J|A|C|S
ARTICLES

Published on Web 06/11/2008

Artificial DNA Made Exclusively of Nonnatural C-Nucleosides with Four Types of Nonnatural Bases

Yasuhiro Doi, Junya Chiba,* Tomoyuki Morikawa, and Masahiko Inouye*

Graduate School of Pharmaceutical Sciences, University of Toyama, Sugitani 2630, Toyama 930-0194, Japan

Received February 12, 2008; E-mail: chiba@pha.u-toyama.ac.jp; inouye@pha.u-toyama.ac.jp

Abstract: We describe a new class of DNA-like oligomers made exclusively of nonnatural, stable C-nucleosides. The nucleosides comprise four types of nonnatural bases attached to a deoxyribose through an acetylene bond with the β-configuration. The artificial DNA forms right-handed duplexes and triplexes with the complementary artificial DNA. The hybridization occurs spontaneously and sequence-selectively, and the resulting duplexes have thermal stabilities very close to those of natural duplexes. The artificial DNA might be applied to a future extracellular genetic system with information storage and amplifiable abilities.

Introduction

Apart from its biological significance, the beauty of the DNA duplex structure has been inspiring chemists to create artificial analogues in their own right. This type of research is primarily motivated by pure scientific exploration and eventually directed toward biomedical applications. Creation of the artificial analogues is roughly categorized into the following two modifications of natural DNA: the sugar−phosphodiester backbone and the Watson−Crick base pairing. The former is represented by peptide nucleic acids (PNA),[1] locked nucleic acids (LNA),[2] *threo*-furanosyl nucleic acids (TNA),[3] and 2′,5′-linked nucleic acids.[4] These analogues usually possess natural bases because of their potential use in antisense and antigene strategies. On the other hand, examples of the latter are properly aimed at their applications to nanotechnology,[5] biotechnology,[6] and chemical biology.[7] Until recently, however, those analogues have contained only one or a few nonnatural bases linked with deoxyribose. Eschenmoser and co-workers[8] reported notable exceptions for entirely modified oligonucleotide analogues composed of nonnatural bases and various sugars. Furthermore, Kool and co-workers[9] disclosed artificial duplexes in which all of the natural base pairs are replaced by pairs of nonnatural

(1) For a recent review, see Nielsen, P. E *Chem. Biodiversity* **2007**, *4*, 1996–2002.

(2) For a recent review. see Wengel, J.; Petersen, M.; Frieden, M.; Koch, T. *Lett. Pept. Sci.* **2004**, *10*, 237–253.

(3) (a) Heuberger, B. D.; Switzer, C. *Org. Lett.* **2006**, *8*, 5809–5811. (b) Horhota, A.; Zou, K.; Ichida, J. K.; Yu, B.; McLaughlin, L. W.; Szostak, J. W.; Chaput, J. C. *J. Am. Chem. Soc.* **2005**, *127*, 7427–7434. (c) Ichida, J. K.; Zou, K.; Horhota, A.; Yu, B.; McLaughlin, L. W.; Szostak, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 2802–2803. (d) Ichida, J. K.; Horhota, A.; Zou, K.; McLaughlin, L. W.; Szostak, J. W. *Nucleic Acids Res.* **2005**, *33*, 5219–5225. (e) Zou, K.; Horhota, A.; Yu, B.; Szostak, J. W.; McLaughlin, L. W. *Org. Lett.* **2005**, *7*, 1485–1487. (f) Eschenmoser, A. *Origin Life Evol. Biosphere* **2004**, *34*, 277–306. (g) Ferencic, M.; Reddy, G.; Wu, X.; Guntha, S.; Nandy, J.; Krishnamurthy, R.; Eschenmoser, A *Chem. Biodiversity* **2004**, *1*, 939–979, and references therein.

(4) (a) Horowitz, E. D.; Hud, N. V. *J. Am. Chem. Soc.* **2006**, *128*, 15380–15381. (b) Prakash, T. P.; Kraynack, B.; Baker, B. F.; Swayze, E. E.; Bhat, B. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 3238–3240. (c) Plevnik, M.; Gdaniec, Z.; Plavec, J. *Nucleic Acids Res.* **2005**, *33*, 1749–1759. (d) Sinha, S.; Kim, P. H.; Switzer, C. *J. Am. Chem. Soc.* **2004**, *126*, 40–41. (e) Denisov, A. Y.; Hannoush, R. N.; Gehring, K.; Damha, M. J. *J. Am. Chem. Soc.* **2003**, *125*, 11525–11531, and references therein.

(5) For recent reports: (a) Heuberger, B. D.; Shin, D.; Switzer, C *Org. Lett.* **2008**, *10*, 1091–1094. (b) Clever, G. H.; Kaul, C.; Carell, T. *Angew. Chem., Int. Ed.* **2007**, *46*, 6226–6236. (c) Polonius, F.-A.; Müller; J., *Angew. Chem., Int. Ed.* **2007**, *46*, 5602–5604. (d) Shin, D.; Switzer, C. *Chem. Commun.* **2007**, 4401–4403. (e) Clever, G. H.; Carell, T. *Angew. Chem., Int. Ed.* **2007**, *46*, 250–253. (f) Tanaka, Y.; Oda, S.; Yamaguchi, H.; Kondo, Y.; Kojima, C.; Ono, A. *J. Am. Chem. Soc.* **2007**, *129*, 244–245. (g) Tanaka, K.; Clever, G. H.; Takezawa, Y.; Yamada, Y.; Kaul, C.; Shionoya, M.; Carell, T. *Nat. Nanotechnol.* **2006**, *1*, 190–194. (h) Miyake, Y.; Togashi, H.; Tashiro, M.; Yamaguchi, H.; Oda, S.; Kudo, M.; Tanaka, Y.; Kondo, Y.; Sawa, R.; Fujimoto, T.; Machinami, T.; Ono, A. *J. Am. Chem. Soc.* **2006**, *128*, 2172–2173. (i) Clever, G. H.; Polborn, K.; Carell, T. *Angew. Chem., Int. Ed.* **2005**, *44*, 7204–7208. (j) Switzer, C.; Sinha, S.; Kim, P. H.; Heuberger, B. D. *Angew. Chem., Int. Ed.* **2005**, *44*, 1529–1532. (k) Switzer, C.; Shin, D. *Chem. Commun.* **2005**, 1342–1344. (l) Ono, A.; Togashi, H. *Angew. Chem., Int. Ed.* **2004**, *43*, 4300–4302. (m) Zimmermann, N.; Meggers, E.; Schultz, P. G. *Bioorg. Chem.* **2004**, *32*, 13–25. (n) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shionoya, M. *Science* **2003**, *299*, 1212–1213.

