Chapter 8

Next-Generation Programmable Oligomers for DNA Minor Groove Recognition

The text of this chapter was taken in part from a manuscript coauthored with Raymond M. Doss and Professor Peter B. Dervan (Caltech)

(Marques, M. A.; Doss, R. M. and Dervan, P. B. "*Next-Generation Programmable Oligomers for DNA Minor Groove Recognition*" *Journal of the American Chemical Society* **2005 –** In Preparation)

Communication.

The ability to reprogram cells that present aberrant gene expression using specific anti-gene small molecules is one approach to treating a wide variety of genetic diseases.[1-3] Traditional polyamides based on the natural products netropsin and distamycin offer one such approach.[4, 5] We have previously described many different polyamide motifs that code for specific DNA sequences using the anti parallel cofacial pairings of 5-membered heterocyclic carboxamides N-methylpyrrole (**Py**), N-methylimidazole (**Im**), and N-methylhydroxypyrrole (**Hp**), such that an Im/Py pair targets G●C, Py/Im targets C●G, Hp/Py targets T●A and Py/Hp targets A●T.[6-11] While these polyamides bind with high selectivity and affinity comparable to naturally occurring DNA-binding proteins, we became interested in developing new recognition elements to expand the suite of tools available for DNA recognition. Early studies employing small rationally designed libraries yielded a small handful of new 5-membered heterocyclic amino acids that successfully bound DNA at high affinity and imparted previously unrealized specificity.[12-16] After working with a large sampling of 5-membered heterocyles, a search for novel DNA recognition scaffolds, namely, fused 6-5 bicyclic benzimidazole analogues, was initiated. In addition to the structural and electronic variation provided by the fused ring systems, the ability of these benzimidazole analogues to impart favorable pharmacological profiles with respect to nuclear uptake was also a consideration.

Initial research incorporating one or two benzimidazole analogues into an 8-ring hairpin demonstrated that this new architecture successfully mimicked the classic 5-membered carboxamide design, with benzimidazole (**Bi**), imadazopyridine (**Ip**), and hydroxybenzimidazole (**Hz**) mimicking pyrrole (**Py**), imidazole (**Im**), and

hydroxypyrrole (**Hp**), respectively (Figure 8.1).[17-20] The new fused bicycles offer several benefits. First, the benzimidazole analogues bind DNA with good specificity and increased affinity. The increase in affinity may be attributed to a number of factors including a larger hydrophobic surface area and higher degree of preorganization, reducing the entropic cost of complexation. Synthesis of polyamides containing the benzimidazole analogues is also more convergent, as they are coupled onto the resin as dimers. The dimeric nature of the fused bicycles replace the formal carboxamide bonds that link adjacent 5-membered heterocycles, as in the case of classic polyamides, with an exocyclic N-H that is still able to hydrogen bond to the DNA minor groove. The use of benzimidazole dimers also changes the design aspect of how DNA sequences are targeted. More specifically, with individual 5-membered heterocyles, each heterocycle targets a specific DNA base. With the use of benzimidazole dimers, the DNA is coded for two bases at a time, as each section of the benzimidazole dimmer codes for an individual base. Taking advantage of the benefits imparted by the benzimidazole technology, we envisioned developing a specific DNA binding ligand that was no longer a polyamide, but a programmable oligomer composed entirely of fused benzimidazole analogues. To test the DNA recognition profile of this new class of compounds, oligomer **1** NoHz-PyBi-γ*-ImHz-

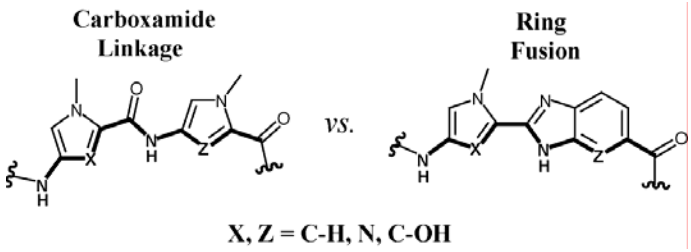

Figure 8.1. Comparison of classic 5-membered heterocycles connected by carboxamide linkages versus fused bicyclic benzimidazole analogues. In both sets of structures the same chemical functionality is presented to the DNA minor groove.

