Chapter 4: Oligomerization Route to Polyamide Macrocycles

The text of this chapter was taken in part from a manuscript coauthored with Daniel A. Harki and Peter B. Dervan (Caltech)*

(Chenoweth, D.M., Harki, D.A., Dervan, P. B. *In Preparation*.)

Abstract

Cyclic pyrrole–imidazole polyamides are sequence-specific DNA-binding small molecules that are cell permeable and can regulate endogenous gene expression. Syntheses of cyclic polyamides have been achieved by solid-phase and more recently, solution-phase methods. We report a rapid solution-phase oligomerization approach to cyclic polyamides that yields 8, 12, and 16 membered macrocycles. A remarkable preference for DNA binding by the 8 and 16 membered oligomers was observed over the 12-ring macrocycle, which we attributed to a conformational constraint not present in the smaller and larger systems.

4.1 Introduction

Pyrrole–imidazole polyamides are a class of cell-permeable small molecules that bind to the minor groove of DNA in a sequence specific manner.[1,2] Antiparallel arrangements of *N*-methylpyrrole (Py) and *N*-methylimidazole (Im) carboxamides (Im/Py) recognize G•C from C•G base pairs, whereas Py/Py specifies for both T•A and A•T.[3] Hairpin Py-Im polyamides have been programmed for a broad repertoire of DNA sequences with high affinities.[4] These cell permeable[5] ligands can influence gene transcription by disrupting protein-DNA interfaces,[2] and have been shown to control transcription of genes important in human disease.[6] Py-Im polyamides have also been used for a variety of applications ranging from DNA detection[7] and transcriptional activation[8] to the self-assembly of higher order structures.[9] Conjugation of polyamides to functional domains have yielded artificial transcription factor mimics,[8] fluorescent sequence-specific DNA probes,[7] and DNA nano-architectures.[9]

We recently reported solution-phase methods for the synthesis of hairpin[10] and cyclic polyamides.[11] Key to the cyclic polyamide synthesis was a highly efficient macrocyclization that yielded polyamide **1z** (Figure 4.1) from an acyclic precursor. Activation of the C-terminal Py amino acid of **1z** as a pentafluorophenyl ester allowed efficient macrocyclization by the γ-NH_2 on the turn moiety under dilute reaction conditions. Our studies of polyamide **1**, the deprotected analogue of **1z**, revealed it possessed extremely high DNA binding affinities, was cell permeable, and could disrupt androgen receptor-activated gene expression in cell culture. Additionally, preliminary studies of the in vitro ADMET properties of **1** revealed excellent metabolic stability.[11,12]

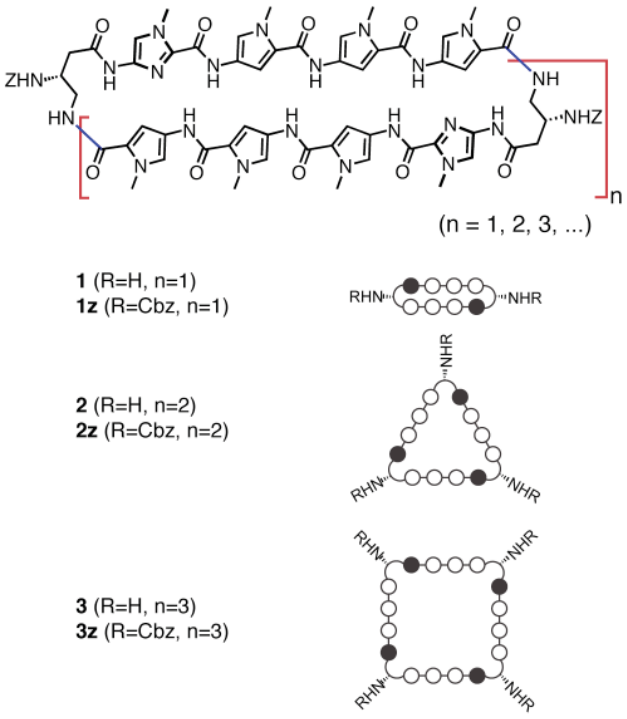


Figure 4.1 Structures of macrocyclic polyamides **1z**–**3z** and **1**–**3**, and their ball-and-stick models. Polyamide shorthand code: closed circles, Im; open circles, Py monomers. (Z = benzyl carbamate protecting group)

An orthogonal polymerization/oligomerization strategy for the synthesis of **1** and related polyamides is reported here. This method delivers symmetrical Py-

Im polyamide macrocycles from simple Py-Im building blocks in a rapid and convergent manner (Figure 4.2). Separable higher-order oligomers, such as the 12-membered (**2**) and 16-membered (**3**) cyclic polyamides are also produced by this method. In addition to describing the synthetic chemistry to prepare **1**–**3**, we report for the first time the DNA binding properties of such expanded polyamide macrocycles.