(6) For recent reports: (a) Leconte, A. M.; Hwang, G. T.; Matsuda, S.; Capek, P.; Hari, Y.; Romesberg, F. E. *J. Am. Chem. Soc.* **2008**, *130*, 2336–2343. (b) Hirao, I.; Mitsui, T.; Kimoto, M.; Yokoyama, S. *J. Am. Chem. Soc.* **2007**, *129*, 15549–15555. (c) Matsuda, S.; Leconte, A. M.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 5551–5557. (d) Yang, Z.; Sismour, A. M.; Sheng, P.; Puskar, N. L.; Benner, S. A. *Nucleic Acids Res.* **2007**, *35*, 4238–4249. (e) Yang, Z.; Sismour, A. M.; Benner, S. A. *Nucleic Acids Res.* **2007**, *35*, 3118–3127. (f) Meena; Sam, M.; Pierce, K.; Szostak, J. W.; McLaughlin, L. W *Org. Lett.* **2007**, *9*, 1161–1163. (g) Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. *Nat. Methods* **2006**, *3*, 729–735. (h) Imoto, S.; Patro, J. N.; Jiang, Y. L.; Oka, N.; Greenberg, M. M. *J. Am. Chem. Soc.* **2006**, *128*, 14606–14611. (i) Matsuda, S.; Henry, A. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 6369–6375. (j) Hirao, I. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622–627. (k) Kim, Y.; Leconte, A. M.; Hari, Y.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7809–7812. (l) Leconte, A. M.; Matsuda, S.; Hwang, G. T.; Romesberg, F. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 4326–4329. (m) Yang, Z.; Hutter, D.; Sheng, P.; Sismour, A. M.; Benner, S. A. *Nucleic Acids Res.* **2006**, *34*, 6095–6101. (n) Hikishima, S.; Minakawa, N.; Kuramoto, K.; Fujisawa, Y.; Ogawa, M.; Matsuda, A. *Angew. Chem., Int. Ed.* **2005**, *44*, 596–598. (o) Zhang, L.; Peritz, A.; Meggers, E. *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175.

10.1021/ja801058h CCC: $40.75

and natural bases. The question, however, still remains whether structural variations for phosphodiester-based duplexes can be further extended. Here we describe a new class of such base-pairing right-handed duplexes and triplexes, which can be formed only from artificial DNA strands containing only nonnatural nucleoside residues possessing four types of nonnatural bases through nonnatural C-glycoside linkers. The hybridization occurs spontaneously and sequence-selectively, and the resulting duplexes have net thermal stabilities very close to those of natural duplexes in spite of their differences for the geometries and hydrogen-bonding patterns. These findings provide further proof that a genetic phosphodiester-based molecular framework is not limited to the natural DNA one and can be extended to many types of nonbiological entities.

Results and Discussion

The molecular design for a new three-point hydrogen-bonding pair was based on our previous studies of molecular recognition for 2-amino-3*H*-pyrimidin-4-one derivatives.[10] These heterocyclic compounds have the hydrogen-bonding array of donor−donor−acceptor (DDA). Although the skeleton can tautomerize to its 1*H* form, bearing a DAA hydrogen-bonding pattern, our results showed that the skeleton selectively interacts with a cytosine derivative (AAD type) among all kinds of natural bases in less polar organic media. This finding led us to choose this heterocycle as a skeleton for a G base analogue. Its actual structure is depicted in Figure 1a and referred to as G* for the free nonnatural base (similarly, C*, A*, and T* for the remaining three nonnatural bases). The counterpart of G* is C*, which possesses an AAD hydrogen-bonding pattern. The structure of C* is derived from 4-amino-1*H*-pyrimidin-2-one by replacing the hydrogen atom at the N1 position by a methyl group. This substitution inhibits the original skeleton from tautomerizing to its 3*H* form that can ambiguously behave as an ADD hydrogen-bonding module as G* does. The position at the bases connected to deoxyribose was selected to be on carbon and not to significantly alter the natural DNA framework. This decision necessarily results in a reversed hydrogen-bonding pattern for the G*/C* combination relative to that of the natural G/C pair. Therefore, the abbreviations of *i*G* (*iso*G*) and *i*C* (*iso*C*) are used instead of G* and C* when the artificial bases are incorporated into nucleoside units. The structures of another pair of nonnatural bases, A* and T*, were determined in order that the pair has a two-point, DA-type hydrogen-bonding complementarity and has the same geometry as the *i*G*/*i*C* pair. The pattern of hydrogen bonds between A* and T* must be incompatible with that between *i*G* and *i*C* when the A*/T* pair is superimposed on the *i*G*/*i*C* one, being set the glycoside bonds to the same directions.

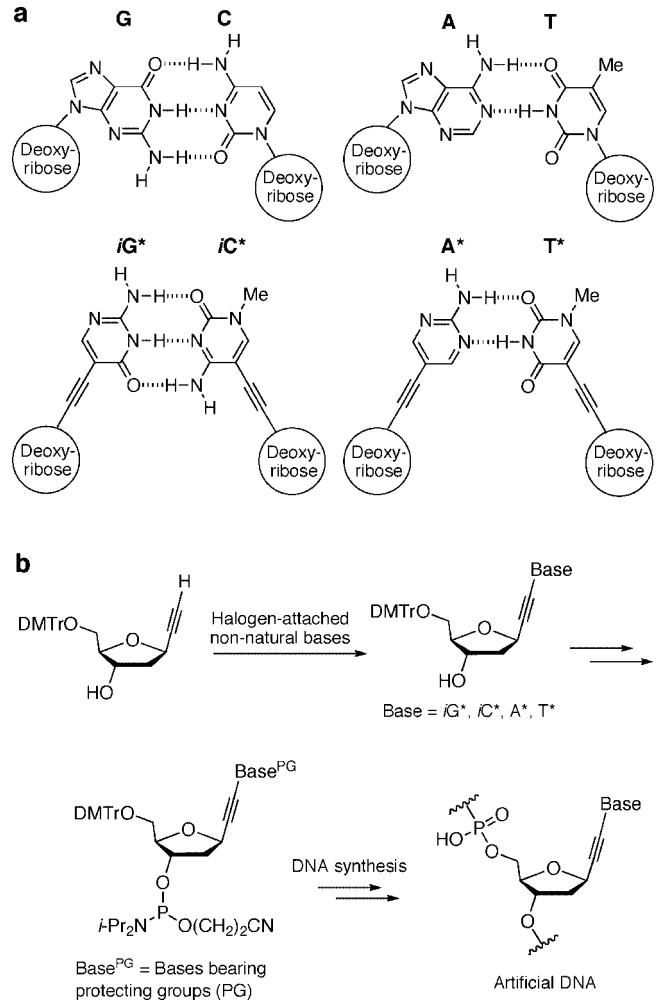

Figure 1. Chemical structures and synthesis of our artificial DNA. (a) Chemical structures of two types of natural and nonnatural base pairs. (b) Synthetic strategy for the artificial DNA.

