ORGANIC LETTERS

2009
Vol. 11, No. 6
1377−1379

Synthesis and Reaction of DNA Oligomers Containing Modified Cytosines Related to Bisulfite Sequencing

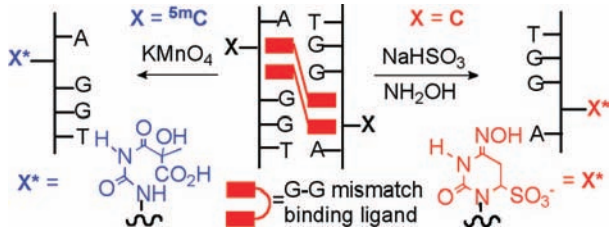
Yoshimi Oka,[†,‡] Tao Peng,[†] Fumie Takei,[†] and Kazuhiko Nakatani*[,†]

Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan, and Career Design Laboratory for Women, Osaka University, 1-1 Yamadaoka, Suita, Osaka 567-0047, Japan

nakatani@sanken.osaka-u.ac.jp

Received January 30, 2009

ABSTRACT



The synthesis of DNA oligomers containing *N*⁴-hydroxy-5,6-dihydrocytosine-6-sulfonate by using ligand-induced base flipping of cytosine followed by the simultaneous addition of bisulfite and hydroxylamine is reported. In contrast to C, the flipped-out 5-methylcytosine was selectively oxidized over thymines and cytosines in the duplex by potassium permanganate. Ligand-induced base flipping is a convenient and powerful strategy for the synthesis of modified cytosines and 5-methylcytosines related to bisulfite sequencing at the predetermined site of DNA.

Methylation of the cytosine at the 5′ position adjacent to guanine has a repressive effect on the expression of many eukaryotic genes. Failure of epigenetic controls of gene expression is associated with many human diseases, including neurodevelopmental disorders and cancer.[1−3] Hayatsu[4,5] and Shapiro[6,7] independently reported that cytosine moieties in single-stranded DNA (ssDNA) are effectively deaminated to uracil by bisulfite, whereas 5-methylcytosines are resistant to the reaction and remain unchanged.[4−8] This cytosine-specific deamination reaction constitutes the chemical basis of bisulfite genomic sequencing, which indirectly determines the site of 5-methylcytosines from C-to-T mutations.[9−12] One of the issues in the bisulfite sequencing is the low recovery of genomic DNA due to the use of a high concentration of bisulfite (more than 3 M) under low pH (typically at pH 5.0) at high temperature (at above 50 °C).[13] Furthermore, the

[†] Institute of Scientific and Industrial Research.
[‡] Career Design Laboratory for Women.

(1) Jones, P. A.; Baylin, S. B. *Nat. Rev. Genet.* **2002**, *3*, 415–428.
(2) Egger, G.; Liang, G.; Aparicio, A.; Jones, P. A. *Nature* **2004**, *429*, 457–463.
(3) Esteller, M. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 629–656.
(4) Hayatsu, H.; Wataya, Y.; Kai, K.; Iida, S. *Biochemistry* **1970**, *9*, 2858–2865.
(5) Hayatsu, H. *J. Mol. Biol.* **1977**, *115*, 19–31.
(6) Shapiro, R.; Servis, R. E.; Welcher, M. *J. Am. Chem. Soc.* **1970**, *92*, 422–424.
(7) Shapiro, R. *Biochem. Biophys. Res. Commun.* **1970**, *40*, 839–843.
(8) Wang, R. Y.-H.; Gehrke, C. W.; Ehrlich, M. *Nucleic Acids Res.* **1980**, *8*, 4777–4790.
(9) Chen, H.; Shaw, B. R. *Biochemistry* **1994**, *33*, 4121–4129.
(10) Frommer, M.; McDonald, L. E.; Millar, D. S.; Collis, C. M.; Watt, F.; Grigg, G. W.; Molloy, P. L.; Paul, C. L. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 1827–1831.
(11) Herman, J. G.; Graff, J. R.; Myöhänen, S.; Nelkin, B. D.; Baylin, S. B. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 9821–9826.
(12) Ehrich, M.; Nelson, M. R.; Stanssens, P.; Zabeau, M.; Liloglou, T.; Xinarianos, G.; Cantor, C. R.; Field, J. K.; van den Boon, D. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15785–15790.

10.1021/ol900195z CCC: $40.75
Published on Web 02/26/2009

intermediates of bisulfite reaction remained unclear. To improve the bisulfite sequencing reaction, oligonucleotides possessing modified cytosines with bisulfite at the predetermined site would be necessary for the study on the precise reaction mechanism and transformation. Sodium bisulfite is known to add to the C6 of cytosine to produce sulfonic acid salt. Due to the facile elimination forming a C5−C6 double bond under basic conditions and the very polar nature of the C6-sulfonic derivative of cytosine, automated DNA synthesis using the phosphoramidite reagents of modified cytosine bases with bisulfite was not available.

