Rational Design of Sequence-Specific DNA Alkylating Agents Based on Duocarmycin A and Pyrrole−Imidazole Hairpin Polyamides

Zhi-Fu Tao,[†] Tsuyoshi Fujiwara,[†] Isao Saito,[‡] and Hiroshi Sugiyama*,[†]

Contribution from the Division of Biofunctional Molecules, Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University, 2-3-10 Surugadai, Kanda, Chiyoda, Tokyo 101-0062, Japan, Department of Synthetic Chemistry and Biological Chemistry, Faculty of Engineering, Kyoto University, Yoshida, Sakyo, and CREST, Japan Science and Technology Corporation (JST), Kyoto 606-8501, Japan

Received September 21, 1998. Revised Manuscript Received March 15, 1999

Abstract: Synthesis of novel conjugates between segment A of Duocarmycin A (Du) and *N*-methylimidazole (Im)−*N*-methylpyrrole (Py) hairpin polyamides and their DNA alkylation are described. The conjugates PyPyPyγImPyPyDu (**8a**) and ImPyPyγImPyPyDu (**8b**) were designed to alkylate the target sequences (A/T)G(A/T)$_2$N(A/G) and (A/T)G(A/T)CN(A/G), respectively, according to Dervan's ring-pairing rule. High-resolution denaturing gel electrophoresis indicated that **8a** exclusively alkylated the A of the 5′-TGTAAAA-3′ within a ∼400 bp DNA fragment. Similarly, alkylation by **8b** occurred exclusively at the G of the 5′-AGTCAGA-3′ sequence with efficiency at nanomolar concentration. To better understand the structure of the alkylated DNA by these conjugates, the alkylation of non-self-complementary duplex decanucleotides, ODN1 and ODN2, was investigated. HPLC and ESMS analyses of the reaction of these ODNs with **8a** and **8b** demonstrated that both conjugates efficiently and selectively alkylate N3 of the purine bases of their target sequences.

Introduction

Sequence-specific DNA alkylations have been a topic of much current interest due to their significant potential in molecular biology and human medicine.[1,2] The sequence specificities induced by DNA alkylators usually result from the combination of their inherent reactivities toward purine bases and their binding selectivities for DNA. While alkylation of predetermined sequences has been reported to be achieved by triple-helix formation through a tethering alkylating group to polypyrimidine oligonucleotides,[2] the triplex helix approach is limited to purine tracks which severely restricts the availability of target sites and suffers from poor cellular uptake. Dervan and Wemmer's research team has developed *N*-methylimidazole (Im)−*N*-methylpyrrole (Py) polyamides as a new class of synthetic DNA binding molecules that recognize any desired DNA sequences.[3] These polyamides bind cooperatively as an antiparallel dimer to the minor groove of the DNA helix. They have developed a simple binary code to correlate the reading DNA sequence with the side-by-side pairing between Py and Im carboxyamides, i.e., Im/Py recognizes G·C base pairs and Py/Py recognizes A·T or T·A base pairs.[3a,4] To break the latter degeneracy, Dervan and colleagues recently introduced 3-hydroxypyrrole to discriminate A·T from T·A.[5]

Duocarmycin A (Duo) is a highly potent antitumor antibiotic, which binds to A·T-rich sequences and selectively alkylates N3 of adenine (A) at the 3′ end of three or more consecutive A·T base pairs in DNA.[6] We discovered that the addition of distamycin A (Dist) markedly modulates the Duo alkylation site of DNA fragments wherein the efficient alkylation occurs

[†] Tokyo Medical and Dental University.

[‡] Kyoto University.

(1) For recent examples of sequence-specific DNA alkylating agents, see: (a) Lukhtanov, E. A.; Kutyavin, I. V.; Gorn, V. V.; Reed, M. W.; Adams, A. D.; Lucas, D. D.; Meyer, R. B., Jr. *J. Am. Chem. Soc.* **1997**, *119*, 6214. (b) Nakatani, K.; Okamoto, A.; Saito, I. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2794. (c) Nakatani, K.; Shirai, J.; Sando, S.; Saito, I. *J. Am. Chem. Soc.* **1997**, *119*, 7625. (d) Hartley, J. A. *Molecular Basis of Specificity in Nucleic Acid-Drug Interaction*; Pullman, B., Jortner, J., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1990; pp 513−530. (e) Baker, B. F.; Dervan, P. B. *J. Am. Chem. Soc.* **1985**, *107*, 8266.

(2) (a) Moser, H. E.; Dervan, P. B. *Science*, **1987**, *238*, 645. (b) Thoung, N. T.; Helene, C. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 666. (c) Lukhtanov, E. A.; Mills, A. G.; Kutyavin, I. V.; Gorn, V. V.; Reed, M. W.; Meyer, R. B. *Nucleic Acids Res*. **1997**, *25*, 5077. (d) Belousov, E. S.; Afonina, I. A.; Podyminogin, M A.; Gamper, H. B.; Reed, M. W.; Wydro, R. M.; Meyer, R. B. *Nucleic Acids Res.* **1997**, *25*, 3440. (e) Taylor, M. J.; Dervan, P. B. *Bioconj. Chem*. **1997**, *8*, 354.

(3) For recent reviews, see: (a) Wemmer, D. E.; Dervan, P. B. *Curr. Opin. Struct. Biol.* **1997**, *7*, 355. Nielsen, P. E. *Chem. Eur. J.* **1997**, *3*, 505. (b) Turner, J. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 7636. (c) Lamamie de Clairac, R. P.; Geierstanger, B. H.; Mrksin, M.; Dervan, P. B.; Wemmer, D. E. *J. Am. Chem. Soc.* **1997**, *119*, 7909. (d) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1997**, *119*, 6953. (e) Chen, X.; Ramakrishnan, B.; Sundaralinggam, M. *J. Mol. Biol.* **1997**, *267*, 1157. (f) Swalley, S. E.; Baird, E. E.; Dervan, P. B. *Chem. Eur. J.* **1997**, *3*, 1600. (g) Kielkopf, C. L.; Baird, E. E.; Dervan, P. B.; Rees, D. C. *Nat. Struct. Biol.* **1998**, *5*, 104. (h) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 1421. (i) Herman, D. M.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1998**, *120*, 1382. (j) Bremer, R. E.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1998**, *5*, 119.

(4) (a)White, S.; Baird, E. E.; Dervan, P. B. *Chem. Biol.* **1997**, *4*, 569. (b) White, S.; Baird, E. E.; Dervan, P. B. *Biochemistry* **1996**, *35*, 12532.

(5) (a) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468. (b) White, S.; Turner, J. M.; Szewczyk, J. W.; Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1999**, *121*, 260.

(6) (a) Takahashi, I.; Takahashi, K.; Ichimura, M.; Morimoto, M.; Asano, K.; Kawamoto, I.; Tomita, F.; Nakano, H. *J. Antibiot.* **1988**, *41*, 1915. (b) Ichimura, M.; Muroi, K.; Asano, K.; Kawamoto, I.; Tomita, F.; Morimoto, M.; Nakano, H. *J. Antibiot.* **1988**, *41*, 1285. (c) Yasuzawa, T.; Iida, K.; Muroi, K.; Ichimura, M.; Takahashi, K.; Sano, H. *Chem. Pharm. Bull.* **1988**, *36*, 3728. (d) Ohba, K.; Watabe, H.; Sasaki, T.; Takeuchi, Y.; Komada, Y.; Nakazawa, T.; Yamamoto, H.; Shomura, T.; Sezaki, M.; Kondo, S. *J. Antibiot.* **1988**, *41*, 1515. (e) Sugiyama, H.; Hosoda, M.; Saito, I.; Asai, A.; Saito, H. *Tetrahedron Lett.* **1990**, *31*, 7197. (f) Sugiyama, H.; Ohmori, K.; Chan, K. L.; Hosoda, M.; Asai, A.; Saito, H.; Saito, I. *Tetrahedron Lett.* **1993**, *34*, 2179. (g) Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H.; Kitos, P. A.; Suntornwat, O. *J. Am. Chem. Soc.* **1990**, *112*, 8961. (h) Boger, D. L.; Ishizaki, T.; Zarrinmayeh, H. *J. Am. Chem. Soc.* **1991**, *113*, 6645. (i) Boger, D. *Acc. Chem. Res.* **1995**, *28*, 20. (j) Boger, D. L.; Johnson, D. S. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1438.