(7) For recent reports: (a) Heuberger, B. D.; Switzer, C *J. Am. Chem. Soc.* **2008**, *130*, 412–413. (b) Tsai, C.-H.; Chen, J.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 14598–14603. (c) Matsuda, S.; Fillo, J. D.; Henry, A. A.; Rai, P.; Wilkens, S. J.; Dwyer, T. J.; Geierstanger, B. H.; Wemmer, D. E.; Schultz, P. G.; Spraggon, G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2007**, *129*, 10466–10473. (d) Murata, S.; Mizumura, Y.; Hino, K.; Ueno, Y.; Ichikawa, S.; Matsuda, A. *J. Am. Chem. Soc.* **2007**, *129*, 10300–10301. (e) Taniguchi, Y.; Kool, E. T. *J. Am. Chem. Soc.* **2007**, *129*, 8836–8844. (f) Seio, K.; Sasami, T.; Ohkubo, A.; Ando, K.; Sekine, M. *J. Am. Chem. Soc.* **2007**, *129*, 1026–1027. (g) Xia, J.; Noronha, A.; Toudjarska, I.; Li, F.; Akinc, A.; Braich, R.; Frank-Kamenetsky, M.; Rajeev, K. G.; Egli, M.; Manoharan, M. *ACS Chem. Biol.* **2006**, *1*, 176–183. (h) Krueger, A. T.; Kool, E. T. *Cuur. Opin. Chem. Biol.* **2007**, *11*, 588–594. (i) Kool, E. T.; Sintim, H. O. *Chem. Commun.* **2006**, 3665–3675. (j) Leconte, A. M.; Matsuda, S.; Romesberg, F. E. *J. Am. Chem. Soc.* **2006**, *128*, 6780–6781. (k) Benner, S. A.; Sismour, A. M. *Nat. Rev. Genet.* **2005**, *6*, 533–543. (l) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Biotechnol.* **2005**, *16*, 370–377. (m) Kawai, R.; Kimoto, M.; Ikeda, S.; Mitsui, T.; Endo, M.; Yokoyama, S.; Hirao, I. *J. Am. Chem. Soc.* **2005**, *127*, 17286–17295. (n) Dupradeau, F.-Y.; Case, D. A.; Yu, C.; Jimenez, R.; Romesberg, F. E. *J. Am. Chem. Soc.* **2005**, *127*, 15612–15617. (o) Leconte, A. M.; Chen, L.; Romesberg, F. E. *J. Am. Chem. Soc.* **2005**, *127*, 12470–12471. (p) Jung, K.-H.; Marx, A. *Cell. Mol. Life Sci.* **2005**, *62*, 2080–2091. (q) Benner, S. A. *Science* **2004**, *306*, 625–626. (r) Kimoto, M.; Endo, M.; Mitsui, T.; Okuni, T.; Hirao, I.; Yokoyama, S. *Chem. Biol.* **2004**, *11*, 47–55. (s) Benner, S. A. *Acc. Chem. Res.* **2004**, *37*, 784–797. (t) Hirao, I.; Harada, Y.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Yokoyama, S. *J. Am. Chem. Soc.* **2004**, *126*, 13298–13305.

(8) For homo-DNA, P-RNA, and hexitol nucleic acids: (a) Egli, M.; Pallan, P. S.; Pattanayek, R.; Wilds, C. J.; Lubini, P.; Minasov, G.; Dobler, M.; Leumann, C. J.; Eschenmoser, A. *J. Am. Chem. Soc.* **2006**, *128*, 10847–10856. (b) Eschenmoser, A. *Chimia* **2005**, *59*, 836–850. (c) Pitsch, S.; Wendeborn, S.; Krishnamurthy, R.; Holzner, A.; Minton, M.; Bolli, M.; Miculca, C.; Windhab, N.; Micura, R.; Stanek, M.; Jaun, B.; Eschenmoser, A. *Helv. Chim. Acta* **2003**, *86*, 4270–4363. (d) Jungmann, O.; Beier, M.; Luther, A.; Huynh, H. K.; Ebert, M.-O.; Jaun, B.; Krishnamurthy, R.; Eschenmoser, A. *Helv. Chim. Acta* **2003**, *86*, 1259–1308. (e) Wippo, H.; Reck, F.; Kudick, R.; Ramaseshan, M.; Ceulemans, G.; Bolli, M.; Krishnamurthy, R.; Eschenmoser, A. *Bioorg. Med. Chem.* **2001**, *9*, 2411–2428. (f) Beier, M.; Reck, F.; Wagner, T.; Krishnamurthy, R.; Eschenmoser, A. *Science* **1999**, *283*, 699–703. (g) Eschenmoser, A. *Science* **1999**, *284*, 2118–2124. (h) Groebke, K.; Hunziker, J.; Fraser, W.; Peng, L.; Diederichsen, U.; Zimmermann, K.; Holzner, A.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1998**, *81*, 375–474, and references therein.

All of the halogen-attached nonnatural bases were synthesized from appropriate commercial materials and connected to deoxyribose through an acetylene bond to afford nonnatural, stable C-nucleosides in good overall yields (Figure 1b and Scheme S1 in Supporting Information). Our laboratory already reported the synthesis of 4,4′-dimethoxytrityl- (DMTr-) protected ethynyl *C*-2-deoxy-β-D-ribofuranoside as a key component of nonnatural nucleosides used in this study.[11] By means of the synthetic versatility of the acetylenic function, a variety of candidates for nonnatural bases can be attached to the deoxyribose with the β-configuration of natural nucleosides. The C-nucleosides thus prepared were further protected on the bases and converted into phosphoramidites. The phosphoramidites were then subjected to automated DNA synthesis to give oligomeric nonnatural nucleotides, hereafter called simply "artificial DNA". After having been released from the solid support, the artificial DNAs were deprotected, purified by reverse-phase HPLC, and characterized by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (Figure S1). Preliminary modeling studies revealed that artificial DNA forms a similar duplex as natural DNA and that the deoxyribose−phosphodiester backbone of artificial DNA appears to scarcely be distorted by adopting suitable conformations. This motion may be assisted by increased puckering flexibilities of the deoxyribose with the aid of the acetylene spacers.

