v) Boc-Im-OH **3**, PyBOP, DIEA, DMF; (vi) 50% TFA/DCM; (vii) Boc-Im-OH **3**, PyBOP, DIEA, DMF; (viii) 50% TFA/DCM; (ix) Fmoc-D-Dab(Boc)-OH **4**, PyBOP, DIEA, DMF; (x) 20% TFA/DCM; (xi) Boc-Py-OBt **2**, DIEA, DMF; (xii) 20% TFA/DCM; (xiii) Boc-Py-OBt **2**, DIEA, DMF; (xiv) 20% TFA/DCM; (xv) Boc-Im-OH **3**, PyBOP, DIEA, DMF; (xvi) 50% TFA/DCM; (xvii) Boc-Im-OH **3**, PyBOP, DIEA, DMF; (xviii) 50% TFA/DCM; (xix) Boc-D-Dab(Fmoc)-OH **5**, PyBOP, DIEA, DMF; (xx) piperidine, DMF; (xxi) DMF, 37°C, 24 h; (xxii) 50% TFA/DCM.



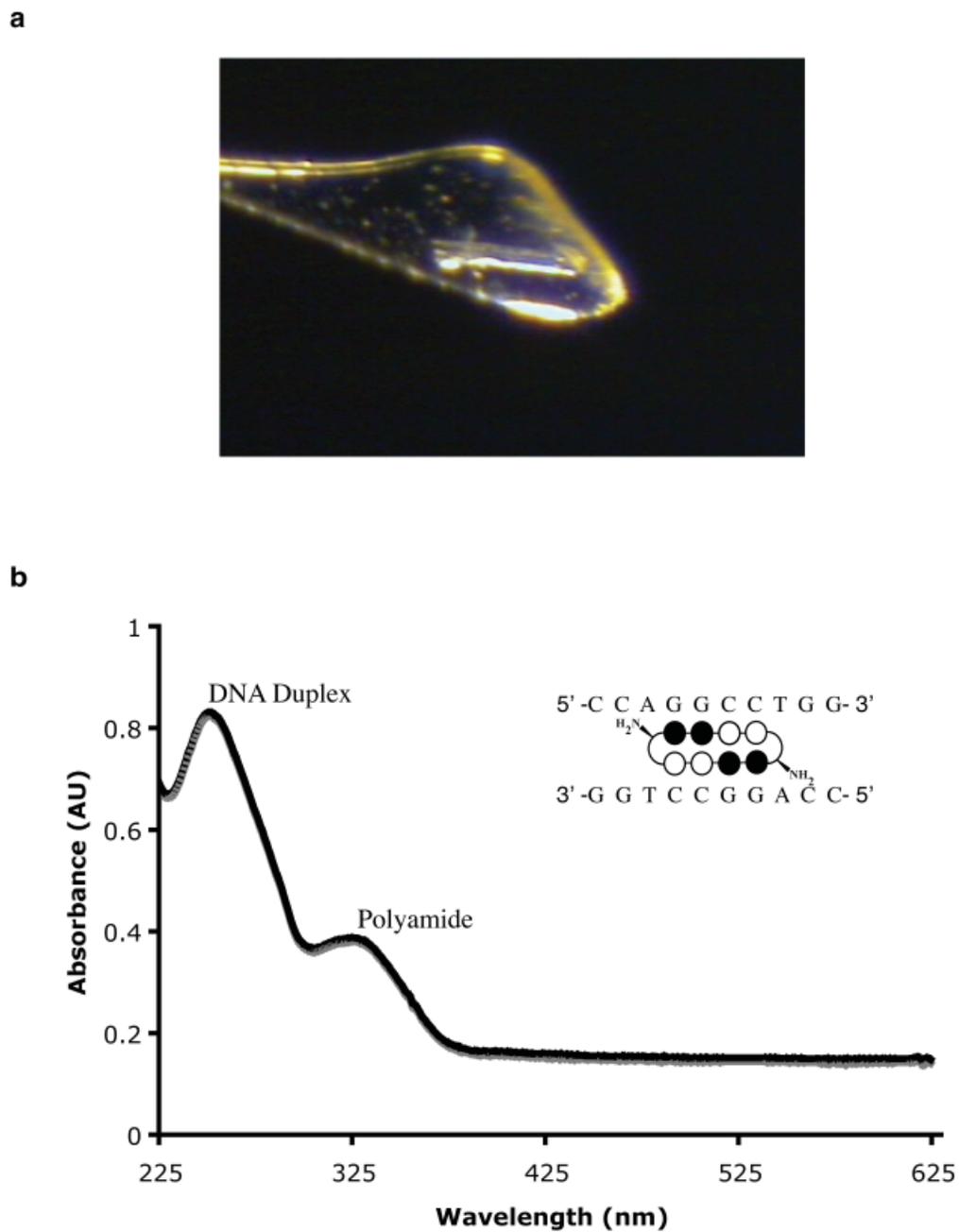
**Figure 5.7** Polyamide analytical data. Analysis of polyamides **6** and **1** by analytical RP-HPLC and MALDI-TOF MS. Wavelength shown is at 310 nm.