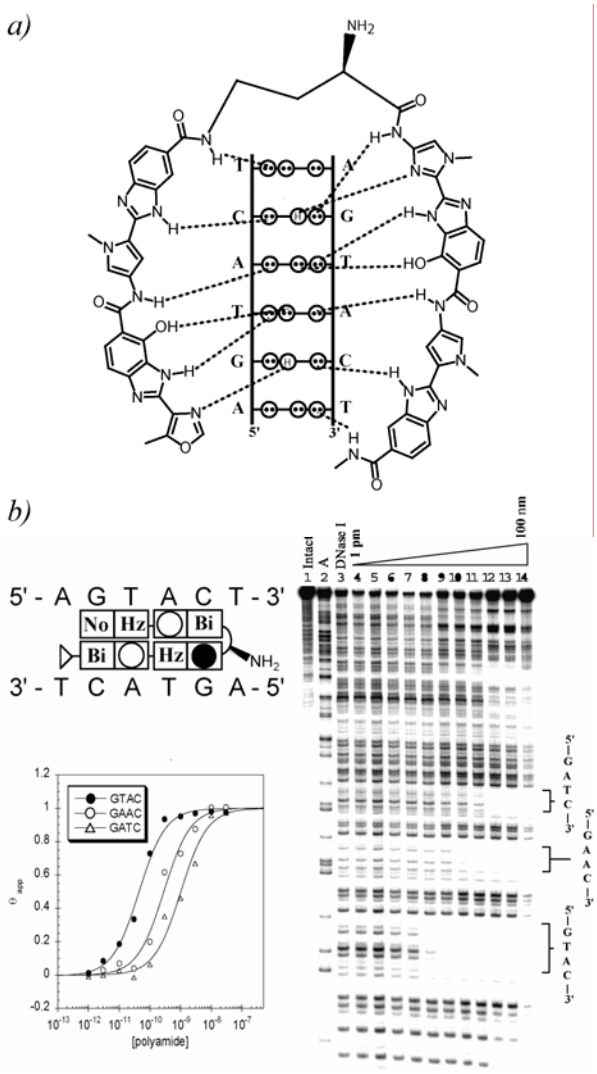

Figure 8.2. (a) Chemical structure of oligomer **1** bound to its respective DNA match site. Oligomer-DNA hydrogen bonds are indicated by a dashed line. (b) Above left: box structure of oligomer **1** showing NoHz, PyBi, and ImHz dimmers. Below left: Binding isotherm showing association constants of oligomer **1** for its targeted match site 5'-aGTACt-3', single base mismatch site 5'-aGAACt-3', and double base mismatch site 5'aGATCt-3'. Right: Quantitative DNase I footprinting experiment for oligomer **1** on the 285 bp, 5'-end-labelled PCR product of plasmid DEH10: lane 1, intact DNA; lane 2, A reaction; lane 3, DNase I standard; lanes 4-14, 1 pM, 3 pM, 10 pM, 30 pM, 100 pM, 300 pM, 1 nM, 3 nM, 10 nM, 30 nM, 100 nM polyamide, respectively.

PyBi-NHMe was synthesized from dimers: NoHz-OH (**2**), Boc-ImHz-OH (**3**), and Boc-PyBi-OH (**4**) using HBTU activation and manual solid phase synthesis protocols.[18-21] Oligomer **1** was then subject to quantitative DNase I footprint titration[22] on plasmid pDEH 10, which contains a match sequence 5'-a**GTAC**t-3', a single base mismatch site 5'-a**GAAC**t-3', and a double base mismatch site 5'-a**GATC**t-3'. The single and double Watson-Crick base pair mismatches place an adenine across from the **Hz** ring. We have recently shown the Hz/Py and Py/Hz pairings to specify for T●A and A●T respectively. Molecular modeling was used to complement the footprint titrations.[23]