10.1021/ja983398w CCC: $18.00
Published on Web 05/12/1999

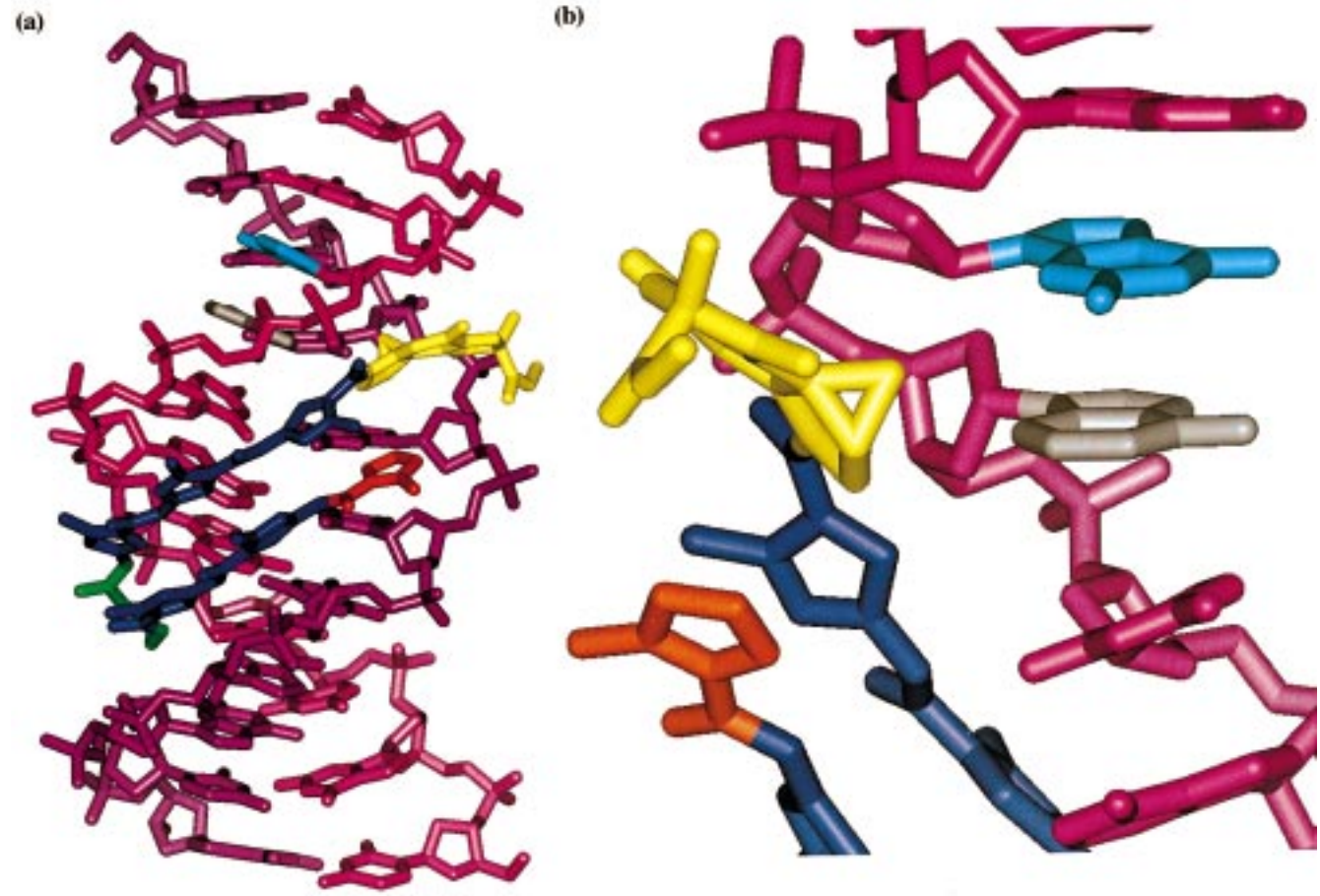


Figure 1. Energy minimized structure of ImPyPyγPyPyPyDu-d(CGCTAACAGGC)/d(GCCTGTTAGCG) complex (a) and a close-up view of the reacting region (b). Minimization was performed in the presence of 20 sodium cations and a 10 Å layer of H_2O by CFF force field. For simplicity, sodium ions and H_2O are not presented. DNA is drawn in purple. Py and Im are drawn in blue and red. The Du moiety of the hybrid is drawn in yellow. The γ-linker region is in green. The putative target G_9 and A_8 bases are in blue and gray.

primarily at the G residues in G·C-rich sequences.[7] We also demonstrated that such highly efficient G alkylation involves a cooperative heterodimer formation between Duo and Dist by high-resolution ^{1}H NMR.[8] The NMR refined structure of a Duo-Dist-deoxyoctanucleotide complex showed that the heterodimer of Duo and Dist neatly binds to the minor groove of the duplex, suggesting that Dist recognizes the complementary strand of the reacting octamer according to a similar binary code of Im/Py polyamides.[8] In fact, our preliminary studies revealed that substitutions of the Py of Dist with Im dramatically modulate the sequence specificity of Duo in a predictable manner. These findings led us to design a new class of sequence-specific DNA alkylating agents through the incorporation of the Py/Im hairpin polyamide subunit into the Duo moiety.

Results and Discussion

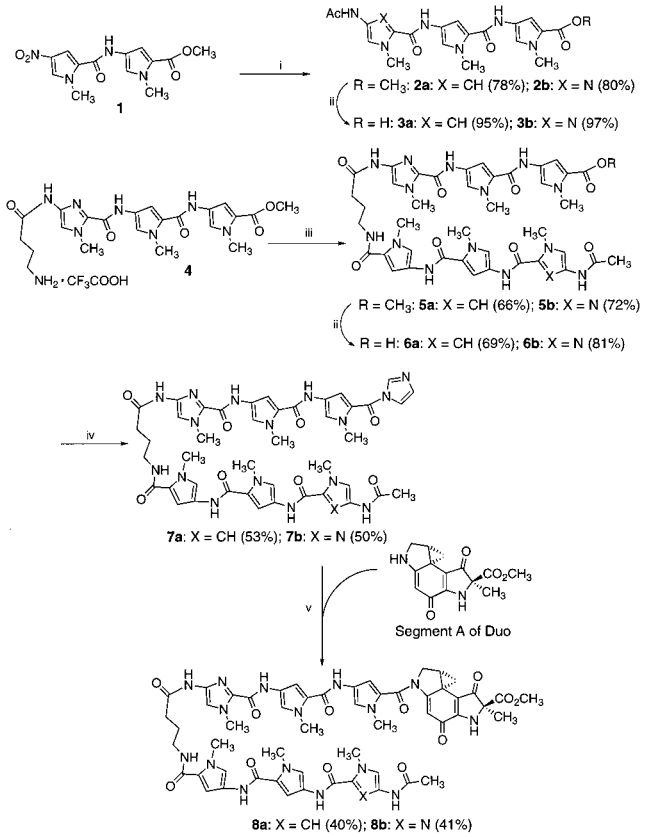
Wemmer and colleagues recently elucidated the structure of the binding complex of ImPyPyγPyPyPy to d(CGCTAA-CAGGC)/d(GCCTGTTAGCG) by ^{1}H NMR;[3c] therefore, we constructed a model of the conjugate of segment A of Duo (Du) and ImPyPyγPyPyPy based on their structure. Figure 1 shows the minimized structure of the ImPyPyγPyPyPyDu-d(CGCTAA-CAGGC)/d(GCCTGTTAGCG) complex. Molecular modeling suggests that the cyclopropane subunit of the Du moiety is located a similar distance from the nucleophilic N3 of the A_8 and G_9 residues and that one nucleotide unit (N) is probably required between the reacting base and the recognition sequence by consideration of the stereoelectronic effect of alkylation.[9]