**Table 5.1** Data collection and refinement statistics.

Data collection		DNA-Polyamide	DNA
	Space group	P1	C2
	Cell dimensions		
	a, b, and c, Å	22.50	31.83
		25.14	25.64
		29.09	34.17
	$\alpha$ , $\beta$ , and $\gamma$ , °	66.53	90
		79.28	116.72
		79.57	90
	Wavelength	0.970	0.999
	Resolution, Å	26.47 - 1.18	15.91 - 0.98
	$R_{\text{merge}}^*$	12.4 (15.4)	5.0 (7.5)
	$I/\sigma I^*$	12.5 (8.9)	18.7 (8.3)
	Completeness, %*	94.6 (90.7)	89.5
	Redundancy	4.0	2.5
Refinement			
	Resolution, Å	26.47 - 1.18	15.91 - 0.98
	No. of reflections	16,139	12,125
	$R_{\text{work}}/R_{\text{free}}^\ddagger$	9.9 / 13.8	10.8 / 14.4
	No. of atoms		
	DNA	441	297
	Polyamide	86	–
	Calcium	6	3
	Water	262	137
	B factors		
	DNA	5.6	6.5
	Polyamide	5.2	–
	Calcium	11.7	33.2
	Water	19.2	20.1
	R.m.s. deviations		
	Bond lengths, Å	0.026	0.033
	Bond angles, °	2.57	3.44

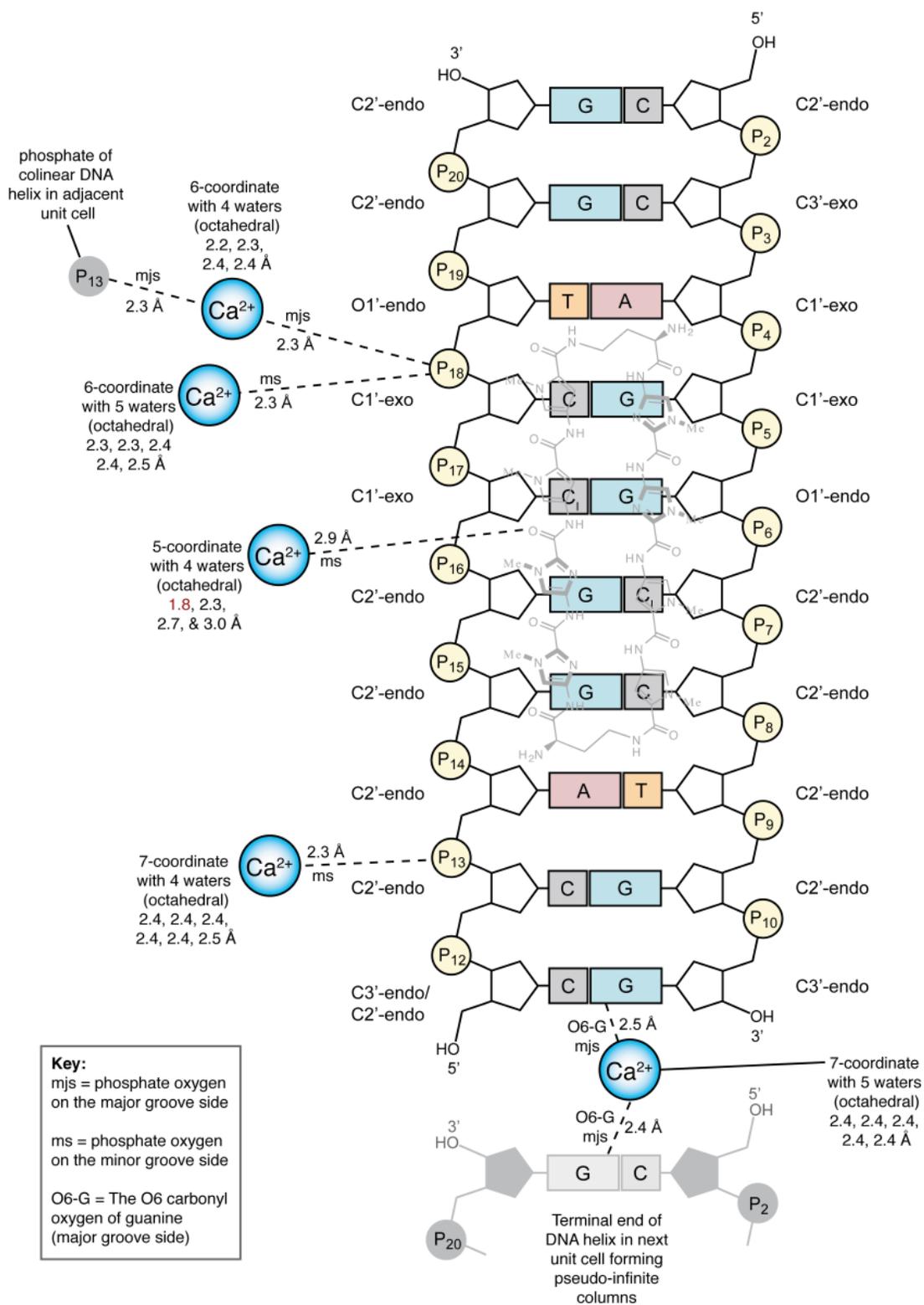
<sup>†</sup>Free  $R$  calculated against 5% of the reflections randomly removed.