Examination of the thermodynamic data shows that oligomer **1** binds its designed DNA

match site 5'-a**GTAC**t-3' with high affinity (K_a 10^{10} M^{-1}) and good specificity over its single and double base pair mismatch sites, $K_a \approx 10^9$ M^{-1} and 10^8 M^{-1} respectively (Figure 8.2 & Table 8.1). Previous attempts at targeting the 5'-a**GTAC**t-3' site using traditional polyamides containing multiple Hp/Py pairs or hybrid polyamides containing two Hz/Py pairs were moderately successful. More specifically, the traditional 8-ring polyamide Im-Hp-Py-Py-γ-Im-Hp-Py-Py-β-Dp **5**, composed of 5-membered heterocyclic pairs, bound the 5'-a**GTAC**t-3' site with much lower affinity.[24] The hybrid polyamide Im-Hz-Py-Py-γ-Im-Hz-Py-Py-β-Dp **6**, containing two hydroxybenzimidazole (**Hz**) dimers, managed to bind the 5'a**GTAC**t-3' sequence with a mild increase in affinity while maintaining similar specificity.[20] In contrast, oligomer **1**, a new class of compound composed only of fused benzimidazole dimers, binds the target match site with a marked increase in

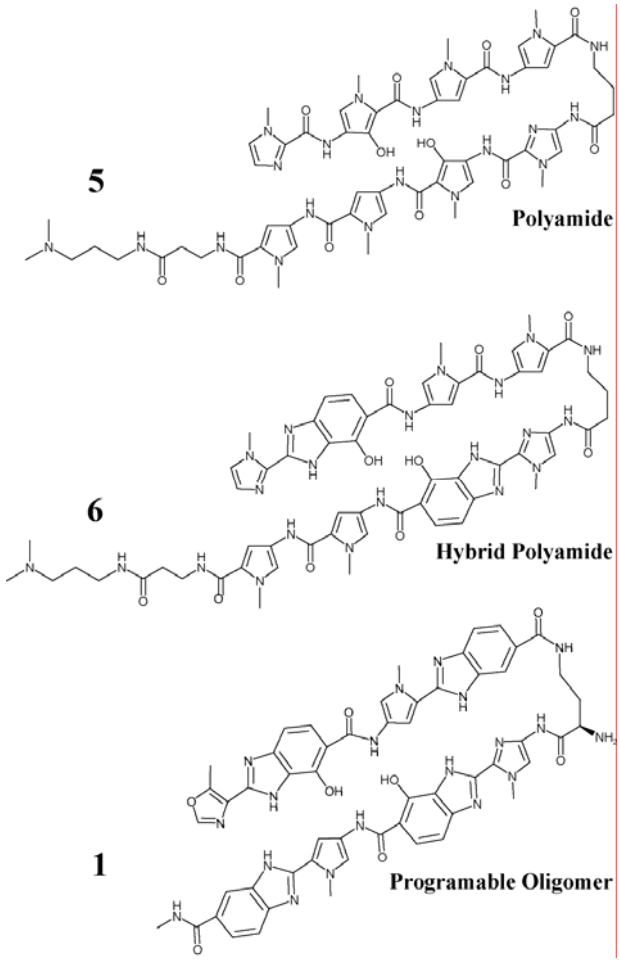

Figure 8.3. Progression of compound design: classic polyamides containing 5-membered heterocyclic amino acids linked as carboxamides **5**; hybrid polyamides containing multiple benzimidazole analogues **6**; programmable oligomers containing only fused benzimidazole analogues and four total carboxamide linkages **1**.

affinity and comparable specificity. Of note, oligomer **1** contains only four carboxamide linkages, in comparison to the eight that are found in compounds such as polyamide **5** (Figure 8.3). The amide linkages that join the 5-membered heterocyclic amino acids have been shown to play a key role in hydrogen bonding to the DNA minor groove floor and there was concern that altering the nature of those linkages to a large extent would disrupt favorable DNA binding. However, molecular modeling of the electronic surface presented to the DNA minor groove floor by the traditional carboxamide linkages versus the exocyclic N-H of the benzimidazole ring shows that the two groups are nearly identical. Molecular modeling of the NoHz-PyBi subunit versus its polyamide comparator Im-Hp-Py-Py provides insight into the differential geometries of the two systems (Figure 8.4). First, the NoHz-PyBi system is substantially less curved. It has been shown that traditional polyamides are overly curved with respect to the DNA helix, limiting the length of the DNA sequence that can be successfully targeted.[25] It can be envisioned that the use of these new fused bicycles may allow access to targeting larger, less degenerate DNA sequences due to their greater complementarity to the DNA minor groove. Greater DNA:oligomer complementarity may also help account for the marked

Table 8.1. Equilibrium Association Constants K_a $[M^{-1}]^{a,b}$

	5'-a**GTAC**t-3'	5'-a**GAAC**t-3'	5'-a**GATC**t-3'
1	$2.3 (\pm 0.3) \times 10^{10}$	$2.5 (\pm 0.7) \times 10^{9}$	$9.7 (\pm 0.2) \times 10^{8}$

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

increase in observed affinity. In addition to differences in ligand geometry, the NoHz-PyBi system is much more conformationally restricted due to the ring fusion. The large degree of ligand preorganization is also likely to play a central role in the higher affinities observed for this class of compounds. Finally, with respect to electronics, the larger hydrophobic surface area of the oligomeric system may make DNA complexation more favorable.