Previously, we reported that the naphthyridine carbamate dimer (**NCD**) binds to the XGG/XGG triad (X = C and T) in dsDNA with a stoichiometry of 2:1 and that nucleotide base X, which is located at the 5′ side of the G−G mismatch, is flipped out of the duplex π-stack (Figure 1).[14,15] The **NCD**

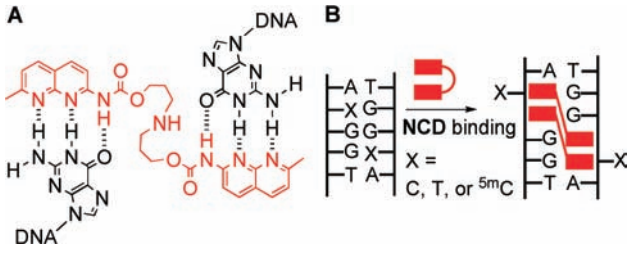


Figure 1. (A) Structure of **NCD** and a hydrogen-bonding pattern to two guanines. (B) The binding scheme of two **NCD** molecules to the XGG/XGG triad (X = C, T, and ^{5m}C).

binding to the XGG/XGG triad offers a unique opportunity to chemically modify the flipped out nucleotide base X in the duplex. Here we report the synthesis and subsequent transformation of DNA containing N^4-hydroxy-5,6-dihydrocytosine-6-sulfonate by the simultaneous addition of bisulfite and hydroxylamine to the flipped out cytosines by ligand binding. Furthermore, the discrimination of flipped out C from ^{5m}C was discussed in detail.

First, we confirmed the binding of **NCD** to the ^{5m}CGG/^{5m}CGG sequence. The melting temperature (T_m) of an 11-mer duplex, 5′-d(TCAA ^{5m}CGG TTGA)-3′/5′-d(TCAA ^{5m}CGG TTGA)-3′ (**^{5m}1**), increased by 22.5 °C in the presence of **NCD**, demonstrating that **NCD** bound to the ^{5m}CGG/^{5m}CGG triad. Cold-spray ionization time-of-flight mass spectrometry of the **^{5m}1**−**^{5m}1** duplex in the presence of **NCD** detected ions with *m*/*z* of 1548.9 and 1936.4, corresponding to the 5^- ([G−G + 2**NCD**]$^{5-}$ calcd 1549.4) and 4^- ([G−G + 2**NCD**]$^{4-}$ calcd 1936.8) ions of the 2:1 complex, respectively (Figure S1, Supporting Information). A large T_m increase and 2:1 binding stoichiometry were common features of **NCD** binding that resulted in base flipping.[14,15]

Having duplexes with a flipped-out ^{5m}C or C in hand, we examined reactions to bisulfite of the 11-mer duplex, 5′-d(TCAA XGG TTGA)-3′/5′-d(TCAA XGG TTGA)-3′, in which X = C or ^{5m}C, under a variety of conditions. The unmethylated duplex, **1**−**1**, was resistant to bisulfite at temperatures less than its melting point (33.0 °C). The presence of hydroxylamine usually accelerates the bisulfite reaction,[5] but the **1**−**1** duplex did not react with bisulfite (0.5 M) and hydroxylamine (0.5 M) at pH 6.0 (Figure S2, Supporting Information). These observations confirm that the C flanking the G−G mismatch is unresponsive to bisulfite. In contrast, in the presence of **NCD** (40 μM), a single major product was obtained from the **1**−**1** duplex (13 μM) in the presence of bisulfite and hydroxylamine. The reaction even proceeded at 0 °C at mild pH 6.0 (Figure 2A). MALDI-TOF MS showed that

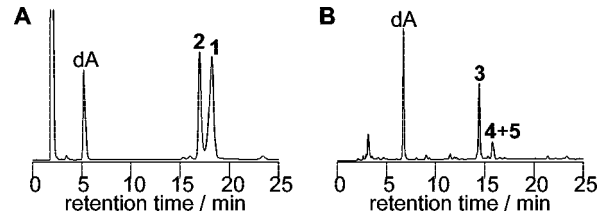


Figure 2. HPLC profiles of **1** under the bisulfite and hydroxylamine treatments. (A) The 11-mer duplex 5′-d(TCAA CGG TTGA)-3′/5′-d(TCAA CGG TTGA)-3′ (**1**−**1**, 13 μM as a duplex) in NaCl (100 mM) was treated with sodium bisulfite (0.5 M, pH 6.0) and hydroxylamine (0.5 M, pH 6.0) in the presence of **NCD** (40 μM) at 0 °C for 12 h. (B) Product **2** was isolated and treated with piperidine at 90 °C for 30 min.