4.2 Results and Discussion

Our strategy for this oligomerization route relied on the palindromic nature of polyamide **1**. Disconnection of **1** at both γ-amino turns affords two identical halves of the molecule. Bimolecular coupling between two molecules, followed by intramolecular ring closure delivers cyclic Py-Im polyamides. Bifunctional monomer **4** contains every atom needed to construct cyclic polyamides **1**–**3** by this process (Figure 4.2).

The pentafluorophenyl ester **4** was prepared in one step from the previously reported carboxylic acid of **4**.[11] Acidic deprotection of the γ-amino functionality of **4** followed by drying *in vacuo* yields intermediate **5** which is the substrate for the homodimerization/ oligomerization reaction. To initiate this sequence, the protected trifluoroacetate salt **5** was diluted with DMSO, then treated with an organic base (DIEA) to unmask the highly nucleophilic primary γ-amine. The ensuing oligomerization/macrocyclization process provides benzylcarbamate protected cyclic polyamides **1z**, **2z**, **3z**, and trace amounts of unisolated higher-order oligomers. A distribution of uncyclized intermediates corresponding to the dimer (8-ring cycle, **1z**), trimer (12-ring cycle, **2z**), tetramer (16-ring cycle, **3z**), and higher order adducts can be observed at early time points, as evidenced by HPLC analysis at 2 hr (Figure 4.3). Extended

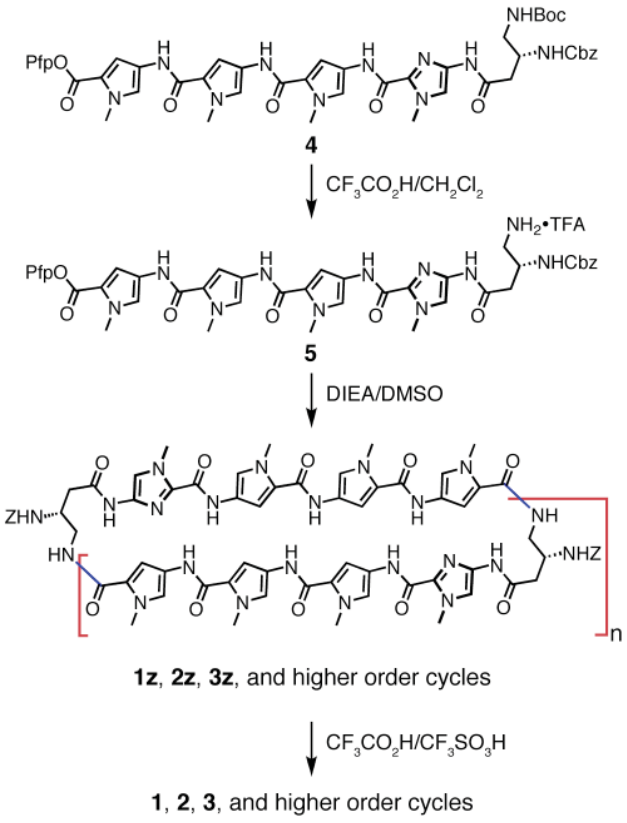


Figure 4.2 Synthesis of macrocyclic polyamides **1**–**3** and higher order cycles by oligomerization of bi-functional intermediate **5**.