First, homooligomers [abbreviated as d(T*)$_y$, where y is the number of nucleoside residues] of artificial DNA were investigated solely on the basis of UV−vis and circular dichroism (CD) spectra in a buffer solution (10 mM Hepes, 10 mM $MgCl_2$, and 100 mM NaCl, pH 7.0). The UV−vis spectra of d(T*)$_{16}$ and d(*i*C*)$_8$ hardly changed in the temperature range from 0 to 70 °C, and their CD spectra were almost silent at 20 °C (Figure S2). Although d(A*)$_{16}$ also showed temperature-insensitive UV−vis absorption, strong Cotton effects were seen for the CD spectrum in the wavelength region where A* absorbs (Figure S3). The CD intensities somewhat weakened by about 40% when the temperature of d(A*)$_{16}$ was raised to 70 °C. Therefore, the emerging Cotton effects would simply be attributed to the chiral arrangement of the A* chromophores and not to the formation of a higher-order structure. On the other hand, the behavior of d(*i*G*)$_8$ is quite interesting, as in the case for natural G homooligomers (Figure 2). Significant Cotton effects of d(*i*G*)$_8$ at room temperature began to weaken upon heating with isodichroic points and finally almost disappeared, accompanied by hyperchromicity of 45% at 310 nm

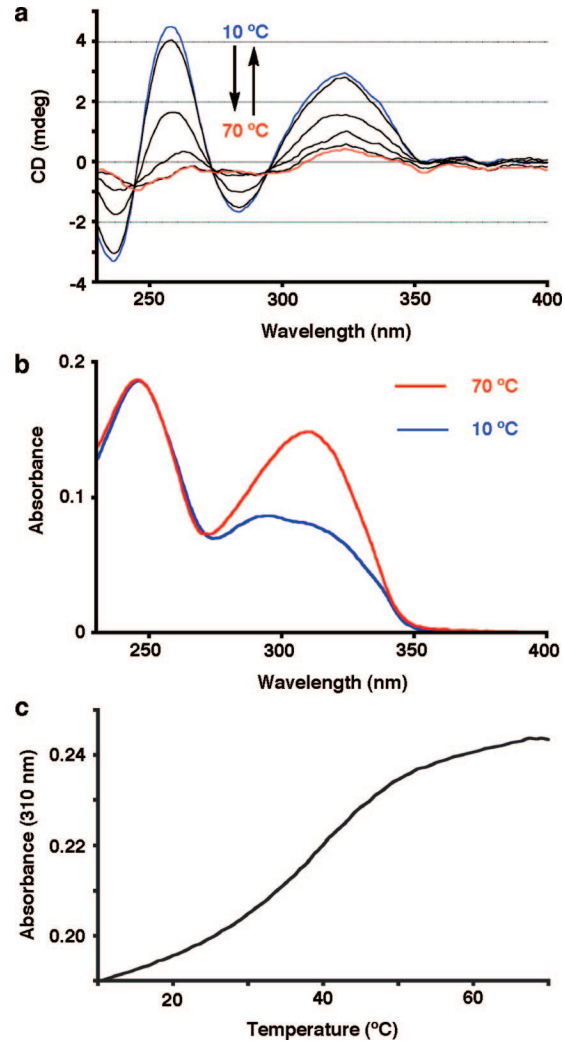

Figure 2. CD and UV−vis data for d(*i*G*)$_8$ (2.0 μM) in 10 mM Hepes (pH 7.0), 10 mM $MgCl_2$, 100 mM NaCl. (a) CD spectra at 10 (blue), 20, 30, 40, 50, 60, and 70 °C (red). (b) UV−vis spectra at 10 (blue) and 70 °C (red). (c) UV melting curve monitored at 310 nm.

in the UV−vis spectra. A cooperative, sigmoidal transition was definitely observed at 40 °C for d(*i*G*)$_8$ by thermal denaturation studies. Cooling of d(*i*G*)$_8$ to room temperature restored the Cotton effects and the hypochromism to the same level of the previous state. This finding means that the *i*G* homooligomer of artificial DNA forms a higher-order structure by itself, possibly resembling the G-quartet structure of natural DNA.

Table 1 summarizes the complementary sequences of artificial and natural DNA used in this study and their melting temperature (T_m) values. The addition of d(*i*C*)$_8$ to a solution of d(*i*G*)$_8$ caused drastic changes in the CD spectra, resulting from the transformation of the self-ordered structure of d(*i*G*)$_8$ into the hybridized one between d(*i*G*)$_8$ and d(*i*C*)$_8$ (entry 3). The Cotton effects of the mixture weakened when the temperature of the mixture was raised to 70 °C and approached that of the sum of the two oligomers at the same temperature (Figure 3a). Thermal denaturation profile of the mixture showed a typical two-state, sigmoidal transition with hyperchromicity of 25% at 305 nm in UV−vis spectra, from which T_m for the combination was estimated to be 59 °C (Figure 3b,c). The 1:1 stoichiometry for this complexation was confirmed on the basis of the CD changes by adding CD-silent d(*i*C*)$_8$ to a solution of d(*i*G*)$_8$,

(9) (a) Krueger, A. T.; Lu, H.; Lee, A. H. F.; Kool, E. T. *Acc. Chem. Res.* **2007**, *40*, 141–150. (b) Lynch, S. R.; Liu, H.; Gao, J.; Kool, E. T. *J. Am. Chem. Soc.* **2006**, *128*, 14704–14711. (c) Lee, A. H. F.; Kool, E. T. *J. Am. Chem. Soc.* **2006**, *128*, 9219–9230. (d) Lee, A. H. F.; Kool, E. T. *J. Am. Chem. Soc.* **2005**, *127*, 3332–3338. (e) Gao, J.; Liu, H.; Kool, E. T. *Angew. Chem., Int. Ed.* **2005**, *44*, 3118–3122. (f) Gao, J.; Liu, H.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 11826–11831. (g) Liu, H.; Lynch, S. R.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 6900–6905. (h) Lu, H.; He, K.; Kool, E. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 5834–5836. (i) Liu, H.; Gao, J.; Maynard, L.; Saito, Y. D.; Kool, E. T. *J. Am. Chem. Soc.* **2004**, *126*, 1102–1109. (j) Liu, H.; Gao, J.; Lynch, S. R.; Saito, Y. D.; Maynard, L.; Kool, E. T. *Science* **2003**, *302*, 868–871.

(10) Abe, H.; Takase, M.; Doi, Y.; Matsumoto, S.; Furusyo, M.; Inouye, M. *Eur. J. Org. Chem.* **2005**, 2931–2940.

(11) (a) Chiba, J.; Takeshima, S.; Mishima, K.; Maeda, H.; Nanai, Y.; Mizuno, K.; Inouye, M. *Chem.—Eur. J.* **2007**, *13*, 8124–8130. (b) Heinrich, D.; Wagner, T.; Diederichsen, U. *Org. Lett.* **2007**, *9*, 5311–5314. (c) Adamo, M. F. A.; Pergoli, R. *Org. Lett.* **2007**, *9*, 4443–4446. (d) Takase, M.; Morikawa, T.; Abe, H.; Inouye, M. *Org. Lett.* **2003**, *5*, 625–628.