\*Highest-resolution shell is shown in parentheses.

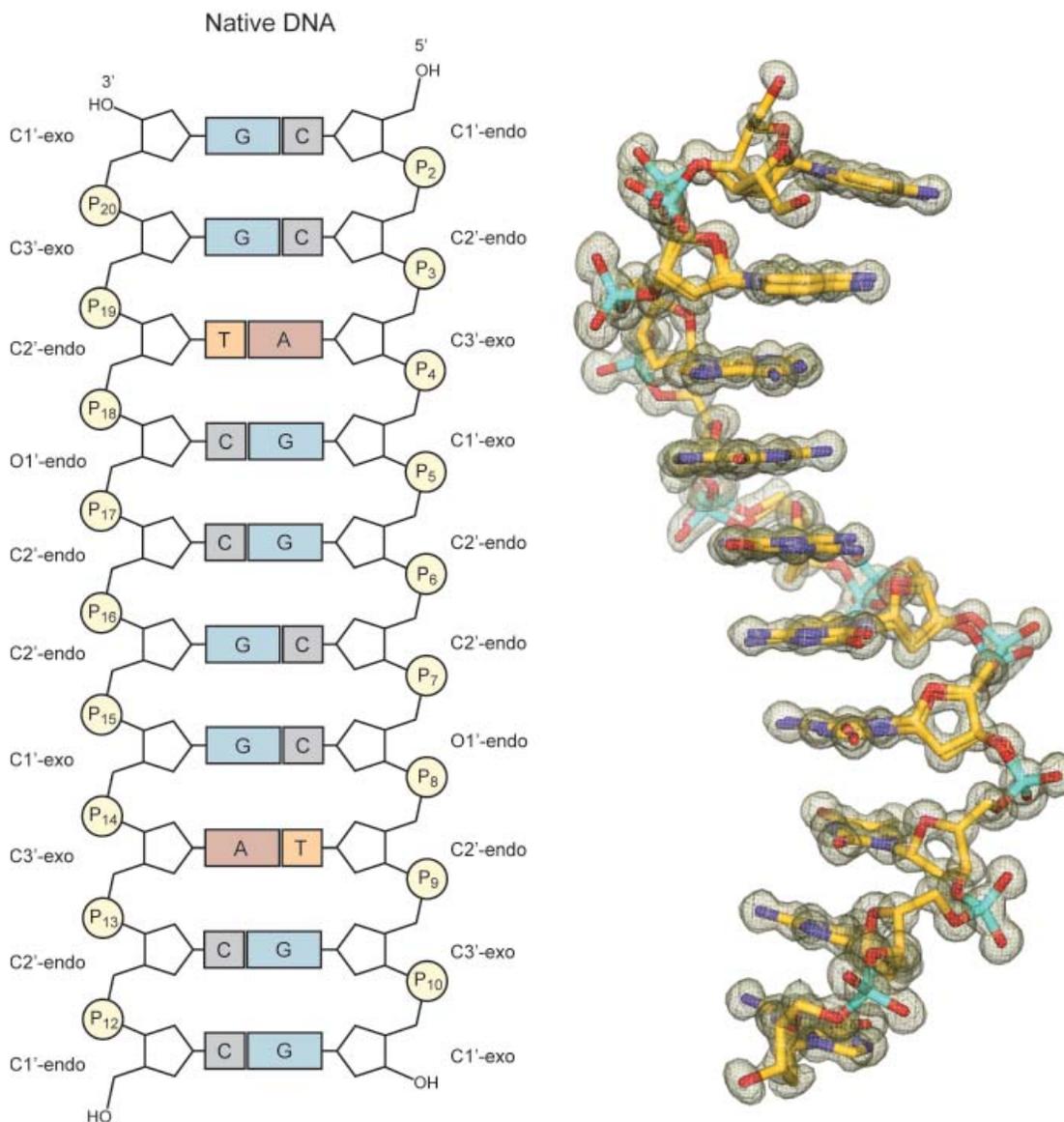


**Figure 5.8** Single crystal of the polyamide-DNA complex. a) Single crystal frozen in a cryoloop for data collection. b) UV-Vis of the single crystal polyamide-DNA complex dissolved in 50  $\mu\text{L}$  of water.