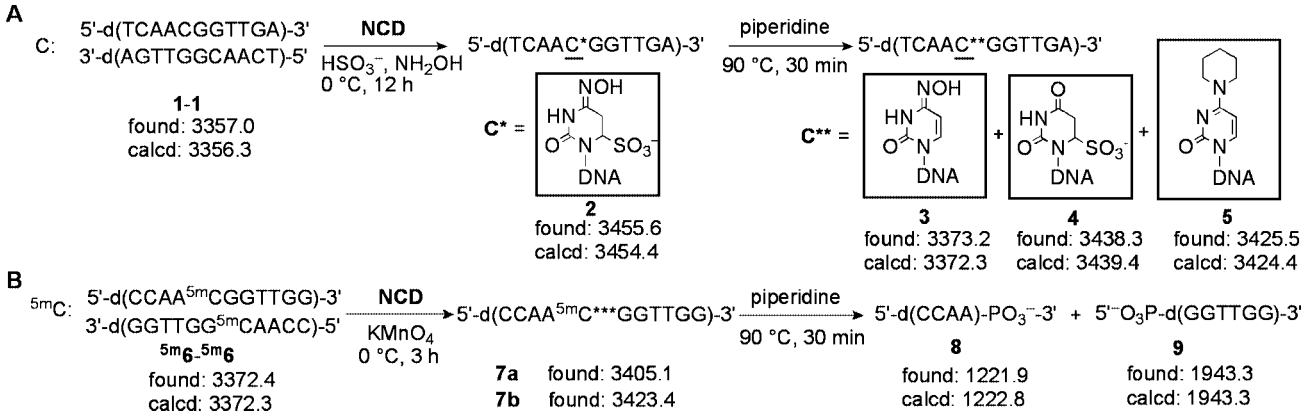
the product, **2** ([M−H]$^-$, found 3455.6, calcd 3454.4), was the adduct of **1** ([M−H]$^-$, found 3357.0, calcd 3356.3) with one bisulfite and one hydroxylamine. According to a description of the reaction of cytosine with bisulfite and hydroxylamine,[16] the C in the triad would most likely have been transformed into N^4-hydroxy-5,6-dihydrocytosine-6-sulfonate (C*), as shown in Scheme 1. The reactivity to bisulfite and hydroxylamine of the C that flipped out of the **1**−**1** duplex was similar to that of the C in ssDNA (Figures S5 and S7, Supporting Information).

The chemical reactivity of the oligomer that contains the C* has not been described in the literature. Piperidine heating of the oligomer that contains the C* at 90 °C for 30 min mainly produced an oligomer having N^4-hydroxycytosine, **3** ([M−H]$^-$, found 3373.2, calcd 3372.3), as confirmed by the MALDI-TOF MS (Figure 2B). N^4-Hydroxycytosine is stable under alkaline conditions.[5] The minor product of the reaction was a mixture of oligomer **4**, which contained 5,6-dihydrouracil-6-sulfonate ([M−H]$^-$, found 3438.3, calcd 3439.4), and oligomer **5** ([M−H]$^-$, found 3425.5, calcd 3424.4), in which piperidine was exchanged for hydroxylamine. Neither the 11-mer duplex **^{5m}1**−**^{5m}1**, which contained 5-methylcytosines, nor the ^{5m}C in the ssDNA reacted with bisulfite and hydroxylamine under the conditions where **2** was produced from the unmethylated duplex (Figure S2C and S7B, Supporting Information).

In contrast to the selective deamination of C in the presence of ^{5m}C,[4−8] selective transformation of ^{5m}C in the

(13) Hayatsu, H. *Mutat. Res.* **2008**, *659*, 77–82.
(14) Peng, T.; Nakatani, K. *Angew. Chem., Int. Ed.* **2005**, *44*, 7280–7283.
(15) Peng, T.; Dohno, C.; Nakatani, K. *Angew. Chem., Int. Ed.* **2006**, *45*, 5623–5626.
(16) Barnes, C. L.; Hawkinson, S. W. *Acta Crystallogr.* **1982**, *B38*, 1907–1911.