Table 1. Artificial and Natural DNAs in This Study and Their Melting Temperatures

entry	duplexes and triplexes	T_m[a] (°C)
1	5′-d(*i*G*)$_8$	40[b]
2	5′-d(*i*G*)$_6$/3′-d(*i*C*)$_6$	41
3	5′-d(*i*G*)$_8$/3′-d(*i*C*)$_8$	59
4	5′-d(*i*G*)$_{10}$/3′-d(*i*C*)$_{10}$	70
5	5′-d(*i*G*)$_6$/3′-d(C)$_6$	33
6	5′-d(*i*G*)$_8$/3′-d(C)$_8$	50
7	5′-d(*i*G*)$_{10}$/3′-d(C)$_{10}$	58
8	5′-d(G)$_6$/3′-d(C)$_6$	30
9	5′-d(G)$_8$/3′-d(C)$_8$	55
10	5′-d(G)$_{10}$/3′-d(C)$_{10}$	58
11	5′-d(*i*G**i*C**i*G**i*C**i*G**i*G**i*C**i*C*)/ 3′-d(*i*C**i*G**i*C**i*G**i*C**i*C**i*G**i*G*)	59
12	3′-d(*i*G**i*C**i*G**i*C**i*G**i*G**i*C**i*C*)/ 3′-d(*i*C**i*G**i*C**i*G**i*C**i*C**i*G**i*G*)	*c*
13	5′-d(GCGCGGCC)/3′-d(CGCGCCGG)	55
14	5′-d(*i*G**i*C**i*G**i*C*A**i*G**i*G**i*C**i*C*)/ 3′-d(*i*C**i*G**i*C**i*G*T**i*C**i*C**i*G**i*G*)	56
15	5′-d(GCGCAGGCC)/3′-d(CGCGTCCGG)	55
16	5′-d(T**i*G**i*C**i*G**i*C**i*G**i*G**i*C**i*C*T*)/ 3′-d(A**i*C**i*G**i*C**i*G**i*C**i*C**i*G**i*G*A*)	66
17	5′-d(TGCGCGGCCT)/3′-d(ACGCGCCGGA)	64
18	5′-d(T**i*G**i*C**i*G**i*<u>C*</u>i*G**i*G**i*C**i*C*T*)/ 3′-d(A**i*C**i*G**i*C**i*<u>T*</u>i*C**i*C**i*G**i*G*A*)[d]	46
19	5′-d(TGCG<u>C</u>GGCCT)/3′-d(ACGC<u>T</u>CCGGA)[d]	35
20	5′-d(T**i*G**i*C**i*G**i*C*A**i*G**i*G**i*C**i*C*T*)/ 3′-d(A**i*C**i*G**i*C**i*G*T**i*C**i*C**i*G**i*G*A*)	64
21	5′-d(TGCGCAGGCCT)/3′-d(ACGCGTCCGGA)	63
22	5′-d(A*)$_{16}$/3′-d(T*)$_{16}$	22[e]
23	5′-d(A*)$_{20}$/3′-d(T*)$_{20}$	30[e]
24	5′-d(A*)$_{24}$/3′-d(T*)$_{24}$	35[e]
25	5′-d(A*A*T*T*A*A*T*T*A*A*T*T*A*A*T*T* A*A*T*T*)	*c*

[a] T_m values were obtained from the maxima of the first derivatives of the melting curves measured in the buffer solution: 2.0 μM duplex, 10 mM Hepes (pH 7.0), 10 mM $MgCl_2$, and 100 mM NaCl. Errors were estimated at ± 1.0 °C. [b] T_m value of self-ordered structure. [c] No apparent sigmoidal transition. [d] *i*C*T* and CT mismatch base pairs are underlined. [e] Exclusive triplex formation was observed: 1.0 μM A* homooligomer and 2.0 μM T* homooligomer.

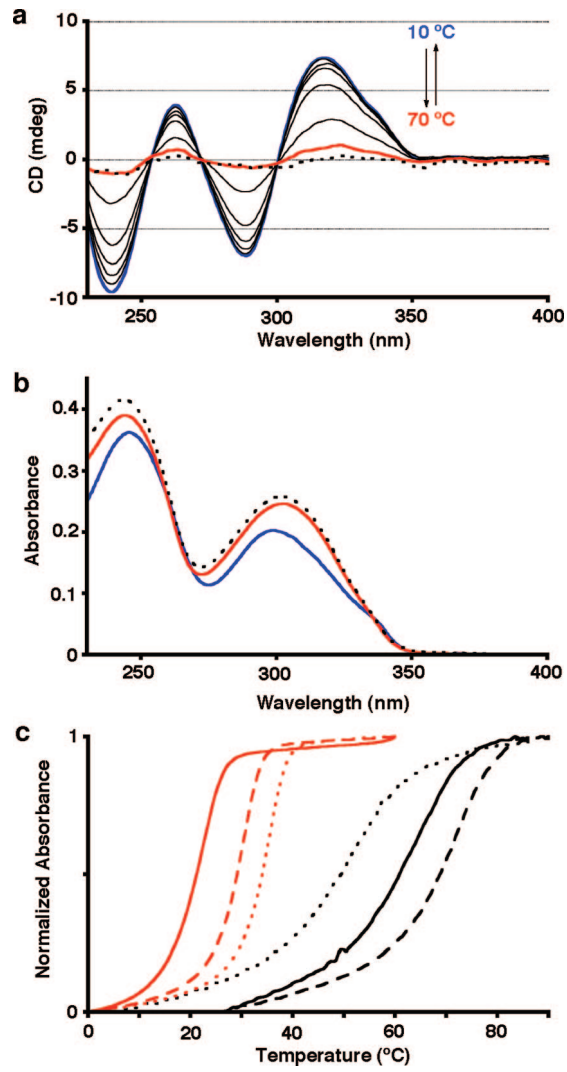

Figure 3. CD and UV−vis measurements were carried out at 2.0 μM duplex in 10 mM Hepes (pH 7.0), 10 mM $MgCl_2$, and 100 mM NaCl unless otherwise noted. (a) CD spectra at 10 (blue), 20, 30, 40, 50, 60, and 70 °C (red) of the duplex of entry 3. Dotted line indicates the sum of the CD spectra at 70 °C for each single strand of d(*i*G*)$_8$ and d(*i*C*)$_8$. (b) UV−vis spectra at 10 °C (blue) and 70 °C (red) of the duplex of entry 3. Dotted line indicates the sum of the UV−vis spectra at 70 °C for each single strand of d(*i*G*)$_8$ and d(*i*C*)$_8$. (c) UV melting curves of entry 3 (305 nm, black solid line), entry 16 (305 nm, black dashed line), entry 18 (305 nm, black dotted line), entry 22 (295 nm, red solid line), entry 23 (295 nm, red dashed line), and entry 24 (295 nm, red dotted line). In the cases of entries 22−24, [A* homooligomer] = 1.0 μM and [T* homooligomer] = 2.0 μM.

which showed an apparent saturation at the molar ratio of 1 (Figure 4a). These observations indeed suggest the formation of a complementary hydrogen-bonding duplex of d(*i*G*/*i*C*)$_8$ between the two strands of the artificial DNAs with π-stacking interactions. Furthermore, the positive−negative pair of the CD bands from the longer wavelength is structurally analogous to that of the natural DNA of d(G/C)$_8$ of entry 9, implying the right-handedness of the artificial duplex (Figure S4).[12] The artificial DNA d(*i*G*)$_8$ also forms a similar two-stranded complex with the natural DNA d(C)$_8$, giving rise to a T_m value of 50 °C, while the higher-order structure for this heterotype duplex remains to be elucidated (entry 6 and Figure S5). Thus, the relative stabilities of the higher-order structures decreased in the following order: duplex of d(*i*G*/*i*C*)$_8$ > duplex of d(G/C)$_8$ > duplex of d(*i*G*/C)$_8$. The highest stability for the artificial−artificial combination is also observed in the cases for d(*i*G*)$_6$ and d(*i*G*)$_{10}$.