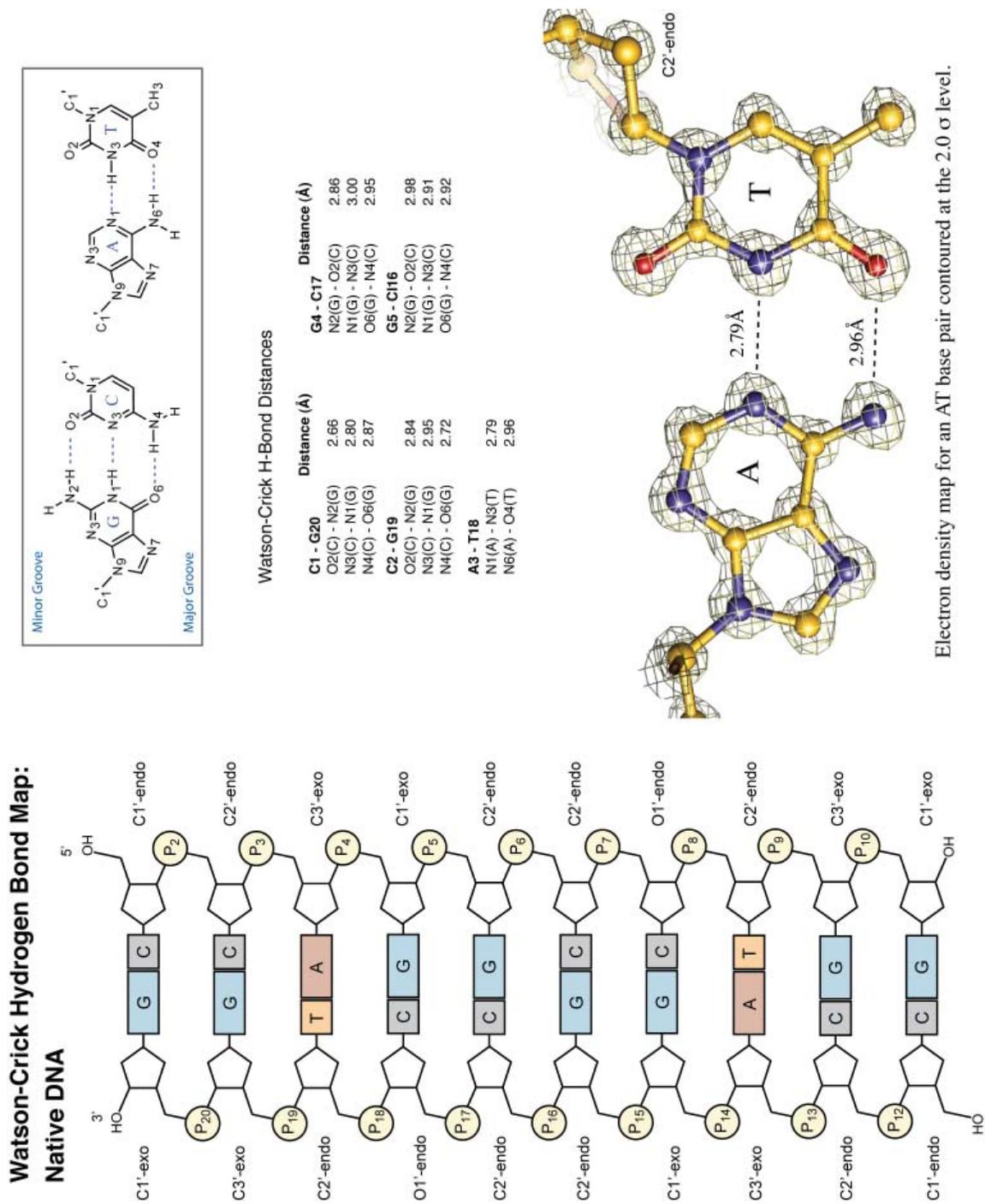




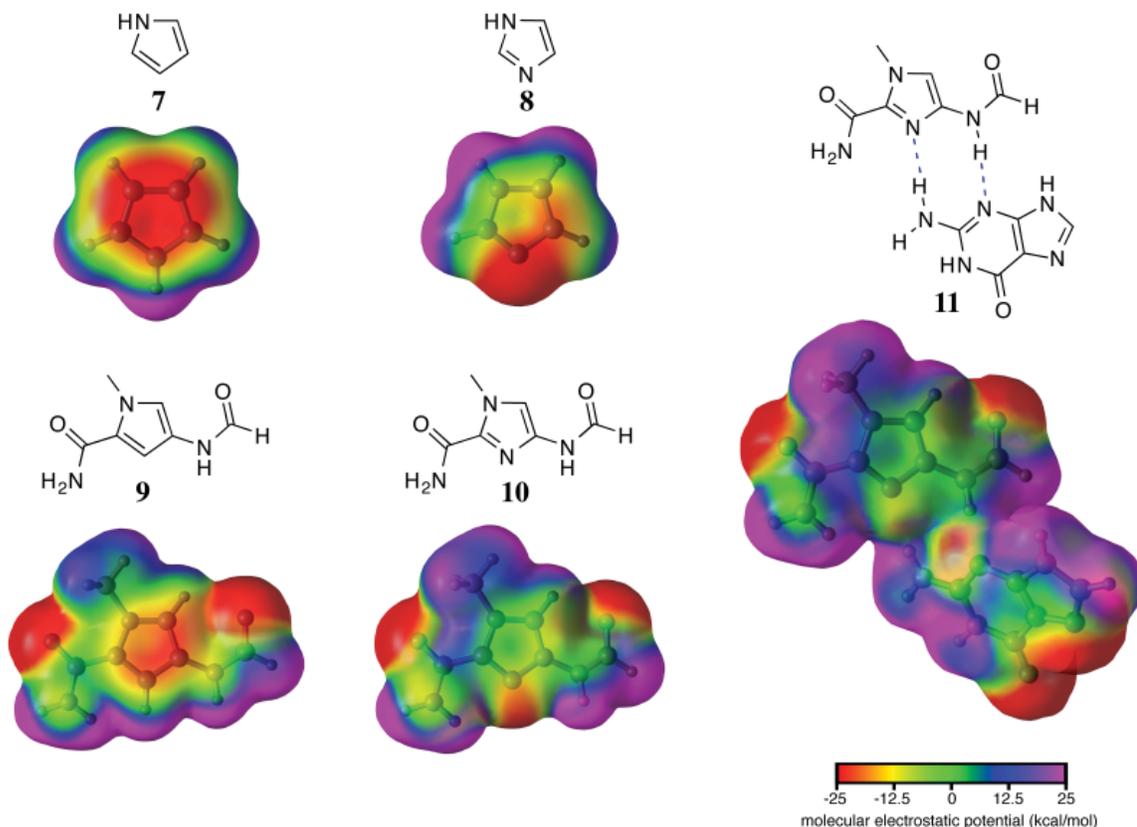
**Figure 5.10** Polyamide/DNA complex map. Schematic diagram of polyamide/DNA crystal structure showing the sugar conformation at each position and calcium ion coordination.



**Figure 5.11** DNA structure map. Schematic diagram of native DNA crystal structure showing the sugar conformation at each position and the electron density map of the DNA asymmetric unit contoured at the 1.0  $\sigma$  level. The structure on the right is oriented 5' to 3' from top to bottom and corresponds to the right hand strand in the structure map.



**Figure 5.12** Hydrogen bond map of DNA crystal structure and electron density of the AT base pair contoured at the 2.0  $\sigma$  level.