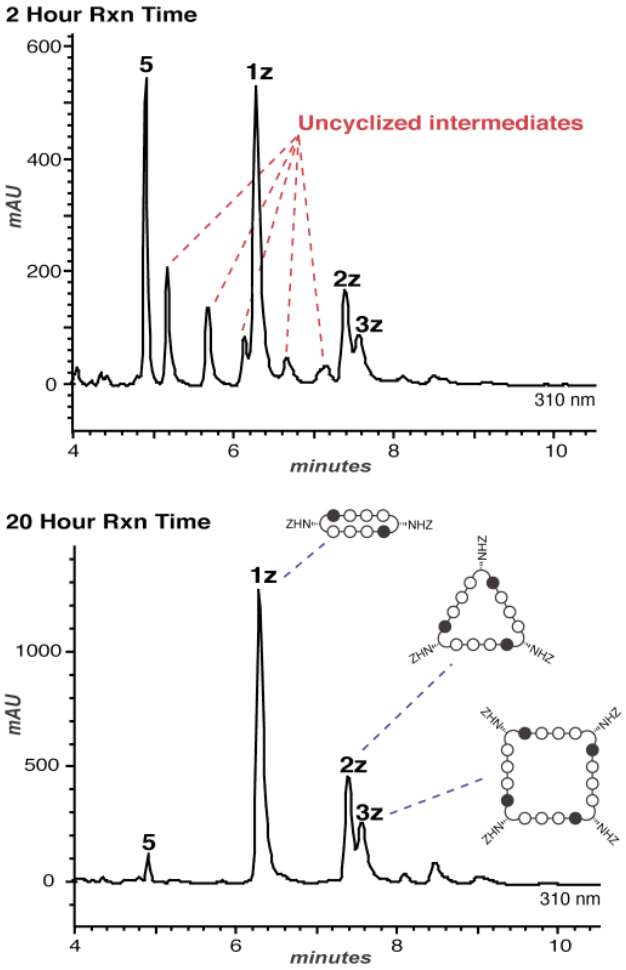


Figure 4.3 (Top) Reverse phase HPLC analysis (2 hr time point) of the oligomerization reaction showing **1z**, **2z**, **3z**, and acyclic intermediates. (Bottom) Analysis (20 h time point) of the oligomerization reaction revealing **1z**, **2z**, **3z**, and consumption of uncyclized intermediates. Peaks were identified by high-resolution mass spectrometry following separation and Cbz-deprotection.

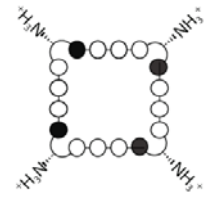
reaction times (20 hours) reveals cyclized polyamides **1z**, **2z**, and **3z** in a ratio of 6.6:2.6:1 almost exclusively (Figure 4.3). Isolation of **1z** (13.9%), **2z** (5.5%), and **3z** (2.1%) by preprative HPLC, followed by Cbz-deprotection under acidic conditions (CF_3CO_2H/CF_3SO_3H) provides polyamide macrocycles **1**–**3**.

With polyamide macrocycles **2** and **3** in hand we evaluated their ability to bind duplex DNA relative to previously reported cycle **1**.[11] Polyamide **1** increases the dsDNA melting temperature by 23.6 ºC. To our surprise, trimeric macrocycle **2** failed to bind its target double stranded DNA sequence as evidenced by the complete lack of ligand-promoted thermal stabilization of duplex DNA melting (Table 4.1). This result is presumably due to inherent geometrical constraints of **2**, preventing the side-by-side antiparallel alignment of the PyPyPyIm strands, a motif well accomodated by the DNA minor groove. In the case of tetrameric macrocycle **3**, dsDNA binding and thermal stabilization was completely restored to a comparable value to dimer **1**. We hypothesize that an even number of PyPyPyIm strands allows **3** to possess a collapsed or folded tetramer geometry, with two adjacent, antiparallel PyPyPyIm strands followed by an identical repeat of this motif linked through two intervening turn units.

4.3 Conclusion

In summary, we have demonstrated that macrocyclic polyamides can by synthesized by

Table 4.1 T_m values for cycles **1**–**3** in the presence of DNA.[a]

dsDNA sequence =	5′-TTGC **TGTTCT** GCAA-3′ 3′-AACG **ACAAGA** CGTT-5′	
Polyamide cycle	T_m / °C	ΔT_m / °C
—	60.0 (±0.3)	—
 (**1**)	83.5 (±0.5)	23.6 (±0.6)
 (**2**)	60.1 (±0.6)	0.1 (±0.6)
 (**3**)	83.0 (±0.3)	23.1 (±0.4)

[a]All values reported are derived from at least three melting temperature experiments with standard deviations indicated in parentheses. Δ*T*m values are given as $Tm^{(DNA/polyamide)} - Tm^{(DNA)}$. The propagated error in Δ*T*m measurements is the square root of the sum of the square of the standard deviations for the *T*m values.

oligomerization of a bifunctional polyamide to yield a distribution of cyclic polyamide oligomers. Additionally, we show that certain cyclic polyamide geometries are completely devoid of the ability to bind dsDNA, a result which could be utilized in the design of highly specific molecules for targeting non-B-form DNA structures or other higher-order nucleic acid motifs.