Duplex formation of artificial DNA proved to have a generality as shown in entries 11−21. Thus, artificial DNA of various mixed sequences formed artificial duplexes with complementary artificial DNA. The thermal denaturation profile for artificial DNA of a mixed sequence was illustrated with the combination of entry 16 (Figure 3c and Figure S6), and its 1:1 stoichiometry was also corroborated (Figure S7). The antiparallel orientation of artificial duplexes was determined by comparing entries 11 (antiparallel complementarity) and 12 (parallel complementarity): no characteristic sigmoidal transition was seen in entry 12. Artificial duplexes revealed a similar relationship between the T_m values and the sequence contexts as for natural duplexes. The T_m values consecutively rose with increasing number of nucleoside residues when the same base pair was added to the duplexes of a single base pair (entries 2−4 vs 8−10). Duplexes containing two types of base pairs behaved in a complicated manner. When an A*/T* base pair was added into the middle of a *i*G*/*i*C* duplex sequence, the T_m values were faintly influenced (entries 11 vs 14 and 16 vs 20). On the other hand, the T_m values of the duplexes substantially increased by adding A*/T* base pairs at the ends of the sequences (entries 11 vs 16). It was also found that the artificial duplexes of mixed sequences have thermal stabilities near the stabilities of

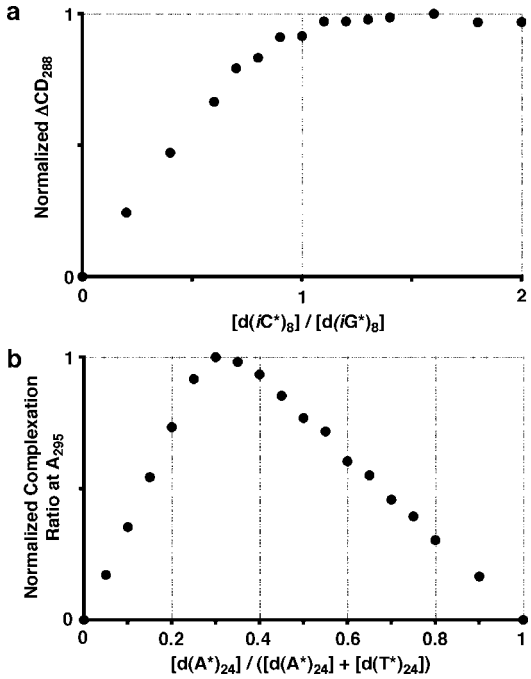

Figure 4. Evaluation of stoichiometry for hybridization events. (a) Titration curve monitored by CD (288 nm, 25 °C) for the duplex of entry 3 at $[d(iG^*)_8] = 2.0\ \mu M$ in 10 mM Hepes (pH 7.0), 10 mM $MgCl_2$, and 100 mM NaCl. (b) Job's plot monitored by UV (295 nm, 0 °C) for the duplex of entry 24 at $[d(A^*)_{24}] + [d(T^*)_{24}] = 2.0\ \mu M$ in 10 mM phosphate buffer (pH 7.0), 0.1 mM EDTA, and 100 mM NaCl.

natural duplexes with analogous sequences (entries 14 vs 15, 16 vs 17, and 20 vs 21). To shed light on this similarity, the thermodynamic parameters were measured for these artificial DNAs by van't Hoff analysis (Figure S13).[13] The ratios of $\Delta H/T\Delta S$ for the duplex formation are about 1.3 and are almost equal to those in natural DNA of the analogous sequences. Artificial DNA was found to compensate a disadvantageous entropy change with an advantageous enthalpy one in the duplex formation, equally as well as natural DNA does.

The thermodynamic stability of duplexes should principally result from two factors: hydrogen-bonding and stacking (and/or hydrophobic) interactions. Although the hydrogen-bonding pattern for the *i*G*/*i*C* combination (DDA/AAD type) in the artificial duplex is reversed from that of the natural G/C one (ADD/DAA type), the total number and the interaction type (i.e., DDA/AAD or DAD/ADA or DDD/AAA) of hydrogen bonds are the same.[14] With respect to the stacking interaction, two different aspects must be considered for comparing the artificial and natural base pairs. The artificial base pair has a pyrimidine/pyrimidine plane that is smaller than a pyrimidine/purine one of the natural base pair. On the other hand, acetylene linkers in the artificial pair may contribute to add an extra stabilization for the stacking as reported in other cases.[15] Thus, the insufficiency for the interaction caused by the smaller plane should be compensated by the presence of the acetylene linkers in the artificial system. These factors appear to play a role in the observation that the artificial duplexes have thermal stabilities near those of the natural duplexes of analogous sequences.

An important feature of natural DNA is its discerning ability between matched and mismatched sequences even in the presence of only a single mismatched base if the duplex is short. Entry 18 represents the combination of such artificial DNA, in which a single mismatched base pair of *i*C*/T* exists near the center of the 10-mer duplex. The T_m value for the mismatched duplex largely decreased by 20 °C compared to that for the corresponding fully matched duplex in entry 16 (Figure 3c and Figure S8). This drop in T_m is noticeable because the size of the mismatched *i*C*/T* base pair is formally identical to that of the matched *i*C*/*i*G* base pair in our artificial duplexes. Pairing selectivity in artificial DNA comes only from complementary hydrogen bonding, not from complementary size. Nevertheless, the ΔT_m is comparable to that of the natural counterpart, which may cause a significant backbone distortion (entries 17 and 19). This finding demonstrates the preciseness of our molecular design for the two types of nonnatural, incompatible base pairs as shown in Figure 1a.

Homooligomers composed of A* and T* were subjected to hybridization experiments (entries 22−24 and Figure S9) and found to behave somewhat unlike the corresponding natural DNA does. The Job's plot between $d(A^*)_{24}$ and $d(T^*)_{24}$ was conducted on the basis of UV−vis spectra at 0 °C after annealing processes (Figure 4b). Mole fraction of $d(A^*)_{24}$ was raised incrementally from 0% to 100% under conditions where the total concentration of $d(A^*)_{24}$ and $d(T^*)_{24}$ was maintained constant at 2.0 μM. In this experiment, we observed that one set of isosbestic points changed to the other set strictly when a mole fraction of $d(A^*)_{24}$ increased beyond 0.33. After that, the absorption bands of $d(A^*)_{24}$ simply grew in the spectra (Figure S10a). This behavior was also seen in the CD Job's (Figure S10b) and titration (Figure S11) experiments. The thermal denaturation profiles revealed a similar two-state sigmoidal transition even when the molar ratio of $d(A^*)_{24}$ was varied (Figure 3c). Furthermore, the slope around the inflection point on the sigmoidal curve is much steeper than any other combination, indicating a strongly cooperative two-state transition (Figure S12). These data unambiguously suggest the exclusive formation of $d(T^*/A^*/T^*)_{24}$-type artificial triplex without $d(A^*/T^*)_{24}$-type duplex even at the low concentration applied.