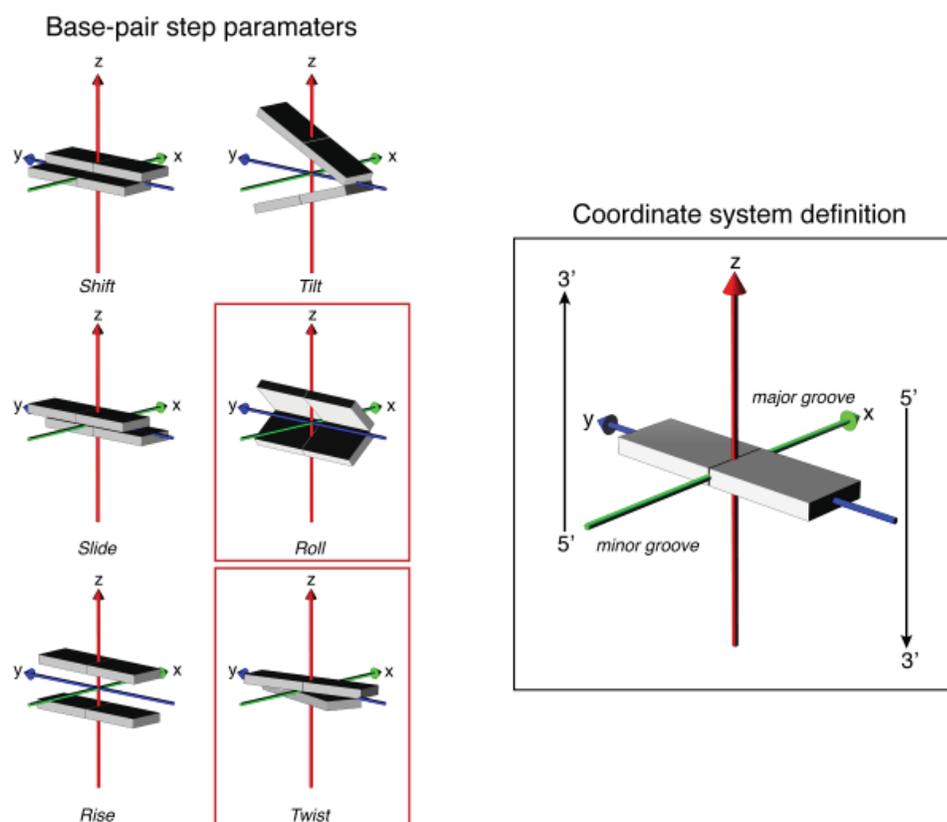


**Figure 5.13** Molecular electrostatic potential maps of compounds **7-10** and complex **11**. Complex **11** coordinates were taken directly from the polyamide/DNA crystal structure and used without geometry optimization. All ab initio calculations reported here were performed using HF/3-231G\* as implemented in the Gamess program.<sup>37-39</sup> Full geometry optimization was performed on all structures except for complex **11** whose coordinates correspond to those of the crystal structure. Electrostatic potential surfaces were generated by mapping the electrostatic potentials onto surfaces of molecular electron density ( $0.002 \text{ electron/\AA}^3$ ) and color-coding, using the Chimera program.<sup>40</sup> The molecular electrostatic potential energy values range from -25 kcal/mol for values of negative potential (red) to +25 kcal/mol for values of positive potential (blue). This range was chosen to emphasize the variations in the aromatic region and some regions of the electrostatic potential associated with heteroatoms may lie beyond the  $\pm 25$  kcal/mol range.

**Local base-pair step parameters**

Parameter*	CC/GG	CA/TG	AG/CT	GG/CC	GC/GC	CC/GG	CT/AG	TG/CA	GG/CC
Shift, Å									
DNA	-0.58	0.21	0.66	-0.34	0.00	0.34	-0.66	-0.21	0.58
PA/DNA	0.00	0.63	-1.96	-1.35	0.13	1.43	1.81	-1.78	0.09
Slide, Å									
DNA	0.69	2.80	0.78	0.24	-0.37	0.24	0.78	2.80	0.69
PA/DNA	-0.47	0.59	0.31	0.94	-0.02	0.84	0.38	1.51	0.29
Rise, Å									
DNA	3.30	3.26	3.08	2.99	4.48	2.99	3.08	3.26	3.30
PA/DNA	3.33	3.08	3.28	3.28	3.04	3.29	3.29	3.08	3.43
Tilt, °									
DNA	1.82	-0.81	-3.57	-5.49	0.00	5.49	3.57	0.81	-1.82
PA/DNA	4.14	1.42	-8.88	-4.81	0.64	5.05	5.38	-2.03	-3.89
Roll, °									
DNA	9.80	-8.73	9.85	-2.96	-4.11	-2.96	9.85	-8.73	9.80
PA/DNA	11.79	7.20	13.46	1.07	-0.20	2.57	16.09	6.54	11.05
Twist, °									
DNA	27.24	50.86	21.53	33.56	52.38	33.56	21.53	50.86	27.24
PA/DNA	36.26	29.92	33.52	38.71	33.84	38.32	31.32	37.22	38.20

\*Relationship between the bases composing the base pair.