According to the ring-pairing rule of Im/Py polyamides, the conjugates PyPyPyγImPyPyDu (**8a**) and ImPyPyγImPyPyDu (**8b**) were expected to alkylate the target sequences (A/T)G(A/T)$_2$N(A/G) and (A/T)G(A/T)CN(A/G), respectively. Both **8a** and **8b** were prepared by the synthetic routes shown in Scheme 1. Du was prepared by hydrolysis of Duocarmycin B_2 as reported previously.[10] The activated hairpin amides **7** were prepared straightforwardly from **1** via **4** and **5**.[11] The key coupling step of the activated amides **7** with Du was effected by sodium hydride in DMF to afford target compounds **8a** and **8b**, whose structures were fully characterized by ^{1}H NMR, electrospray mass spectrum (ESMS).

The sequence-selective alkylation by these conjugates was investigated by high-resolution denaturing gel electrophoresis by using TexasRed labeled DNA fragments and a DNA auto sequencer. The sites of alkylation were visualized by thermal cleavage of the DNA strand at alkylated sites.[12] Under the heating conditions, all purine N3-alkylated sites in DNA showed cleavage bands almost quantitatively on the gel, because

(7) Yamamoto, K.; Sugiyama, H.; Kawanishi, S. *Biochemistry* **1993**, *32*, 1059.
(8) Sugiyama, H.; Lian, C.; Isomura, M.; Saito, I.; Wang, A. H.-J. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14405.
(9) Tao, Z.-F.; Sugiyama, H. To be published elesewhere.
(10) Yasuzawa, T.; Muroi, K.; Ichimura, M.; Takahashi, I.; Ogawa, T.; Takahashi, K.; Sano, H.; Saito, Y. *Chem. Pharm. Bull.* **1995**, *43*, 378.
(11) Baird, E. E.; Dervan, P. B. *J. Am. Chem. Soc.* **1996**, *118*, 6141.
(12) Sugiyama, H.; Fujiwara, T.; Ura, A.; Tashiro, T.; Yamamoto, K.; Kawanishi, S.; Saito, I. *Chem. Res. Toxicol.* **1994**, *7*, 673.

Scheme 1



(i) (a) H_2, Pd/C; (b) DCC, HOBt, DMF, 4-acetoamino-1-methylpyrrole (or 1-methylimidazole)-2-carboxylic acid; (ii) NaOH, MeOH, H_2O; (iii) *N,N*-diisopropylethylamine, DCC, HOBt, DMF, **3**; (iv) 1,1′-carbonyldiimidazole, DMF; (v) NaH, DMF.

subsequent hot piperidine treatment (0.1 M, 90 °C, 20 min) did not enhance the cleavage bands indicating that the present neutral heating condition (94 °C, 20 min) is sufficient to cleave all the DNA-alkylated sites. As shown in Figure 2a, the conjugate ***8a*** *showed a single cleavage band indicating that this molecule exclusively alkylated the A of the 5′-TGTAAAA-3′ sequence within a* ∼400 bp DNA fragment. Similarly, alkylation by **8b** *occurred exclusively at the G of the 5′-AGTCAGA-3′ sequence with efficiency at nanomolar concentration of the agent* as shown in Figure 2b. The densitometric analysis of the cleavage bands indicates that the efficiencies of alkylation by **8a** and **8b** reached 1.6% and 7.4%, respectively, after 7 days. It is noteworthy that there are several other matched sequences for the Dervan ring-pairing rule, which were not alkylated by these hybrid molecules, such as 5′-AGATTGT-3′ (recognition sequence and putative alkylating site are underlined) and 5′-AGAAAAT-3′ for **8a** and 5′-TGACGAG-3′, 5′-CGACAGG-3′, 5′-GGACTAT-3′, 5′-TGTCCGC-3′, 5′-CGACCGC-3′, and 5′-AGTCCGC-3′ for **8b** in the same DNA fragments. The molecular basis for the sequence preference at flanking regions observed in both conjugates awaits high-resolution X-ray or NMR structural study. However, Dervan and colleagues recently demonstrated that the γ-linker of the hairpin binds preferentially to the AT base pair. Therefore, the observed preference for the A·T base pair at the 5′ side of the binding site can be explained by the presence of the γ-linker. A similar preference exists for the A·T base pair at one nucleotide unit between the reacting base and the recognition sequence except for 5′-AGAAAAT-3′ possessing an A-tract which is known to cause DNA bending.[13] These results clearly indicated that the observed high specificity and efficiency were attained as a result of the combination of the inherent reactivity of the Du moiety and the binding specificity of the hairpin polyamides, implying that the Du moiety itself also contributes at least in some part to determining the sequence specificity of alkylation by the hybrids.

To better understand the structure of the alkylated DNA by these conjugates, we investigated the alkylation of non-self-complementary duplex decanucleotides. ODN1 and ODN2 were designed as the target sequences for **8a** and **8b**, respectively, according to the gel experiments described above. Figure 3 shows the HPLC analysis of the reaction mixture of these ODNs with **8a** (a, top) and **8b** (b, bottom) after 19 h of incubation. It was revealed that both conjugates can efficiently and selectively alkylate their own target sequences with atom specificity. Quantitative analysis of the reaction mixture indicated that **8b** has a higher efficiency than **8a**.

The formation of the conjugate−ODN complexes in solution was directly confirmed by the ESMS. Figures 4a and 4b show the ESMS of the **8a**−ODN1 and **8b**−ODN2 complexes, respectively. Direct observation of a −4 molecular ion for the **8a**−ODN1 and of −4 and −5 molecular ions for the **8b**−ODN2 complexes by ESMS indicates that the covalently alkylated strand greatly stabilizes the association of the unreacted strand and suggests that these conjugates were strongly associated with duplex DNA. The relative intensities of these MS peaks suggested that **8b** binds to its target sequence much more strongly than **8a** does, which is consistent with the efficiency of DNA alkylation.

In conclusion, the present study outlines the facile preparation of novel conjugates between Du and Py/Im hairpin polyamides. These conjugates can efficiently alkylate G/A of predetermined DNA sequences of seven base pairs. Recently, Dervan and coworkers have demonstrated that synthetic Py-Im polyamides have a strong affinity and a full range of specificity for DNA sequences which are comparable to naturally occurring DNA binding proteins.[14] These polyamides are cell-permeable and inhibit transcription of specific genes in cell cultures.[15] Recently, Dervan and colleagues demonstrated that combination of the Py/Im hairpin polyamides effectively inhibits HIV replication.[16] Results from the present investigation suggest a promising approach for developing a new type of tailor-made sequence-specific DNA alkylating agent. Further studies on the generality and the optimization of this new class of DNA alkylation systems are currently in progress.