(12) (a) Johnson, W. C. In *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley−VCH: New York, 2000; pp 703−718. (b) Maurizot, J. C. In *Circular Dichroism: Principles and Applications*, 2nd ed.; Berova, N., Nakanishi, K., Woody, R. W., Eds.; Wiley−VCH: New York, 2000; pp 719−739.

(13) Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601–1620.

(14) (a) Chen, G.; Kierzek, R.; Yildirim, I.; Krugh, T. R.; Turner, D. H.; Kennedy, S. D. *J. Phys. Chem. B* **2007**, *111*, 6718–6727. (b) Johnson, S. C.; Sherrill, C. B.; Marshall, D. J.; Moser, M. J.; Prudent, J. R *Nucleic Acids Res.* **2004**, *32*, 1937–1941. (c) Chaput, J. C.; Switzer, C. *J. Am. Chem. Soc.* **2000**, *122*, 12866–12867. (d) Robinson, H.; Gao, Y.-G.; Bauer, C.; Roberts, C.; Switzer, C.; Wang, A. H.-J. *Biochemistry* **1998**, *37*, 10897–10905. (e) Roberts, C.; Bandaru, R.; Switzer, C. *J. Am. Chem. Soc.* **1997**, *119*, 4640–4649. (f) Roberts, C.; Bandaru, R.; Switzer, C. *Tetrahedron Lett.* **1995**, *36*, 3601–3604. (g) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33–37. (h) Switzer, C.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322–8323.

(15) (a) Chaput, J. C.; Sinha, S.; Switzer, C. *Chem. Commun.* **2002**, 1568–1569. (b) Ogawa, A. K.; Wu, Y.; McMinn, D. L.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 3274–3287. (c) McMinn, D. L.; Ogawa, A. K.; Wu, Y.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **1999**, *121*, 11585–11586. (d) Sági, J.; Szemozö, A.; Ébinger, K.; Szabolcs, A.; Sági, G.; Ruff, É.; Ötvös, L. *Tetrahedron Lett.* **1993**, *34*, 2191–2194. (e) Froehler, B. C.; Jones, R. J.; Cao, X.; Terhorst, T. J. *Tetrahedron Lett.* **1993**, *34*, 1003–1006.

This strong tendency for the formation of triplexes over duplexes should primarily be due to the symmetry of A* along the axis from the amino nitrogen to the carbon connected to the acetylene linker. Therefore, interaction of the third $d(T^*)_{24}$ strand with $d(A^*/T^*)_{24}$ produces no energetic loss and rather adds a further stabilization arising from the enforced π-stacking interaction of the enlarged π-plane. This situation should cause the equilibrium constant between $d(T^*)_{24}$ and $d(A^*/T^*)_{24}$ to be larger than that of the $d(A^*/T^*)_{24}$ duplex formation, resulting in the observed two-state transition. In the cases of natural A and T homooligomers, the relative magnitudes of the two equilibrium constants are strongly depend on the concentrations of the oligomers and the existing salts.[16] The selectivity for the formation of triplex over duplex is so high that the thermal denaturation study of the palindromic 20-mer in entry 25 displayed little hyperchromism and no apparent sigmoidal transition. This finding reconfirms that triplexes predominantly are formed rather than duplexes from the artificial DNAs composed only of A* and T*.

Conclusion

Artificial DNA made exclusively of nonnatural nucleosides with four types of nonnatural bases represents a new class of DNA-like synthetic oligomers. The *i*G*/*i*C*-rich artificial DNA forms right-handed duplexes with the complementary artificial DNA in a sequence-specific manner with antiparallel orientation. The thermal stabilities and thermodynamic parameters of the artificial duplexes are very close to those of the natural duplexes in spite of differences in their geometries and hydrogen-bonding patterns. On the other hand, the artificial homooligomeric DNAs consisted only of A* and T* were found to exclusively form triplexes. By means of the synthetic feasibility of the artificial nucleoside unit, a variety of candidates for nonnatural bases can be incorporated into the artificial DNA. Therefore, the present molecular framework has a potential for storing genetic information and for application to enzymatic replication directed toward engineered genetics. Furthermore, the artificial DNA may be a superior building scaffold for constructing nanostructures of materials interest because of the stable C-nucleosides against ubiquitous naturally occurring enzymes such as DNase.

Experimental Section

Synthesis of Artificial DNA Oligomers. Artificial DNA oligomers were synthesized by use of phosphoramidites **8**, **12**, **16**, and **19** (see Supporting Information and Scheme S1) on an Applied Biosystems 392 synthesizer by standard β-cyanoethylphosphoramidite chemistry with a coupling reaction time of 15 min. The solid support (Universal Support II), which allows for 3′ placement of nonnatural nucleosides, was purchased from Glen Research. After automated synthesis, the oligomers were removed from the solid support with 2 M ammonia methanol solution at 30 °C for 30 min and deprotected with concentrated NH_4OH at 40 °C for 8 h. The oligomers were then purified by reverse-phase HPLC on a Chemcobond 5-ODS-H column (10 × 150 mm) with an eluent of 5 mM ammonium formate and a linear gradient of acetonitrile (0−30 min, 0−15% CH_3CN) at a flow rate of 3.0 mL/min.