4.4 Experimental Section

4.4.1 General

Chemicals were purchased from Sigma-Aldrich and were used without further purification. (*R*)-3,4-Cbz-Dbu(Boc)-OH was purchased from Senn Chemicals AG (code number 44159). Bulk grade solvents were from Fisher Scientific. Analytical HPLC analysis was conducted on a Beckman Gold instrument equipped with a Phenomenex Gemini analytical column (250 × 4.6 mm, 5 μm), a diode array detector, and the mobile phase consisted of a gradient of acetonitrile (MeCN) in 0.1% (v/v) aqueous CF_3CO_2H. Preparative HPLC was performed on an Agilent 1200 system equipped with a solvent degasser, diode array detector, and a Phenomenex Gemini column (250 × 21.2 mm, 5 μm). A gradient of MeCN in 0.1% (v/v) aqueous CF_3CO_2H was utilized as the mobile phase. NMR spectroscopy was performed on a Varian instrument operating at 499.8 MHz (for ^1H) at ambient temperature. All NMR analyses were performed in DMSO-d_6, and chemical shifts are reported in parts per million relative to the internal solvent peak referenced to 2.49 (for ^1H). High-resolution mass spectrometry (HRMS) was recorded in positive-ion mode by fast-atom bombardment (FAB+) on a JEOL JMS-600H instrument or by matrix-assisted, LASER desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) on an Applied Biosystems Voyager DE-Pro spectrometer

using α-cyano-4-hydroxycinnamic acid as matrix.

4.4.2 BocHN-(R)$^{\beta\text{-CbzHN}}$γ-ImPyPyPy-CO_2H

Synthesized as previously described in Chapter 3 of this thesis.

*4.4.3 BocHN-(R)$^{\beta\text{-CbzHN}}$γ-ImPyPyPy-CO_2Pfp (**4**)*

A solution of BocHN-(*R*)$^{\beta\text{-CbzHN}}$γ-ImPyPyPy-CO_2H (100 mg, 0.119 mmol) and DCC (49 mg, 0.238 mmol) in CH_2Cl_2 (5.2 mL) was stirred at 23 ºC for 45 min. The solution was then treated with DMAP (1.4 mg, 0.012 mmol) followed by pentafluorophenol (131.2 mg, 0.713 mmol) and stirred at 23 ºC for 12 h. The reaction mixture was then loaded onto a silica gel column with CH_2Cl_2 and eluted with step gradients of 100% CH_2Cl_2 to 100% acetone with incremental steps of 5% acetone. The product was concentrated *in vacuo* to yield BocHN-(*R*)$^{\beta\text{-CbzHN}}$γ-ImPyPyPy-CO_2Pfp (**4**) as an off-white solid (84 mg, 71%). ^{1}H NMR (500 MHz, DMSO-d_6): δ 10.16 (s, 1H), 10.08 (s, 1H), 9.99 (s, 1H), 9.97 (s, 1H), 7.74 (d, J = 1.8 Hz, 1H), 7.44 (s, 1H), 7.33 – 7.30 (m, 5H), 7.29 (d, J = 1.9 Hz, 1H), 7.27 (d, J = 1.8 Hz, 1H), 7.25 (d, J = 1.8 Hz, 1H), 7.14 (d, J = 1.8 Hz, 1H), 7.13 (d, J = 1.8 Hz, 1H), 7.03 (d, J = 8.4 Hz, 1H), 6.80 (t, J = 5.8 Hz, 1H), 4.98 (s, 2H), 3.95 (m, 4H), 3.89 (s, 3H), 3.860 (s, 3H), 3.856 (s, 3H), 3.03 (m, 2H), 2.46 (m, 2H), 1.36 (s, 9H). HRMS (FAB+) calc'd for $C_{46}H_{47}N_{11}O_{10}F_5$ $[M+H]^+$ 1008.343, found 1008.342.