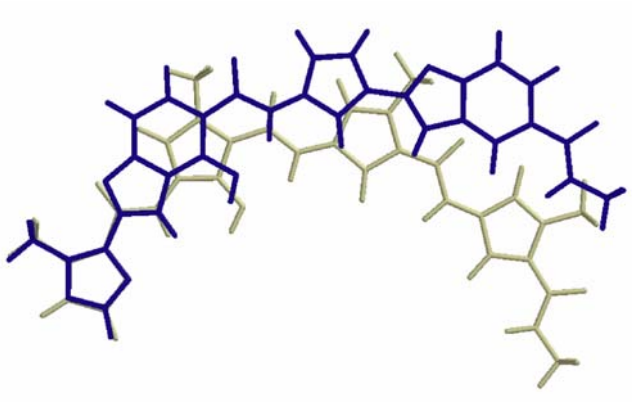

Figure 8.4. Comparison of polyamide and oligomer architectures. The polyamide subunit containing 5-membered heterocyclic amino acids and connected by carboxamide linkages (tan) is shown overlaid with the fused benzimidazole system (blue). *Ab initio* calculations done using a 6-31G* basis set.

We have shown the successful transition from classic polyamides inspired by the natural products netropsin an distamycin, composed of 5-membered heterocyclic carboxamides, to a new class of compound composed entirely of fused benzimidazole analogues. These new programmable oligomers have been shown to bind DNA with high fidelity and affinity, and offer a new scaffold for further design work. Currently, we are interested in replacing the final bridging carboxamide linkage with a trans alkene. Initial molecular modeling is promising and the molecular recognition capabilities of this new series of compounds will be reported in due course. Furthermore, the biological properties of these new compounds are also of interest and cellular trafficking studies are underway.

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

Supplemental Information: Oligomer 1 Synthesis. Oligomer **1** was prepared using Kaiser oxime resin (0.65 meqiv/gram) from Nova Biochem (Figure 8.5). 0.1 grams of resin was added to a solid phase synthesis vessel. The resin was then washed with DCM (15 mL) followed by DMF (15 mL). In a separate vessel was added Boc-PyBi-OH (**4**) (60 mg, 0.168 mmol), HBTU (61 mg, 0.159 mmol), DIEA (43 mg, 60 μL, 0.336 mmol), and DMF (400 μL). The mixture was vortexed and allowed to activate for 20 minutes at room temperature. The activated mixture was then added to the resin. The reaction vessel was shaken at room temperature overnight. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). A solution of 25% TFA in DCM was then flowed over the resin (20 mL), followed by shaking at room temperature for 25 min to deblock the Boc-protected amine of pyrrole. In a separate vessel, a mixture of Boc-ImBi-OH (**3**) was activated using HBTU, DIEA, and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated Boc-ImHz-Obt dimer was then added to the vessel containing (Resin-BiPy-NH_2). The mixture was shaken at room temperature for 2.5 hours. The reaction vessel was then drained and washed with DMF (15 mL) and DCM (15 mL). A solution of 50% TFA in DCM was then washed over the resin (20 mL), followed by shaking at room temperature for 25 minutes. In a separate vessel, a mixture of Boc-γ*-OH (Boc-diaminobutyric acid) was

activated using HBTU, DIEA, and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated Boc-γ*-Obt monomer was then added to the vessel containing