Experimental Section

General Methods. Reagents and solvents were purchased from standard suppliers without further purification. Abbreviations of some reagents follow: DCC, dicyclohexylcarbodiimide; HOBt, hydroxylbenzotriazole; DIEA, *N,N*-diisopropylethylamine; DMF, *N,N*-dimethylformamide. Thin-layer chromatography (TLC) was performed on a silica gel 60 F_{254} precoated plate purchased from Merck. NMR spectra were recorded on a JEOL JNM-A 500 (^{1}H spectra at 500 MHz; ^{13}C spectra at 125 MHz) magnetic resonance spectrometer, and tetramethylsilane was used as the internal standard. Proton NMR spectra were recorded in parts per million (ppm) downfield relative to a tertramethylsilane. The following abbreviations apply to spin multiplicity:

(13) Wu, H. M.; Crothers, D. M. *Nature* **1984**, *308*, 509.

(14) (a) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1996**, *382*, 559. (b) White, S.; Szewczyk, J. W.; Turner, J. M.; Baird, E. E.; Dervan, P. B. *Nature* **1998**, *391*, 468. (c) Trauger, J. W.; Baird, E. E.; Dervan, P. B. *J. Am Chem. Soc.* **1998**, *120*, 3534. (d) Turner, J. M.; Swalley, S. E.; Baird, E. E.; Dervan, P. B. *J. Am Chem. Soc*. **1998**, *120*, 6219.

(15) Gottesfeld, J. M.; Neely, L.; Trauger, J. W.; Baird, E. E.; Dervan, P. B. *Nature* **1997**, *387*, 202.

(16) Dickinson, L. A.; Gulizia, R. J.; Trauger, J. W.; Baird, E. E.; Mosier, D. E.; Gottesfeld, J. M.; Dervan, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12890.

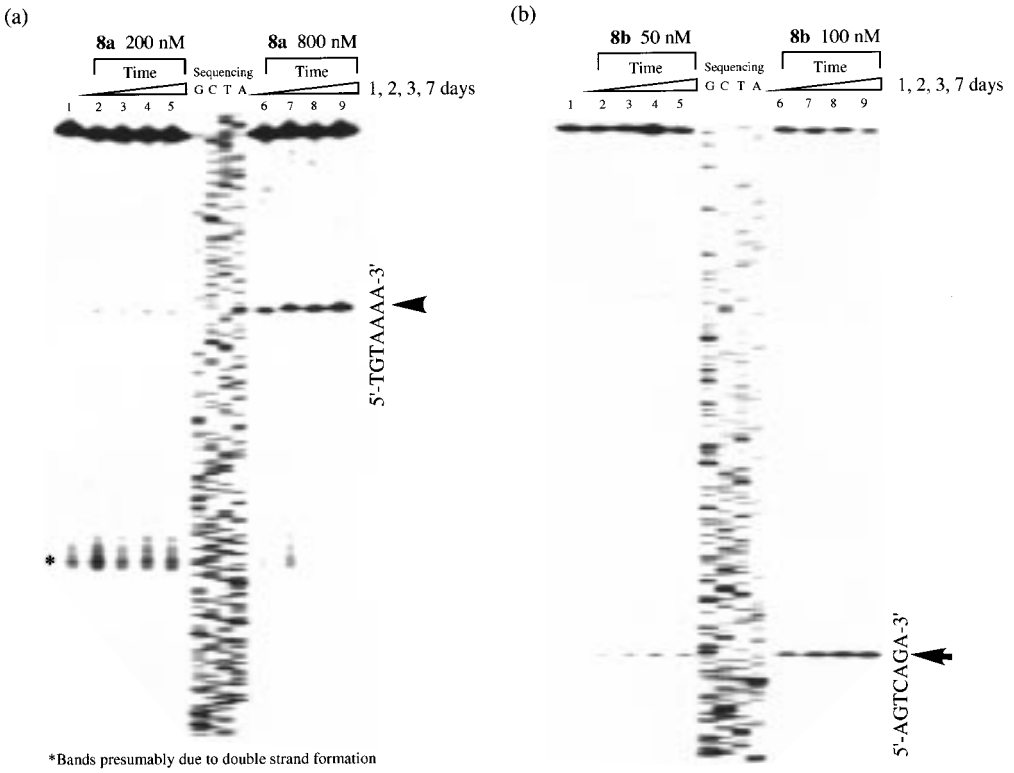


Figure 2. Thermally induced strand cleavage of 5′ TexasRed labeled (a) pUC 18 F 56−481 and (b) pUC 18 F 780−1229 DNA fragments by conjugates **8a** and **8b**, respectively: lane 1, DNA control; lanes 2−5 and 6−9, incubation period of 1, 3, 5, 7 days, respectively. The concentrations of the drugs are indicated in each lane.

s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet), and b (broad). Carbon-13 NMR spectra were recorded with complete heterodecoupling. Electron impact (EI) mass spectra were recorded on a JNM-AX 505 mass spectrometer and electrospray mass spectra (ESMS) on a PE SCIEX API 165 mass spectrometer. Ex Taq DNA polymerase and Filter tube (Suprec-02) were purchased from Takara Co. The Thermo sequenase core sequencing kit and the loading dye (DMF with fushin red) were purchased from Amersham Co., Ltd. 5′ end TexasRed modified DNA primer (18 mer) was purchased from Kurabo Co., Ltd. and 50% Long Ranger gel solution from FMC Bioproducts.

Molecular Modeling Studies. Minimizations were performed with the Discover (MSI, San Diego, CA) program by using CFF force-field parameters. The starting structure was built based on the NMR structure of the ImPyPyγPyPyPy-d(CGCTAACAGGC)/d(GCCTGTTAGCG) complex[3c] and Duo-Dist-octamer complex.[8] The connecting parts between them were built using standard bond lengths and angles. The Du moiety of the assembled initial structure was energy minimized using a distance-dependent dielectric constant of $\epsilon = 4r$ (r stands for the distance between atoms i and j) and the convergence criteria with the root-mean-square gradient of less than 0.001 kcal/mol Å. Twenty Na cations were placed at bifucating positions of the O−P−O angle at a distance of 2.51 Å from the phosphorus atom. The resulting complex was soaked in a 10 Å layer of water. The whole system was minimized without any constraint to the stage where the root-mean-square was less than 0.001 kcal/(mol Å).

Methyl 4-[[4-[[4-(Acetylamino)-1-methylpyrrol-2-yl]carbonylamino]-1-methylpyrrol-2-yl]carbonylamino]-1-methylpyrrole-2-carboxylate (2a). To a solution of 4-acetylamino-1-methylpyrrole-2-carboxylic acid (915 mg, 5.03 mmol) in 15 mL of DMF was added HOBt (744 mg, 5.51 mmol) followed by DCC (1.14 g, 5.52 mmol). The solution was stirred for 60 h and the DCU removed by filtration. Separately, to a solution of **1** (1.50 g, 4.90 mmol) in 30 mL of DMF was added Pd/C catalyst (10%, 1.1 g), and the mixture was stirred at 50 °C under a slight positive pressure of H_2 for 3 d. The catalyst was removed by filtration through Celite and the filtrate immediately added to the OBt ester solution. The reaction mixture was stirred for 36 h and concentrated to a residue under reduced pressure. After purification by flash chromatograph using a mixture of methanol and dichloromathane as eluent, 1.68 g of **2a** was obtained in 78% yield. ^{1}H NMR (DMSO-d_6) δ 9.90 (s, 1H), 9.88 (s, 1H), 9.79 (s, 1H), 7.45 (d, $J = 1.8$ Hz, 1H), 7.22 (d, $J = 1.8$ Hz, 1H), 7.14 (d, $J = 1.8$ Hz, 1H), 7.04 (d, $J = 1.8$ Hz, 1H), 6.90 (d, $J = 1.8$ Hz, 1H), 6.84 (d, $J = 1.8$ Hz, 1H), 3.834 (s, 3H), 3.829 (s, 3H), 3.82 (s, 3H), 3.73 (s, 3H), 1.96 (s, 3H); EIMS *m*/*e* (%) 440 (100), 287 (91.7); HREIMS *m*/*e* calcd for $C_{21}H_{24}N_6O_5$ 440.1808, found 440.1803.