Scheme 1. Reaction of the Flipped Out (A) C and (B) ^{5m}C



presence of C and T does not occur unless the reactive species tethered to the DNA oligomer are delivered to the vicinity of the predetermined site.[17−19] To gain insight into the reactivity of the flipped-out ^{5m}C and discrimination of ^{5m}C from C and T in the duplex, the flipped-out ^{5m}C was treated with potassium permanganate under a variety of conditions. Preliminary experiments showed that the T at the 5′ end of the **^{5m}1**−**^{5m}1** duplex was oxidized by potassium permanganate in preference to the flipped-out ^{5m}C via partial denaturation of the duplex (Figure S9, Supporting Information). To investigate the oxidation of ^{5m}C in detail, we used the oligomer 5′-d(CCAA ^{5m}CGG TTGG)-3′ (**^{5m}6**), in which the 5′-end T in **^{5m}1** was replaced with C. In the presence of **NCD** (40 μM), a single product **7** was obtained from the **^{5m}6**−**^{5m}6** duplex (13 μM) when treated with potassium permanganate (0.2 mM) in sodium cacodylate buffer (10 mM, pH 6.0) at 0 °C (Figure 3A). Piperidine heating of

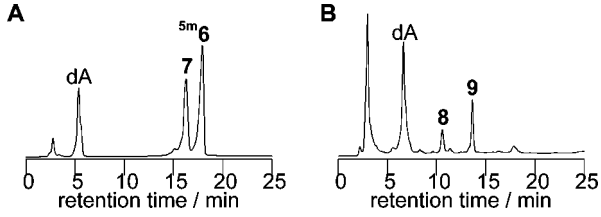


Figure 3. HPLC profiles of **^{5m}6** under the potassium permanganate treatment. (A) The 11-mer duplex 5′-d(CCAA ^{5m}CGG TTGG)-3′/5′-d(CCAA ^{5m}CGG TTGG)-3′ (**^{5m}6**-**^{5m}6**, 13 μM as a duplex) in NaCl (100 mM) and sodium cacodylate buffer (10 mM, pH 6.0) was treated with potassium permanganate (0.2 mM) in the presence of **NCD** (40 μM) at 0 °C for 3 h. (B) Product **7** was isolated and treated with piperidine at 90 °C for 30 min.

isolated **7** produced **8** and **9**, which were identified by MALDI-TOF MS analysis as 5′-d(CCAA)-PO_3^--3′ ([M−H]$^-$, found 1221.9, calcd 1222.8) and 5′-$^-O_3$P-d(GGTTGG)-3′ ([M−H]$^-$, found 1943.3, calcd 1943.3), respectively. The primary oxidized product, **7**, was a mixture of at least two molecules with molecular weights of 3405.1 (**7a**) and 3423.4 (**7b**). According to the literature[20,21] and its subsequent transformation to **8** and **9**, **7a** contained 5-methylcytosine glycol ([M−H]$^-$, calcd 3405.2) and/or thymine glycol ([M−H]$^-$, calcd 3406.2), and **7b** contained the more oxidized product, *N*-(2-deoxyribosyl)-*N*′-methyltartronylurea ([M−H]$^-$, calcd 3423.2). In contrast to the high reactivity of the flipped-out ^{5m}C, the reaction of the flipped-out C and the ^{5m}C in the duplex with potassium permanganate was remarkably slow (Figure S11, Supporting Information), as reported for the deoxycytidine.[20]

The results described here showed that ligand-induced base flipping offers convenient and powerful methods for the synthesis of modified cytosines and 5-methylcytosines related to the bisulfite sequencing at the predetermined site of DNA. Facile access to these oligomers with modified cytosines provides opportunities to study the transformtion of C* in detail at a molecular level.

Acknowledgment. This work was supported by Grant in Aid for Scientific Research (S) (18105006) from the Japan Society for the Promotion of Science (JSPS). This work was supported in part by Special Coordination Funds for Promoting Science and Technology from the Ministry of Education, Culture, Sports, Science and Technology for the Osaka University Program for the Support of Networking among Present and Future Women Researchers.

Supporting Information Available: Experimental details and additional HPLC and MS data. This material is available free of charge via the Internet at http://pubs.acs.org.

OL900195Z

(17) Nakatani, K.; Hagihara, S.; Sando, S.; Miyazaki, H.; Tanabe, K; Saito, I. *J. Am. Chem. Soc.* **2000**, *122*, 6309–6310.
(18) Tanaka, K.; Tainaka, K.; Umemoto, T.; Nomura, A.; Okamoto, A. *J. Am. Chem. Soc.* **2007**, *129*, 14511–14517.
(19) Bareyt, S.; Carell, T. *Angew. Chem., Int. Ed.* **2008**, *47*, 181–184.
(20) Iida, S.; Hayatsu, H. *Biochim. Biophys. Acta* **1971**, *240*, 370–375.
(21) Tremblay, S.; Wagner, J. R. *Nucleic Acids Res.* **2008**, *36*, 284–293.