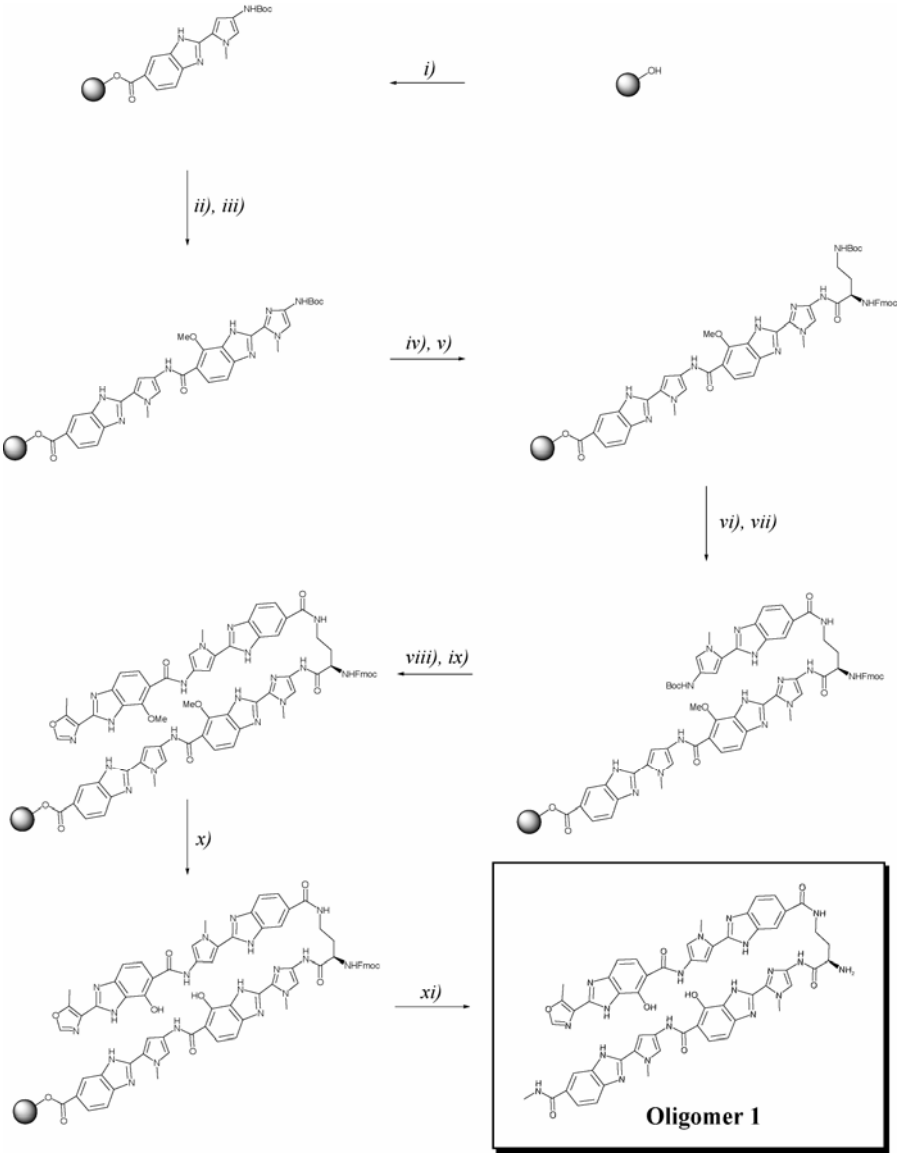

Figure 8.5. Solid-phase synthesis of oligomer 1. i) Boc-PyBi-OH, HBTU, DIEA and DMF; ii) 25% TFA in DCM; iii) Boc-ImHz-OH, HBTU, DIEA, DMF; iv) 50% TFA in DCM; v) Boc-γ*-OH, HBTU, DIEA, DMF; vi) 25% TFA in DCM; vii) Boc-PyBi-OH, HBTU, DIEA, DMF; viii) 25% TFA in DCM; ix) NoHz-OH; x) BCl_3, DCM; xi) $MeNH_2$, THF.