Methyl 4-[[4-[[4-(Acetylamino)-1-methylimidazol-2-yl]carbonylamino]-1-methylpyrrole-2-yl]carbonylamino]-1-methylpyrrole-2-carboxylate (2b). A similar synthetic procedure as that for **2a** was followed for preparation of **2b**, yield 80%. ^{1}H NMR (DMSO-d_6) δ 10.22 (s, 1H), 9.96 (s, 1H), 9.92 (s, 1H), 7.46 (d, $J = 1.8$ Hz, 1H), 7.41 (s, 1H), 7.26 (d, $J = 1.8$ Hz, 1H), 7.14 (d, $J = 1.8$ Hz, 1H), 6.90 (d, $J = 1.8$ Hz, 1H), 3.94 (s, 3H), 3.84 (s, 3H), 3.80 (s, 3H), 3.73 (s, 3H), 2.02 (s, 3H); EIMS *m*/*e* (%) 441 (100), 288 (77.6). HREIMS *m*/*e* calcd for $C_{20}H_{23}N_7O_5$ 441.1760, found 441.1761.

4-[[4-[[4-(Acetylamino)-1-methylpyrrol-2-yl]carbonylamino]-1-methylpyrrol-2-yl]carbonylamino]-1-methylpyrrole-2-carboxylic Acid (3a). To a solution of **2a** (997 mg, 2.27 mmol) in 15 mL of methanol was added 15 mL of 2 N aqueous NaOH. The mixture was stirred at room temperature for 40 h. After evaporation of methanol, the aqueous solution was cooled to 4 °C and acidified to pH 2 with diluted hydrochloric acid. The precipitate was collected by filtration, washed with water, and dried to give 917 mg of **3a** in 95% yield. ^{1}H NMR (DMSO-d_6) δ 10.00 (s, 1H), 9.95 (s, 1H), 9.92 (s, 1H), 7.41 (d, $J = 1.8$ Hz, 1H), 7.23 (d, $J = 1.8$ Hz, 1H), 7.14 (d, $J = 1.8$ Hz, 1H), 7.07 (d, $J = 1.8$ Hz, 1H), 6.89 (d, $J = 1.8$ Hz, 1H), 6.86 (d, $J = 1.8$ Hz, 1H), 3.82 (s, 3H), 3.81 (s, 3H), 3.80 (s, 3H), 1.96 (s, 3H); ^{13}C NMR (DMSO-d_6) δ 166.55, 161.93, 158.48, 122.71, 122.68, 122.56, 122.22, 122.13, 120.26, 119.50, 118.58, 118.12, 108.46, 104.89, 104.02. ESMS *m*/*e* calcd for $C_{20}H_{21}N_6O_5$ (M−H) 425.2, found 425.1.

4-[[4-[[4-(Acetylamino)-1-methylimidazol-2-yl]carbonylamino]-1-methylpyrrol-2-yl]carbonylamino]-1-methylpyrrole-2-carboxylic acid (3b). A similar synthetic procedure as that for **3a** was followed for preparation of **3b**, yield 97%. ^{1}H NMR (DMSO-d_6) δ 10.25 (s, 1H), 9.94 (s, 2H), 7.42 (s, 2H), 7.26 (s,1H), 7.14 (s, 1H), 6.85 (s, 1H), 3.94 (s, 3H), 3.84 (s, 3H), 3.81 (s, 3H), 2.02 (s, 3H); ESMS calcd for $C_{19}H_{21}N_7O_5$ (M−H) 426.2, found 426.2.

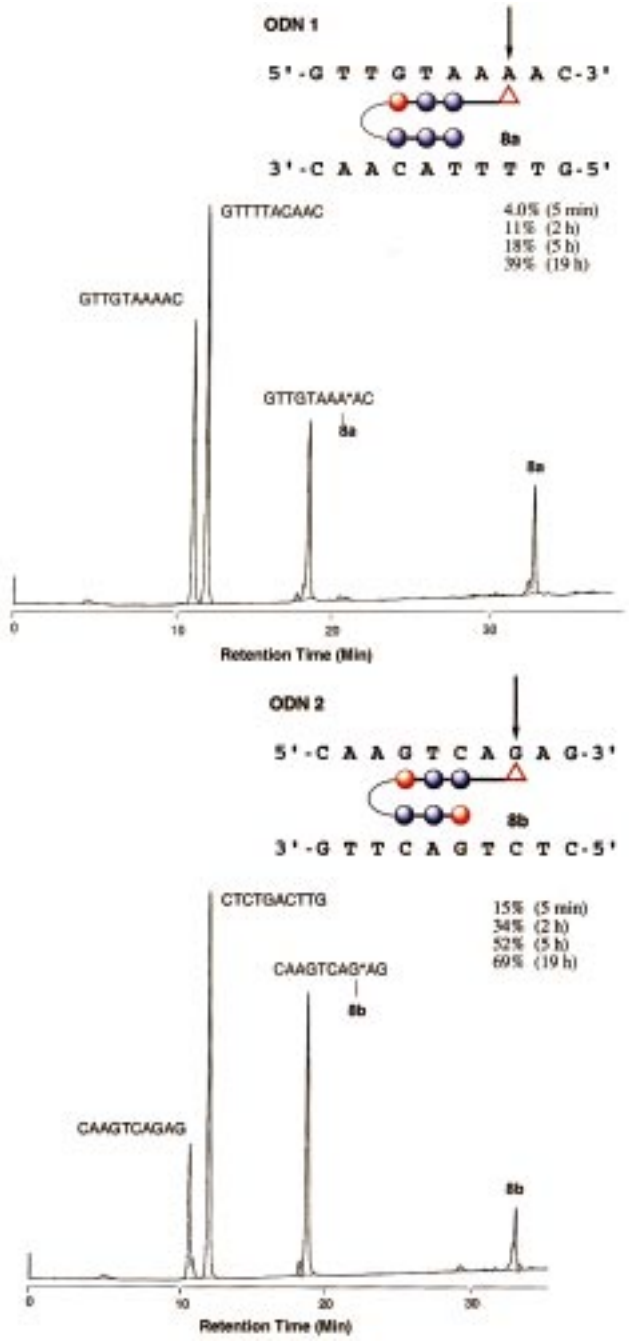


Figure 3. HPLC profile of (a) ODN1 by **8a** and (b) ODN2 by **8b**. A schematic representation of alkylation of the ODNs by conjugates **8a** and **8b** is also included in the figure. The arrows represent the location and the extent of alkylation after the indicated incubation periods. The imidazole and pyrrole rings are represented by red and blue circles, respectively; the curved line connecting the polyamide subunits represents γ-aminobutyric acid, and the red triangle represents Du.