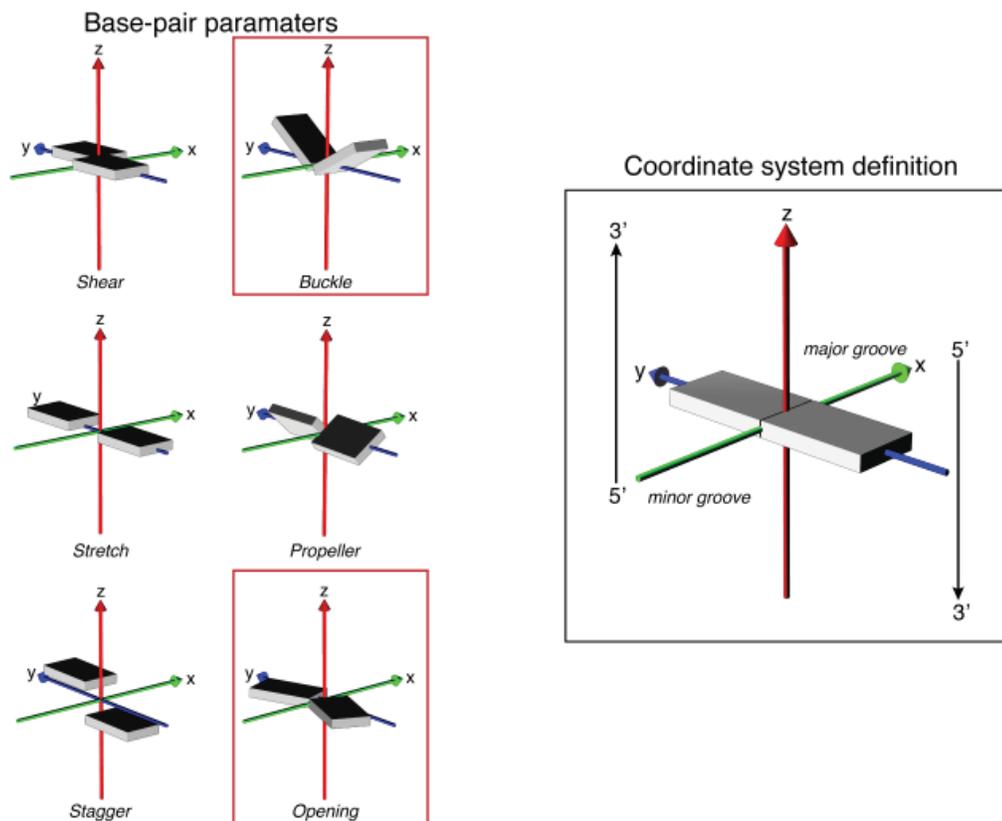


**Figure 5.14** Comparison of Local base-pair step parameters for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.

## Local base-pair parameters and sugar pucker

Parameter*	C•G	C•G	A•T	G•C	G•C	C•G	C•G	T•A	G•C	G•C
Shear, Å										
DNA	0.04	0.16	0.02	-0.33	-0.32	0.32	0.33	-0.02	-0.16	-0.04
PA/DNA	0.15	0.10	0.16	-0.07	-0.17	0.11	0.10	-0.04	-0.37	0.05
Stretch, Å										
DNA	-0.21	-0.20	-0.13	-0.19	0.02	0.02	-0.19	-0.13	-0.20	-0.21
PA/DNA	-0.14	-0.12	-0.07	-0.10	-0.15	-0.13	-0.11	-0.13	-0.26	-0.09
Stagger, Å										
DNA	0.07	0.08	0.13	0.31	0.86	0.86	0.31	0.13	0.08	0.07
PA/DNA	0.16	0.09	-0.10	0.08	0.07	-0.02	-0.09	-0.21	0.03	0.17
Buckle, °										
DNA	-5.52	-2.64	0.23	13.52	22.43	-22.43	-13.52	-0.23	2.64	5.52
PA/DNA	-5.30	-9.09	5.62	-1.10	-2.90	1.60	1.81	0.42	11.29	5.28
Propeller, °										
DNA	-16.32	-9.30	-4.01	-9.51	-12.64	-12.64	-9.51	-4.01	-9.30	-16.32
PA/DNA	-21.09	-7.15	-12.99	-6.10	-16.87	-15.12	-1.60	-5.10	-9.33	-24.06
Opening, °										
DNA	2.09	-2.60	1.17	0.49	4.88	4.88	0.49	1.17	-2.60	2.09
PA/DNA	-0.69	-0.93	4.73	-1.94	-4.39	-4.48	-0.99	6.11	0.60	2.78
Sugar pucker										
DNA	C1'-exo C1'-exo	C2'-endo C3'-exo	C2'-endo C2'-endo	C1'-exo C1'-exo	C2'-endo C1'-exo	C1'-exo C2'-endo	C1'-exo C1'-exo	C2'-endo C2'-endo	C3'-exo C2'-endo	C1'-exo C1'-exo