4.4.4 Oligomerization procedure

A glass vial (1 dram) was charged with **4** (5.0 mg, 4.96 μmole) and treated with a solution of CF_3CO_2H in CH_2Cl_2 (1:1 CF_3CO_2H:CH_2Cl_2, 1 mL) and stirred at 23 °C for 10 min. The solvent was removed in vacuo and the residual solid was dried under high vacuum for 20 min. The solid was diluted with DMSO (500 μL) followed by DIEA (80 μL) and the solution was stirred at 23 °C for 20 h. After 20 h the reaction was complete by analytical HPLC analysis. The reaction was diluted to a final volume of 10 mL by addition of a solution of DMF in aqueous CF_3CO_2H (2:3 DMF:0.1% aqueous CF_3CO_2H). NOTE: A small amount of yellow insoluble material was observed and discarded. Purification by RP-HPLC yielded **1z** (13.9% yield), **2z** (5.5% yield), and **3z** (2.1% yield). The yield of **1z** is calculated from the mass of the purified and isolated material (0.5 mg). Yields for **2z** and **3z** were calculated based on **1z** using the relative product distribution as measured by integration of the preparative HPLC chromatogram at 310 nm (product distribution: 6.6:2.6:1.0 ratio of **1z**:**2z**:**3z**; UV integral values were normalized to the number of ImPyPyPy strands contained

in each cyclic oligomer). The benzyl carbamate (Cbz) protecting groups of **1z**-**3z** were removed as previously described.[5g] Characterization data for dimer **1** has been reported previously in Chapter 3 of this thesis.[11] Trimer **2** HRMS (MALDI-TOF) calc'd for $C_{81}H_{94}N_{33}O_{15}$ $[M+H]^+$ 1768.7607, found 1768.7566. Tetramer **3** HRMS (MALDI-TOF) calc'd for $C_{108}H_{125}N_{44}O_{20}$ $[M+H]^+$ 2358.0112, found 2358.0143.

4.4.5 UV Absorption Spectrophotometry

Melting temperature analysis was performed on a Varian Cary 100 spectrophotometer equipped with a thermo-controlled cell holder possessing a cell path length of 1 cm. A degassed aqueous solution of 10 mM sodium cacodylate, 10 mM KCl, 10 mM $MgCl_2$, and 5 mM $CaCl_2$ at pH 7.0 was used as analysis buffer. DNA duplexes and hairpin polyamides were mixed in 1:1 stoichiometry to a final concentration of 2 μM for each experiment. Prior to analysis, samples were heated to 90 °C and cooled to a starting temperature of 23 °C with a heating rate of 5 °C/min for each ramp. Denaturation profiles were recorded at λ = 260 nm from 23 °C to 90 °C with a heating rate of 0.5 °C/min. The reported melting temperatures were defined as the maximum of the first derivative of the denaturation profile.

4.5 References and Notes

1. Dervan, P. B. *Bioorg. Med. Chem.* **2001**, *9*, 2215–2235.

2. Dervan, P. B., and Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284–299.

3. (a) Trauger, J. W., Baird, E. E., and Dervan, P. B. *Nature* **1996**, *382*, 559–561. (b) White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., and Dervan, P. B. *Nature* **1998**, *391*, 468–470. (c) Kielkopf, C. L., Baird, E. E., Dervan, P. B., and Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104–109. (d) Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B., and Rees, D. C. *Science* **1998**, *282*, 111–115.

4. Hsu, C. F., Phillips, J. W., Trauger, J. W., Farkas, M. E., Belitsky, J. M., Heckel, A., Olenyuk, B. Z., Puckett, J. W., Wang, C. C. C., and Dervan, P. B. *Tetrahedron* **2007**, *63*, 6146–6151.