Methyl 4-[[4-[[4-(Acetylamino)-1-methylimidazol-2-yl]carbonyl-amino]-1-methylpyrrol-2-yl]carbonylamino]-1-methylpyrrole-2-carboxylate (4). To a solution of Boc-protected precursor of **4** (1.120 g, 2.31 mmol) in 20 mL of CH_2Cl_2 was added 4.9 mL of trifluroacetic acid at nitrogen atmosphere. The mixture was stirred for 3 h. After removal of the solvent and purification by flash chromatography using a mixture of CH_2Cl_2 and CH_3OH as eluent, **4** was obtained as light yellow solid in 92% yield. ^{1}H NMR (DMSO-d_6) δ 10.32 (s, 1H), 9.94 (s, 2H), 7.80 (s, br, 3H), 7.46 (d, J = 1.8 Hz, 1H), 7.43 (s, 1H), 7.26 (d, J = 1.8 Hz, 1H), 7.14 (d, J = 1.8 Hz, 1H), 6.90 (d, J = 1.8 Hz, 1H), 3.95 (s, 3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.73 (s, 3H), 2.82 (m, 2H), 2.41 (t, J = 7.3 Hz, 2H), 1.83 (m, 2H). ESMS *m*/*e* calcd for $C_{22}H_{27}N_8O_5$ (M−H) 483.2, found 483.3.

PyPyPyγImPyPyCOOCH$_3$ (5a). To a solution of **3a** (276 mg, 0.65 mmol) in 6 mL of DMF was added HOBt (96.2 mg, 0.71 mmol) followed by DCC (147 mg, 0.71 mmol). The solution was stirred overnight and the DCU removed by filtration. To the OBt ester solution was added **4** (400 mg, 0.67 mmol). The mixture was stirred for 3 h and concentrated to a residue, which was subjected to flash chromatography (eluent: CH_2Cl_2:CH_3OH = 10:1) to afford **5a** in 66% yield. ^{1}H NMR (DMSO-d_6) δ 10.24 (s, 1H), 9.96 (s, 1H), 9.92 (s, 1H), 9.872 (s, 1H), 9.866 (s, 1H), 9.79 (s, 1H), 8.00 (t, J = 5.5 Hz, 1H), 7.45 (d, J = 1.8 Hz, 2H), 7.25 (d, J = 1.8 Hz, 1H), 7.22 (s, 1H), 7.16 (d, J = 1.8 Hz, 1H), 7.13 (d, J = 1.8 Hz, 1H), 7.03 (s, J = 1.8 Hz, 1H), 6.90 (d, J = 1.8 Hz, 1H), 6.89 (d, J = 1.8 Hz, 1H), 6.85 (d, J = 1.8 Hz, 1H), 3.94 (s, 3H), 3.84 (s, 6H), 3.83 (s, 3H), 3.82 (s, 3H), 3.79 (s, 3H), 3.73 (s, 3H), 3.18 (m, 2H), 2.36 (t, J = 7.3 Hz, 2H), 1.96 (s, 3H), 1.79 (m, 2H). ESMS *m*/*e* calcd for $C_{42}H_{47}N_{14}O_9$ (M−H) 891.4, found 891.5.

ImPyPyγImPyPyCOOCH$_3$ (5b). A similar synthetic procedure as that for **5a** was followed for preparation of **5b**, yield 72%. ^{1}H NMR (DMSO-d_6) δ 10.24 (s, 1H), 10.22 (s, 1H), 9.96 (s, 1H), 9.93 (s, 1H), 9.89 (s, 1H), 9.88 (s, 1H), 8.01 (t, J = 5.5 Hz, 1H), 7.46 (d, J = 1.8 Hz, 1H), 7.45 (s, 1H), 7.42 (s, 1H), 7.26 (s, 2H), 7.17 (d, J = 1.8 Hz, 1H), 7.14 (s, J = 1.8 Hz, 1H), 7.12 (d, J = 1.8 Hz, 1H), 6.90 (d, J = 1.8 Hz, 1H), 6.89 (d, J = 1.8 Hz, 1H), 3.943(s, 3H), 3.939 (s, 3H), 3.844 (s, 3H), 3.838 (s, 3H), 3.83 (s, 3H), 3.79 (s, 3H), 3.73 (s, 3H), 3.19 (m, 2H), 2.36 (t, J = 7.4 Hz, 2H), 2.01 (s, 3H), 1.78 (m, 2H); ESMS *m*/*e* calcd for $C_{41}H_{46}N_{15}O_9$ (M−H) 892.4, found 892.4.

PyPyPyγImPyPyCOOH (6a). To a suspension of **5a** (220 mg, 0.25 mmol) in 90 mL of methanol was added 1 N aqueous NaOH (4 mL). The mixture was stirred at 55 °C for 3.5 days. After removal of CH_3-OH and addition of H_2O (50 mL), the aqueous solution was acidified to pH 2. The precipitate was collected by filtration, washed with water, and dried to afford 150 mg of **6a** in 69% yield. ^{1}H NMR (DMSO-d_6) δ 10.24 (s, 1H), 9.95 (s, 1H), 9.89 (s, 1H), 9.873 (s, 1H), 9.868 (s, 1H), 9.80 (s, 1H), 8.01 (t, J = 5.5 Hz, 1H), 7.45 (s, H), 7.41 (d, J = 1.8 Hz, 1H), 7.25 (d, J = 1.8 Hz, 1H), 7.22 (d, J = 1.8 Hz, 1H), 7.16 (d, J = 1.8 Hz, 1H), 7.13 (s, 2H), 7.03 (d, J = 1.8 Hz, 1H), 6.89 (d, J = 1.8 Hz, 1H), 6.84 (t, J = 3.0 and 1.8 Hz, 2H), 3.94 (s, 3H), 3.84 (s, 6H), 3.82 (s, 3H), 3.81 (s, 3H), 3.79 (s, 3H), 3.20 (m, 2H), 2.35 (t, J = 7.3 Hz, 2H), 1.96 (s, 3H), 1.78 (m, 2H). ESMS *m*/*e* calcd for $C_{41}H_{45}N_{14}O_9$ (M−H) 877.3, found 877.4.

ImPyPyγImPyPyCOOH (6b). A similar synthetic procedure as that for **6a** was followed for preparation of **6b**, yield 81%.^{1}H NMR (DMSO-d_6) δ 10.24 (s, 1H),10.22 (s, 1H), 9.96 (s, 1H), 9.93 (s, 1H), 9.89 (s, 1H), 9.88 (s, 1H), 8.01 (t, J = 5.5 Hz, 1H), 7.45 (s, 1H), 7.41 (s, 2H), 7.26 (t, J = 1.8 Hz, 2H), 7.16 (d, J = 1.2 Hz, 1H), 7.13 (d, J = 1.8 Hz, 1H), 7.12 (d, J = 1.8 Hz, 1H), 6.89 (d, J = 1.8 Hz, 1H), 6.84 (d, J = 2.4 Hz, 1H), 3.941(s, 3H), 3.939 (s, 3H), 3.844 (s, 3H), 3.838 (s, 3H), 3.81 (s, 3H), 3.79 (s, 3H), 3.19 (m, 2H), 2.36 (t, J = 7.4 Hz, 2H), 2.01 (s, 3H), 1.78 (m, 2H). ESMS *m*/*e* calcd for $C_{40}H_{44}N_{15}O_9$ (M−H) 878.3, found 878.4.