PA/DNA	C3'-endo C2'-endo	C2'-endo C2'-endo	C2'-endo O4'-endo	C2'-endo C1'-exo	C2'-endo C1'-exo	C1'-exo C2'-endo	C1'-exo C2'-endo	C1'-exo C2'-endo	C3'-exo C2'-endo	C2'-endo C3'-endo

\*Relationship between the bases composing the base pair.

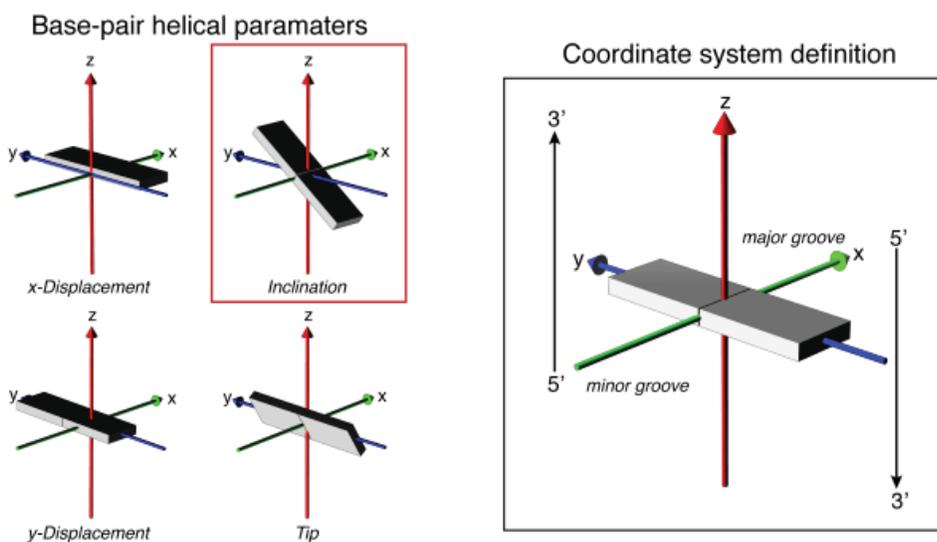


**Figure 5.15** Comparison of Local base-pair parameters and sugar conformations for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.

**Local base-pair helical parameters**

Parameter*	CC/GG	CA/TG	AG/CT	GG/CC	GC/GC	CC/GG	CT/AG	TG/CA	GG/CC
X-displacement, Å									
DNA	-0.87	3.77	-1.26	0.83	-0.05	0.83	-1.26	3.77	-0.87
PA/DNA	-2.19	-0.26	-1.50	1.28	-0.01	0.94	-1.93	1.49	-0.93
Y-displacement, Å									
DNA	1.57	-0.30	-2.74	-0.22	0.00	0.22	2.74	0.30	-1.57
PA/DNA	0.52	-0.91	1.77	1.42	-0.12	-1.51	-2.12	2.48	-0.60
Inclination, °									
DNA	19.97	-10.08	24.57	-5.07	-4.65	-5.07	24.57	-10.08	19.97
PA/DNA	18.29	13.69	21.84	1.61	-0.34	3.89	27.43	10.14	16.42
Tip, °									
DNA	-3.71	0.94	8.91	9.41	0.00	-9.41	-8.91	-0.94	3.71
PA/DNA	-6.42	-2.70	14.41	7.22	-1.09	-7.64	-9.18	3.15	5.78

\*Relationship between the bases composing the base pair.



**Figure 5.16** Comparison of Local base-pair helical parameters for DNA with and without polyamide complexed. Parameters were calculated using 3DNA.