5. (a) Belitsky, J. M., Leslie, S. J., Arora, P. S., Beerman, T. A., and Dervan, P. B. *Bioorg. Med. Chem.* **2002**, *10*, 3313–3318. (b) Crowley, K. S., Phillion, D. P., Woodard, S. S., Scheitzer, B. A., Singh, M., Shabany, H., Burnette, B., Hippenmeyer, P., Heitmeier, M., and Bashkin, J. K. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1565–1570. (c) Best, T. P., Edelson, B. S., Nickols, N. G., and Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 12063–12068. (d) Edelson, B. S., Best, T. P., Olenyuk, B., Nickols, N. G., Doss, R. M., Foister, S., Heckel, A., and Dervan, P. B. *Nucleic Acids Res.* **2004**, *32*, 2802–2818. (e) Xiao, X., Yu, P., Lim, H. S., Sikder, D., and Kodadek, T. *Angew. Chem. Int. Ed.* **2007**, *46*, 2865–2868. (f) Nickols, N. G., Jacobs, C. S., Farkas, M. E., and Dervan, P. B. *Nucleic Acids Res.* **2007**, *35*, 363–370. (g) Dose, C., Farkas, M. E., Chenoweth, D. M., and Dervan, P. B. *J. Am. Chem. Soc.* **2008**, *130*, 6859–6866. (h) Hsu, C. F., and Dervan, P. B. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5851–5855.

6. (a) Olenyuk, B. Z., Zhang, G. J., Klco, J. M., Nickols, N. G., Kaelin, Jr., W. G., and Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16768–16773. (b) Kageyama, Y., Sugiyama, H., Ayame, H., Iwai, A., Fujii, Y., Huang, L. E., Kizaka-Kondoh, S., Hiraoka, M., and Kihara, K. *Acta. Oncol.* **2006**, *45*, 317–324. (c) Nickols, N. G., Jacobs, C. S., Farkas, M. E., and Dervan, P. B. *ACS Chem. Biol.* **2007**, *2*, 561–571. (d) Nickols, N. G., and Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 10418–10423. (e) Matsuda, H., Fukuda, N., Ueno, T., Tahira, Y., Ayame, H., Zhang, W., Bando, T., Sugiyama, H., Saito, S., Matsumoto, K., and others, O. *J. Am. Soc. Nephrol.* **2006**, *17*, 422–432. (f) Yao, E. H., Fukuda, N., Ueno, T., Matsuda, H., Matsumoto, K., Nagase, H., Matsumoto, Y., Takasaka, A., Serie, K., Sugiyama, H., and Sawamura, T. *Hypertension* **2008**, *52*, 86–92.

7. (a) Rucker V. C.; Foister S.; Melander C.; Dervan P. B. *J. Am. Chem. Soc.* **2003**, *125*, 1195–1202; (b) Chenoweth, D. M.; Viger, A.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2216–2217.

8. (a) Arndt, H. D.; Hauschild, K. E.; Sullivan, D. P.; Lake, K.; Dervan, P. B.; Ansari, A. Z. *J. Am. Chem. Soc.* **2003**, **125**, 13322–13323; (b) Kwonj, Y.; Arndt, H. D.; Qian, M.; Choi, Y.; Kawazoe, Y.; Dervan, P. B.; Uesugi, M. *J. Am. Chem. Soc.* **2004**, *126*, 15940–15941; (c) Hauschild, K. E.; Metzler, R. E.; Arndt, H. D.; Moretti, R.; Raffaelle, M.; Dervan, P. B.; Ansari, A. Z. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 5008–5013; (d) Stafford, R. L.; Arndt, H. D.; Brezinski, M. L.; Ansari, A. Z.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 2660–2668; (e) Stafford, R. L.; Dervan, P. B. *J. Am. Chem. Soc.* **2007**, *129*, 14026–14033; (f) Xiao, X.; Yu, P.; Lim, H. S.; Sikder, D.; Kodadek, T. *Angew. Chem. Int.* **2007**, *46*, 2865–2868.

9. (a) Cohen, J. D.; Sadowski, J. P.; Dervan, P. B. *Angew. Chem. Int. Ed.* **2007**, *46*, 7956–7959; (b) Schmidt, T. L.; Nandi, C. K.; Rasched, G.; Parui, P. P.; Brutschy, B.; Famulok, M.; Heckel, A. *Angew. Chem. Int. Ed.* **2007**, *46*, 4382–4384; (c) Cohen J. D.; Sadowski J. P.; Dervan P. B. *J. Am. Chem. Soc.* **2008**, *130*, 402–403.