PyPyPyγImPyPyCOIm (7a). To a solution of **6a** (47 mg, 0.053 mmol) in 2 mL of DMF was added 1,1′-carbonyldiimidazole (17 mg, 0.105 mmol). The reaction mixture was stirred at room temperature for 9 h, then 10 mL of ice water was added. The precipitate was collected by filtration, washed with water, and dried to afford 25 mg of **7a** in 53% yield. ^{1}H NMR (DMSO-d_6) δ 10.25 (s, 1H), 10.06 (s, 1H), 10.01 (s, 1H), 9.89 (s, 1H), 9.88 (s, 1H), 9.81 (s, 1H), 8.26 (s, 1H), 8.02 (t, J = 5.5 Hz, 1H), 7.78 (s, br, H), 7.70 (s, br, 1H), 7.46 (s, 1H), 7.27 (d, J = 2.0 Hz, 1H), 7.23 (d, J = 1.2 Hz, 1H), 7.21 (d, J = 2.0 Hz, 1H), 7.17 (d, J = 1.5 Hz, 1H), 7.14 (d, J = 2.0 Hz, 2H), 7.05 (d, J = 1.5 Hz, 2H), 6.96 (d, J = 1.5 Hz, 1H), 6.90 (s, br, 1H), 6.86 (d, J = 2.0 Hz, 1H), 3.95 (s, 3H), 3.92 (s, 3H), 3.86 (s, 3H), 3.85 (s, 3H), 3.83 (s, 3H), 3.80 (s, 3H), 3.21 (m, 2H), 2.36 (t, J = 7.3 Hz, 2H), 1.97 (s, 3H), 1.79 (m, 2H). ESMS *m*/*e* calcd for $C_{44}H_{47}N_{16}O_8$ (M−H) 927.4, found 927.5.

ImPyPyγImPyPyCOIm (7b). A similar synthetic procedure of **7a** was followed for preparation of **7b**, yield 50%. ^{1}H NMR (DMSO-d_6) δ 10.23 (s, 1H), 10.21 (s, 1H), 10.04 (s, 1H), 10.00 (s, 1H), 9.92 (s, 1H), 9.87 (s, 1H), 8.25 (s, 1H), 8.00 (t, J = 5.5 Hz, 1H), 7.77 (d, J = 2.0 Hz, 1H), 7.69 (s, 1H), 7.45 (s, 1H), 7.41 (s, 1H), 7.26 (d, J = 2.0 Hz, 1H), 7.25 (d, J = 2.0 Hz, 1H), 7.19 (d, J = 2.0 Hz, 1H), 7.16 (d, J = 2.0 Hz, 1H), 7.12 (t, J = 2.0 Hz, 2H), 6.95 (d, J = 2.0 Hz, 1H),

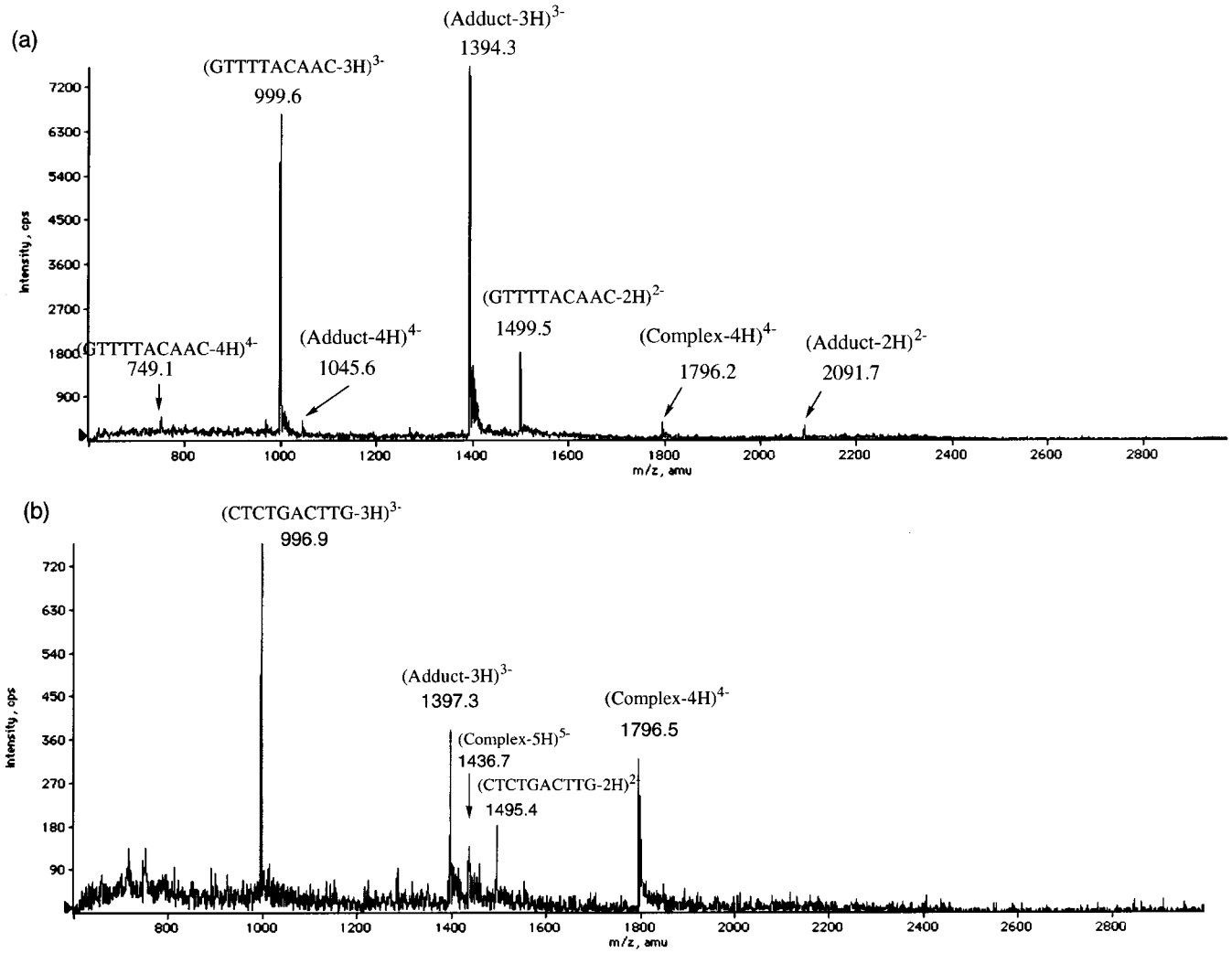


Figure 4. ESMS spectrum of (a) **8a**−ODN1 and (b) **8b**−ODN2 complexes. The *y*-axis indicates the relative peak intensity.

6.89 (d, J = 2.0 Hz, 1H), 3.94 (s, 6H), 3.90 (s, 3H), 3.85 (s, 3H), 3.84 (s, 3H), 3.79 (s, 3H), 3.20 (m, 3H), 2.35 (t, J = 7.0 Hz, 2H), 2.01 (s, 3H), 1.79 (m, 2H). ESMS *m*/*e* calcd for $C_{43}H_{46}N_{15}O_9$ (M−H) 928.4, found 928.5.