MALDI-TOF Mass Measurements. MALDI-TOF mass spectra were recorded on a Bruker-Daltonics-Autoflex mass spectrometer operating in the positive ion mode with 3-hydroxypicolinic acid as a matrix (see Figure S1). $d(iG^*)_6$: calcd for MH^+ ($C_{66}H_{73}N_{18}O_{34}P_5$), 1817.33; found, 1817.92. $d(iG^*)_8$: calcd for MH^+ ($C_{88}H_{97}N_{24}O_{46}P_7$), 2444.43; found, 2444.64. $d(iG^*)_{10}$: calcd for MH^+ ($C_{110}H_{121}N_{30}O_{58}P_9$), 3070.52; found, 3070.88. $d(iC^*)_6$: calcd for MH^+ ($C_{72}H_{85}N_{18}O_{34}P_5$), 1901.42; found, 1902.44. $d(iC^*)_8$: calcd for MH^+ ($C_{96}H_{113}N_{24}O_{46}P_7$), 2556.55; found, 2556.28. $d(iC^*)_{10}$: calcd for MH^+ ($C_{120}H_{141}N_{30}O_{58}P_9$), 3210.68; found, 3210.97. $d(A^*)_{16}$: calcd for MH^+ ($C_{176}H_{193}N_{48}O_{78}P_{15}$), 4693.88; found, 4694.84. $d(A^*)_{20}$: calcd for MH^+ ($C_{220}H_{241}N_{60}O_{98}P_{19}$), 5882.09; found, 5884.77. $d(A^*)_{24}$: calcd for MH^+ ($C_{264}H_{289}N_{72}O_{118}P_{23}$), 7071.30; found, 7071.71. $d(T^*)_{16}$: calcd for MH^+ ($C_{192}H_{209}N_{32}O_{110}P_{15}$), 5189.79; found, 5191.35. $d(T^*)_{20}$: calcd for MH^+ ($C_{240}H_{261}N_{40}O_{138}P_{19}$), 6502.98; found, 6503.89. $d(T^*)_{24}$: calcd for MH^+ ($C_{288}H_{313}N_{48}O_{166}P_{23}$), 7815.17; found, 7816.01. 5′-d(*i*G*C*iG*C*iG*C*iG*C*iC*C*)-3′: calcd for MH^+ ($C_{92}H_{105}N_{24}O_{46}P_7$), 2500.49; found, 2500.60. 3′-d(C*iG*C*iG*C*iG*C*i-C*iG*C*iG*)-5′: calcd for MH^+ ($C_{92}H_{105}N_{24}O_{46}P_7$), 2500.49; found, 2500.50. 3′-d(*i*G*C*iG*C*iG*C*iG*C*iC*C*)-5′: calcd for MH^+ ($C_{92}H_{105}N_{24}O_{46}P_7$), 2500.49; found, 2500.66. 5′-d(*i*G*C*iG*C*A*iG*iG*C*iC*)-3′: calcd for MH^+ ($C_{103}H_{117}N_{27}O_{51}P_8$), 2797.54; found, 2798.50. 3′-d(*i*C*-iG*iC*iG*T*iC*iC*iG*iG*)-5′: calcd for MH^+ ($C_{104}H_{118}N_{26}O_{53}P_8$), 2828.53; found, 2828.88, 5′-d(T*iG*iC*iG*iC*iG*iG*iC*iC*T*)-3′: calcd for MH^+ ($C_{116}H_{131}N_{28}O_{60}P_9$), 3156.58; found, 3156.72. 3′-d(A*iC*-iG*iC*iG*iC*iC*iG*iG*A*)-5′: calcd for MH^+ ($C_{114}H_{129}N_{30}O_{56}P_9$), 3094.59; found, 3093.61. 3′-d(A*iC*iG*iC*T*iC*iC*iG*iG*A*)-5′: calcd for MH^+ ($C_{115}H_{130}N_{29}O_{57}P_9$), 3109.59; found, 3110.05. 5′-d(T*iG*iC*iG*-iC*A*iG*iG*iC*iC*T*)-3′: calcd for MH^+ ($C_{127}H_{143}N_{31}O_{65}P_{10}$), 3453.63; found, 3453.34. 3′-d(A*iC*iG*iC*iG*T*iC*iC*iG*iG*A*)-5′: calcd for MH^+ ($C_{126}H_{142}N_{32}O_{63}P_{10}$), 3422.64; found, 3421.71. 5′-d(A*A*T*T*)$_5$-3′: calcd for MH^+ ($C_{230}H_{251}N_{50}O_{118}P_{19}$), 6130.04; found, 6129.23.

UV and T_m Measurements. UV−vis spectra and T_m melting curves (1.0 °C/1.0 min) were obtained on a Jasco V-560 UV/vis spectrophotometer with a Peltier and a temperature controller in a temperature range from 0 to 90 °C. The T_m values were determined from the maxima of the first derivatives of the melting curves measured in a buffer solution: 10 mM Hepes (pH 7.0), 10 mM $MgCl_2$, and 100 mM NaCl. Errors were estimated at ±1.0 °C. Concentrations of the solutions containing artificial DNAs were determined on the basis of molar extinction coefficients at 260 nm (ε_{260}) of the artificial nucleoside monomers **9**, **13**, **17**, and **20** (see Supporting Information).

CD Measurements. CD spectra were recorded on a Jasco-J-720WI spectropolarimeter with a temperature controller at 0, 10, 20, 30, 40, 50, 60, and 70 °C in a buffer solution: 10 mM Hepes (pH 7.0), 10 mM $MgCl_2$, and 100 mM NaCl.

Titration Experiments. Titration curves for artificial DNAs were obtained by monitoring a specified wavelength of CD or UV. In entry 3, for example, 3.0 mL of a $d(iG^*)_8$ solution (2.0 μM with 10 mM Hepes (pH 7.0), 10 mM $MgCl_2$, and 100 mM NaCl) was prepared, and CD measurement of the solution was carried out at 25 °C in a quartz cell of 1 cm path length. Separately, 150 μL of a $d(iC^*)_8$ solution (100 μM in the same buffer) was then prepared, and 12.0 μL of the $d(iC^*)_8$ solution [0.2 equiv relative to $d(iG^*)_8$] was added to the $d(iG^*)_8$ solution in the quartz cell. The mixed solution was stirred for 15 min at room temperature, and then CD measurement was performed. A series of the operations was repeated for all the ratios of $[d(iC^*)_8]/[d(iG^*)_8] = 0$, 0.2, 0.4, 0.6, 0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.6, 1.8, and 2.0. The normalized CD changes at 288 nm were plotted against $[d(iC^*)_8]/[d(iG^*)_8]$ (Figure 4a).

Job's Plots. Job's plots for artificial DNAs were obtained by monitoring a specified wavelength of CD or UV. In entry 24, for example, 20 mL of a $d(A^*)_{24}$ solution [2.0 μM with 10 mM phosphate (pH 7.0), 0.1 mM EDTA, and 100 mM NaCl] and 20 mL of a $d(T^*)_{24}$ solution (2.0 μM in the same buffer) were prepared. The $d(A^*)_{24}$ solutions of 0, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 900, and 1000 μL were mixed in micro test tubes with 1000, 950, 900, 850, 800, 750, 700, 650, 600, 550, 500, 450, 400, 350, 300, 250, 200, 100, and 0 μL of the $d(T^*)_{24}$ solutions, respectively. All the mixed solutions were heated to 80 °C, followed by slow cooling to 25 °C over 60 min,

(16) Roberts, R. W.; Crothers, D. M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 4320–4325.

and then kept in a refrigerator (4 °C) for 1 h. UV−vis measurements of all of the solutions were carried out at 0 °C in a quartz cell of 1 cm path length. Against $[d(A^*)_{24}]/([d(A^*)_{24}] + [d(T^*)_{24}])$ were plotted the normalized UV changes for the mixtures of $d(A^*)_{24}$ and $d(T^*)_{24}$ at 295 nm. The changes were corrected by subtracting the sum of the intensities for each strand at the same concentrations (Figure 4b).

Acknowledgment. This study was performed through Special Coordination Funds for Promoting Science and Technology of the Ministry of Education, Culture, Sports Science and Technology, the Japanese Government.

Supporting Information Available: Synthetic details, Scheme S1, and Figures S1−13. This material is available free of charge via the Internet at http://pubs.acs.org.

JA801058H