10. Chenoweth, D. M.; Harki, D. A.; Dervan, P. B. *J. Am. Chem. Soc.* **2009**, In Press. (See Chapter 2 of this thesis)

11. Chenoweth, D. M.; Harki, D. A.; Phillips, J. W.; Dose, C.; Dervan, P. B. *J. Am. Chem. Soc.* **2009**, In Press. (See Chapter 3 of this thesis)

12. Compound **1** targets the DNA sequence 5'-WGWWCW-3', where W =A/T or T/A.

4.6 Spectra and Supplemental Information

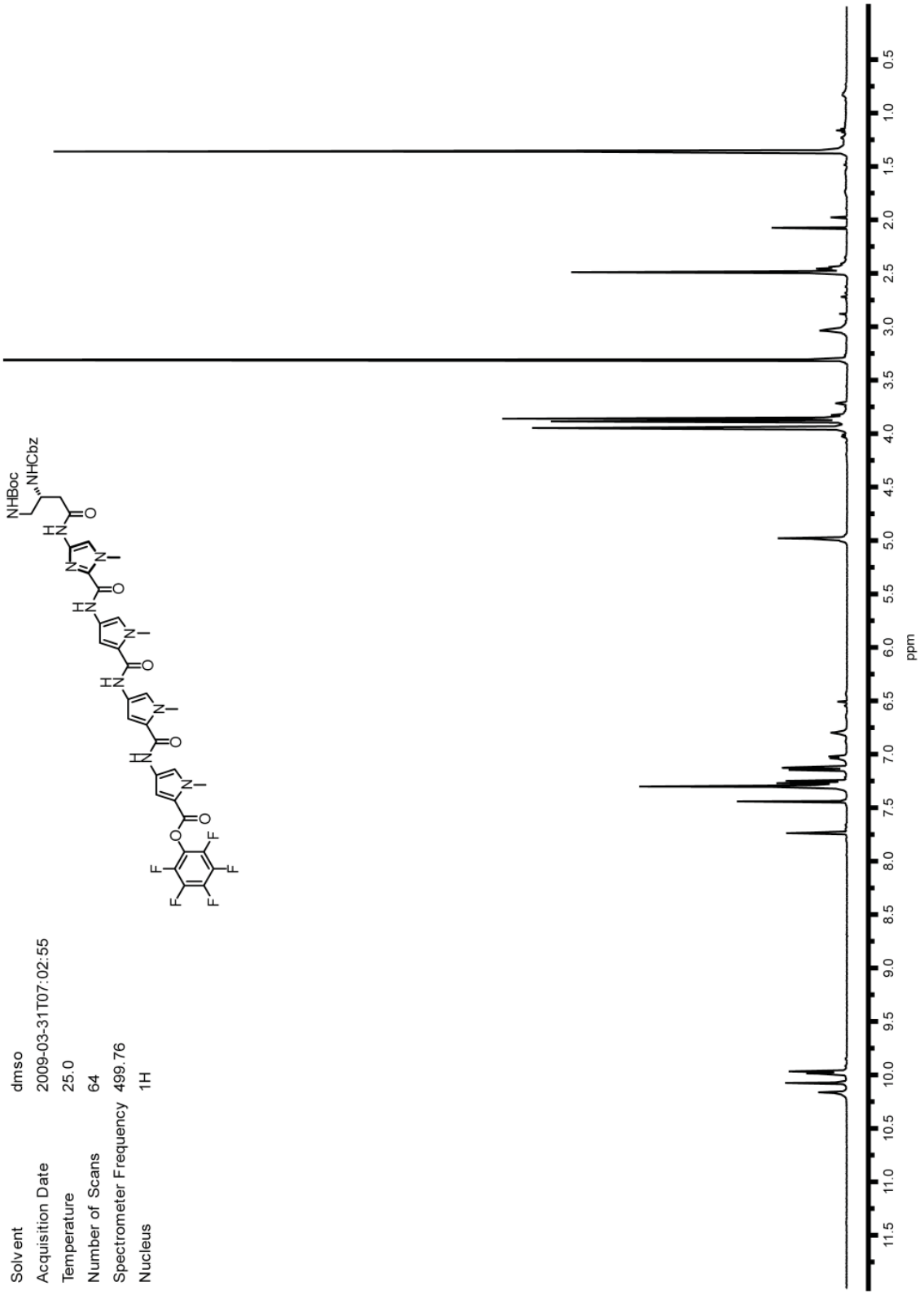


Figure 4.4 ^{1}H NMR BocHN-(*R*)$^{\beta\text{-CbzHN}}$γ-ImPyPyPy-CO_2Pfp (**4**)