PyPyPy*γ*ImPyPyDu (8a). Under nitrogen atmosphere, to a mixture of 2 mg (0.087 mmol) of 60% NaH in 0.3 mL of DMF was added dropwise 7.5 mg (0.027 mmol) of segment A of duocarmycin A in 0.3 mL of DMF at −60 °C. The mixture was stirred at −60 to −30 °C for 2 h and then cooled to −50 °C; 10 mg (0.011 mmol) of **7a** in 0.3 mL of DMF was injected dropwise. The reaction mixture was stirred at −50 to −30 °C for 1 h. The reaction mixture was directly subject to TLC using a mixture of dichloromethane and methanol as eluent to give 4.9 mg of **8a** in 40% yield. ^{1}H NMR (DMSO-d_6) δ 10.24 (s, 1H), 9.96 (s, 1H), 9.93 (s, 1H), 9.88 (s, 1H), 9.87 (s, 1H), 9.80 (s, 1H), 8.69 (s, 1H), 8.01 (t, J = 5.5 Hz, 1H), 7.49 (d, J = 1.8 Hz, 1H), 7.45 (s, 1H), 7.25 (d, J = 1.8 Hz, 1H), 7.22 (d, J = 1.2 Hz, 1H), 7.16 (d, J = 1.8 Hz, 2H), 7.13 (d, J = 1.2 Hz, 1H), 7.03 (d, J = 1.8 Hz, 1H), 6.89 (d, J = 1.8 Hz, 2H), 6.85 (d, J = 1.8 Hz, 1H), 6.74 (d, J = 1.2 Hz, 1H), 6.36 (s, 1H), 4.28 (dd, J = 4.9 Hz, 1H), 4.08 (d, J = 11.0 Hz, 1H), 3.93 (s, 3H), 3.84 (s, 3H), 3.83 (s, 3H), 3.82 (s, 3H), 3.79 (s, 3H), 3.75 (s, 3H), 3.59 (s, 3H), 3.19 (m, 2H), 2.93 (m, 1H), 2.35 (m, 2H), 1.98 (m, 1H), 1.96 (s, 3H), 1.78 (m, 2H), 1.44 (s, 3H), 1.40 (t, J = 4.3 and 3.7 Hz, 1H). ESMS *m*/*e* calcd for $C_{55}H_{57}N_{16}O_{12}$ (M−H) 1133.4, found 1133.5.

ImPyPy*γ*ImPyPyDu (8b). A similar synthetic procedure as that for **8a** was followed for preparation of **8b**, yield 41%.^{1}H NMR (DMSO-d_6) δ 10.24 (s, 1H), 10.21 (s, 1H), 9.95 (s, 1H), 9.93 (s, 2H), 9.87 (s, 1H), 8.68 (s, 1H), 8.00 (t, J = 5.5 Hz, 1H), 7.49 (d, J = 2.0 Hz, 1H), 7.41 (s, 1H), 7.25 (d, J = 2.0 Hz, 1H), 7.24 (d, J = 2.0 Hz, 1H), 7.16 (d, J = 2.0 Hz, 1H), 7.15 (d, J = 2.0 Hz, 1H), 7.12 (d, J = 2.0 Hz, 1H), 6.88 (d, J = 2.0 Hz, 1H), 6.74 (d, J = 2.0 Hz, 1H), 6.36 (s, 1H), 4.28 (dd, J = 4.0 and 4.5 Hz, 1H), 4.08 (d, J = 11.0 Hz, 1H), 3.94 (s, 6H), 3.84 (s, 6H), 3.79 (s, 3H), 3.75 (s, 3H), 3.59 (s, 3H), 3.19 (m, 2H), 2.93 (m, 1H), 2.35 (m, 2H), 2.01 (s, 3H), 1.98 (dd, J = 3.0 Hz, 1H), 1.78 (m, 2H), 1.44 (s, 3H), 1.39 (t, J = 4.5 and 3.5 Hz, 1H). ESMS *m*/*e* calcd for $C_{54}H_{56}N_{17}O_{12}$ (M−H) 1134.4, found 1134.5.

Preparation of 5′-TexasRed-End-Modified 426- and 450-Base Pair DNA Fragments. 5′-TexasRed-labeled 450-base pair fragment (pUC 18 forward 780−1229) was prepared by PCR using 5′-TexasRed modified 5′-AGAATCAGGGGATAACGCAG-3′ (pUC 18 forward 780−799) and 5′-TTACCAGTGGCTGCTGCCAG-3′ (pUC 18 reverse 1469−1488) as primers.[17] The asterisk indicates Texas Red modification and the nucleotide numbering starts with the replication site. 5′-TexasRed-labeled 426-base pair fragment (pUC 18 forward 56−481) was prepared by PCR using 5′-TexasRed modified 5′-GGTCACAGCT-TGTCTGTAAG-3′ (pUC 18 forward 56−75) and 5′-CAGGAAA-CAGCTATGAC-3′ (pUC 18 reverse 2207−2223) as primers. DNA fragments were purified by filtration using Suprec-02, and the concentration was determined by ethidium bromide staining.

Alkylation of 450- and 426-Base Pair DNA Fragments by Conjugate 8a or 8b. 5′-TexasRed labeled DNA fragment (60 nM) was alkylated in 10 μL of 10 mM Na phosphate buffer (pH 7.0) containing 10% DMF at room temperature for the indicated period. The reaction was quenched by addition of calf thymus DNA (5 mM, 1 μL) and heated for 5 min at 90 °C. The DNA was collected by ethanol precipitation. The pellet was resolved in 8 μL of loading dye

(17) Liang, L.; Pardee, A. B. *Current protocols in molecular biology*; Ausubel, F. M., Brent, R., Kinstong, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., Struhl, K., Eds.; John Wiley & Sons: New York, 1996; Chapter 15.

(formamide with fushin red) and heated at 94 °C for 20 min and was immediately cooled to 0 °C to denature DNA. A 2 μL aliquot was electrophoresed on a 6% denaturing Long Ranger polyacrylamide gel using a Hitachi 5500-S DNA sequencer system.

Alkylation of Deoxydecanucleotides by Conjugate 8a or 8b As Monitored by HPLC. Deoxydecanucleotides were synthesized on an automated DNA synthesizer. The alkylation of DNA oligonucleotides was carried out according to the following procedure. A reaction mixture (50 μL) containing conjugates **8a** and **8b** (each 0.1 mM) and the duplex DNA decamer (1 mM base concentration) in 50 mM sodium cacodylate buffer (pH 7.0) was incubated at 0 °C for indicated periods. The progress of the reaction was monitored by HPLC using a Chemcobond 5-ODS−H column (4.6 × 150 mm). Elution was performed with 0.05 M ammonium formate and a 0−50% acetonitrile linear gradient (0−40 min) at a flow rate of 1.0 mL/min. Detection was at 254 nm. The production of conjugate-guanine and -adenine adducts and the abasic site-containing nucleotides upon heating (90 °C for 5 min) was confirmed according to the previous reported procedures.[6e] HPLC profiles of the reaction mixture were shown in Figure 3.

ESMS Detection of 8a−ODN1 and 8b−ODN2 Complexes. A reaction mixture (50 μL) containing a new compound (0.1 mM) and the DNA decamer (1 mM base concentration) in 50 mM sodium cacodylate buffer (pH 7.0) was incubated at 0 °C for 2 days and then subjected to HPLC as above. The covalent adducts GTTGTAAAAC−**8a** and CAAGTCAGAG−**8b** were collected, respectively, then their complementary oligonucleotide strands were added to the corresponding covalent adduct solution. The resulting reconstruction complex solutions were immediately subject to electrospray mass spectrometry (ESMS) are shown in Figure 4.

Acknowledgment. We thank Dr. David E. Wemmer for providing the coordinate of the NMR refined structure of the ImPyPyγPyPyPy-d(CGCTAACAGGC)/d(GCCTGTTAGCG) complex. The authors also thank Dr. Peter B. Dervan for providing the detailed experimental procedure for the synthesis of the compound **4**.

JA983398W