(Resin-BiPy-HzIm-NH_2). The mixture was shaken for 2 hours at room temperature. The

reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). A solution of 25% TFA in DCM was then flowed over the resin (20 mL), followed by shaking at room temperature for 25 min to deblock the Boc-protected alkyl amine. In a separate vessel, a mixture of Boc-PyBi-OH (**3**) was activated using HBTU, DIEA and DMF as described above. The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated Boc-PyBi-Obt dimer was then added to the vessel containing (Resin-BiPy-HzIm-γ*-NH_2). The mixture was shaken at room temperature for 2.5 hours. The reaction vessel was then drained and the resin washed with DMF (15 mL), followed by DCM (15 mL). A solution of 25% TFA in DCM was then flowed over the resin (20 mL), followed by shaking at room temperature for 25 min to deblock the Boc-protected amine of pyrrole. In a separate vessel, a mixture of NoHz-OH (**2**) was activated using HBTU, DIEA and DMF as described above for (**3**). The reaction vessel containing the resin was then drained and washed with DCM (15 mL) and 10% DIEA in DMF. The activated NoHz-Obt dimer was then added to the vessel containing (Resin-BiPy-HzIm-γ*-BiPyNH_2). The mixture was shaken at room temperature for 2.5 hours. The reaction vessel was then drained and washed with DMF (15 mL) and DCM (15 mL) to provide (Resin-BiPy-HzIm-γ*-BiPy-HzNo).

O-Methyl Deprotection: To the synthesis vessel containing oligomer **1** was added DCM (800 μL) and BCl_3 (400 μL), (2M in heptanes). The mixture was shaken at room temperature for 2.5 hours. The vessel was then drained and washed with DCM (15 mL). Deprotection was determined to be complete by analytical HPLC (Figure 8.6).

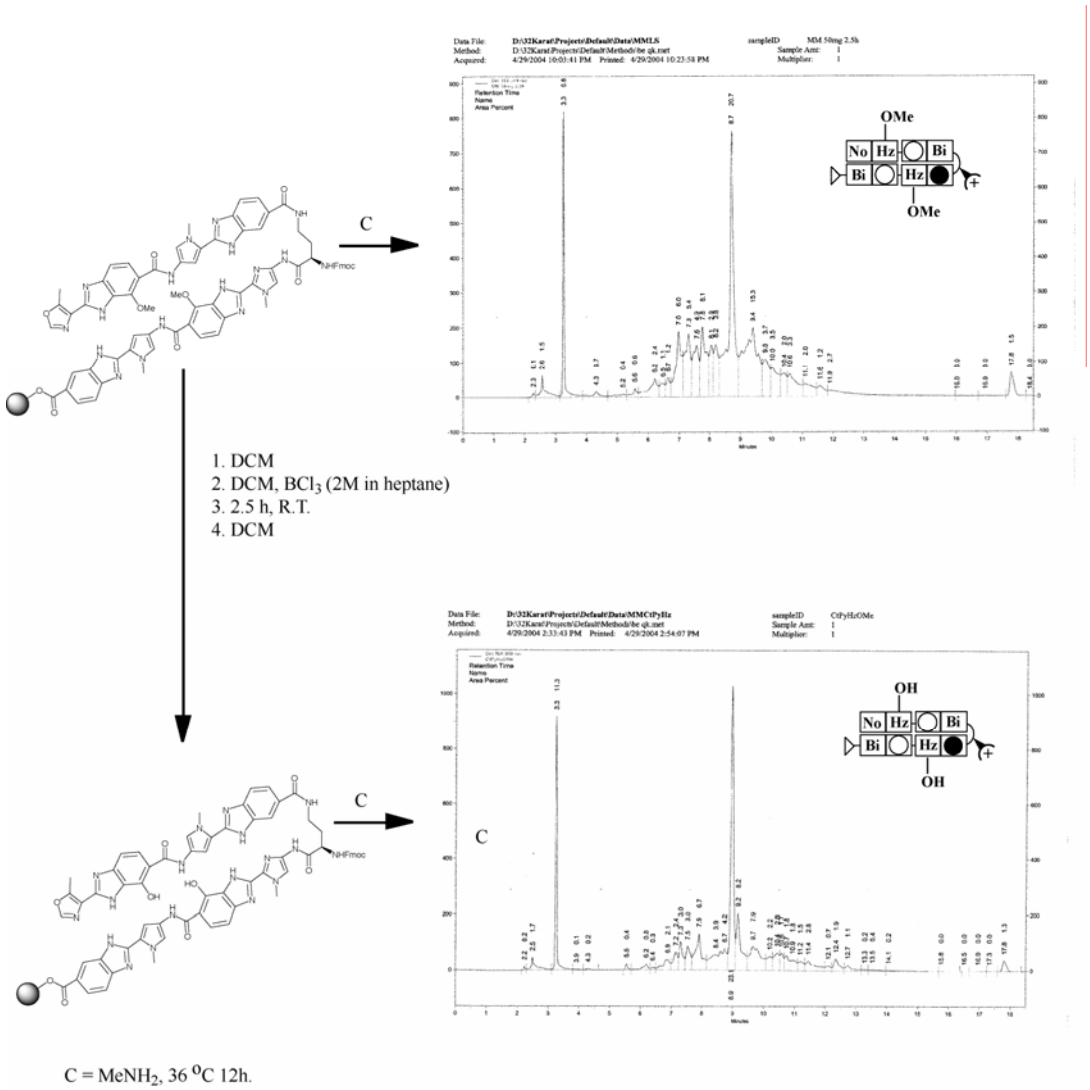

Figure 8.6. Deprotection of aryl O-methoxy group using boron trichloride protocol. deprotection was monitored by analytical HPLC.

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

References:

[1] Gottesfeld, J. M.; Turner, J. M.; Dervan, P. B., *Gene Expression* **2000,** 9, (1-2), 77-91.

[2] Pandolfi, P. P., *Oncogene* **2001,** 20, (24), 3116-3127.

[3] Darnell, J. E., *Nature Reviews Cancer* **2002,** 2, (10), 740-749.

[4] Arcamone, F.; Nicolell.V; Penco, S.; Orezzi, P.; Pirelli, A., *Nature* **1964,** 203, (494), 1064.

[5] Finlay, A. C.; Hochstein, F. A.; Sobin, B. A.; Murphy, F. X., *Journal of the American Chemical Society* **1951,** 73, (1), 341-343.

[6] Geierstanger, B. H.; Mrksich, M.; Dervan, P. B.; Wemmer, D. E., *Science* **1994,** 266, (5185), 646-650.

[7] Dervan, P. B.; Edelson, B. S., *Current Opinion in Structural Biology* **2003,** 13, (3), 284-299.

[8] Dervan, P. B.; Burli, R. W., *Current Opinion in Chemical Biology* **1999,** 3, (6), 688-693.

[9] Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2001,** 9, (9), 2215-2235.

[10] Kielkopf, C. L.; White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B.; Rees, D. C., *Science* **1998,** 282, (5386), 111-115.

[11] Kielkopf, C. L.; Baird, E. E.; Dervan, P. D.; Rees, D. C., *Nature Structural Biology* **1998,** 5, (2), 104-109.

[12] Ellervik, U.; Wang, C. C. C.; Dervan, P. B., *Journal of the American Chemical Society* **2000,** 122, (39), 9354-9360.

[13] Foister, S.; Marques, M. A.; Doss, R. M.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2003,** 11, (20), 4333-4340.

[14] Marques, M. A.; Doss, R. M.; Urbach, A. R.; Dervan, P. B., *Helvetica Chimica Acta* **2002,** 85, (12), 4485-4517.

[15] Doss, R. M.; Marques, M. A.; Foister, S.; Dervan, P. B., *Chemistry & Biodiversity* **2004,** 1, (6), 886-899.

[16] Nguyen, D. H.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B., *Bioorganic & Medicinal Chemistry* **2001,** 9, (1), 7-17.

[17] Minehan, T. G.; Gottwald, K.; Dervan, P. B., *Helvetica Chimica Acta* **2000,** 83, (9), 2197-2213.

[18] Briehn, C. A.; Weyermann, P.; Dervan, P. B., *Chemistry-a European Journal* **2003,** 9, (9), 2110-2122.

[19] Renneberg, D.; Dervan, P. B., *Journal of the American Chemical Society* **2003,** 125, (19), 5707-5716.

[20] Marques, M. A.; Doss, R. M.; Foister, S.; Dervan, P. B., *Journal of the American Chemical Society* **2004,** 126, (33), 10339-10349.

[21] Baird, E. E.; Dervan, P. B., *Journal of the American Chemical Society* **1996,** 118, (26), 6141-6146.

[22] Trauger, J. W.; Dervan, P. B., *Methods in Enzymology* **2001,** 340, 450-466.

[23] *Sparten Essential*, Wavefunction Inc: 2000.

[24] Herman, D. E. California Institute of Technology, Pasadena, 2001.

[25] Kelly, J. J.; Baird, E. E.; Dervan, P. B., *Proceedings of the National Academy of Sciences of the United States of America* **1996,** 93, (14), 6981-6985.