ORGANIC LETTERS

2011
Vol. 13, No. 13
3446–3449

Syntheses of 5-Formyl- and 5-Carboxyl-dC Containing DNA Oligos as Potential Oxidation Products of 5-Hydroxymethylcytosine in DNA

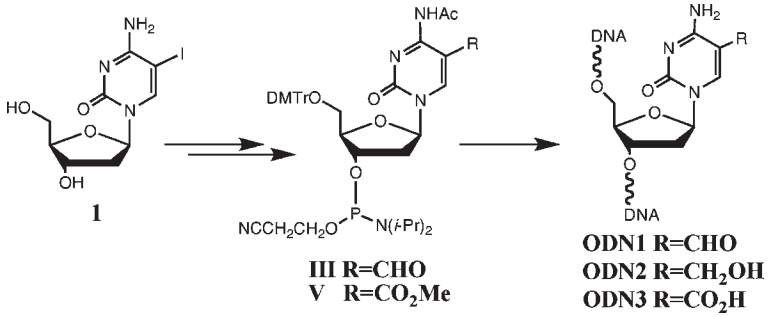
Qing Dai* and Chuan He

Department of Chemistry and Institute for Biophyscial Dynamics, The University of Chicago, 929 East 57th Street, Chicago, Illinois 60637, United States

daiqing@uchicago.edu

Received May 4, 2011

ABSTRACT



To investigate the potential oxidation products of 5-hydroxymethylcytosine (5-hmC)-containing DNA, we present here efficient syntheses of 5-formyl- and 5-methoxycarbonyl-2′-deoxycytidine phosphoramidites. The 5-formyl group in III was easy to introduce and was compatible with phosphoramidite and DNA syntheses. An additional treatment of ODN1 with $NaBH_4$ produced the corresponding ODN2 quantitatively. Phosphoramidite V was also incorporated into DNA, and the methyl ester could be hydrolyzed under mild basic conditions to afford ODN3.

As one of the best characterized epigenetic modifications, 5-methylcytosine (5-mC) plays an important role in many biological processes that include regulation of gene expression, genomic imprinting, X chromosome inactivation, and suppression of transposable element propagation.[1] Important progress in recent studies of DNA modification is the discovery of 5-hmC (the "sixth" base) in the mammalian genome, which is produced through the oxidation of 5-mC by a group of TET (ten eleven translocation) dioxygenases.[2] Recent studies have shown that 5-hmC is a widespread DNA modification in various tissues and cells.[3,4] We have developed a chemical-labeling method to enrich and sequence the distribution of 5-hmC in the mammalian genome.[4] Despite such exciting progress, little is known about how 5-hmC is metabolized in vivo. It has been hypothesized that the TET protein-catalyzed conversion of 5-mC DNA to 5-hmC may represent the first step of multiple-step reactions for DNA demethylation (Figure 1).[5] Further oxidation of 5-hmC in DNA may result in the formation of 5-formylcytosine (5-fC) DNA. It is possible that 5-fC is a transient intermediate, which can be quickly converted to 5-carboxylcytosine (5-caC) in vivo and cytosine by further decarboxylation (Figure 1).[6] Thus, it is important to

(1) (a) Feng, S.; Jacobsen, S. E.; Reik, W. *Science* **2010**, *330*, 622. (b) Ooi, S. K.; O'Donnell, A. H.; Bestor, T. H. *J. Cell Sci.* **2009**, *122*, 2787.

(2) (a) Kriaucionis, S.; Heintz, N. *Science* **2009**, *324*, 929. (b) Tahiliani, M.; Koh, K. P.; Shen, Y.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A. *Science* **2009**, *324*, 930. (c) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129.

(3) (a) Münzel, M.; Globisch, D.; Brückl, T.; Wagner, M.; Welzmiller, V.; Michalakis, S.; Müller, M.; Biel, M.; Carell, T. *Angew. Chem., Int. Ed.* **2010**, *49*, 5375. (b) Szwagierczak, A.; Bultmann, S.; Schmidt, C. X.; Spada, F.; Leonhardt, H. *Nucleic Acids Res.* **2010**, *38*, 181.

(4) Song, C.-X.; Szulwach, K. E.; Fu, Y.; Dai, Q.; Yi, C.; Li, X.; Li, Y.; Chen, C.-H.; Zhang, W.; Jian, X.; Wang, J.; Zhang, L.; Looney, T. J.; Zhang, B.; Godley, L. A.; Hicks, L. M.; Lahn, B. T.; Jin, P.; He, C. *Nat. Biotechnol.* **2011**, *29*, 68.

(5) Wu, S. C.; Zhang, Y. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 607.

(6) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90.

10.1021/ol201189n
Published on Web 06/07/2011

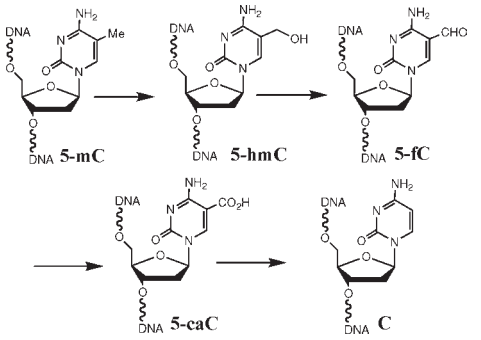


Figure 1. Possible demethylation pathway of 5-mC through 5-hmC, 5-fC, and 5-caC.

explore efficient and convenient syntheses of 5-fC and 5-caC-containing DNA as models for biological studies. These modified nucleosides can also serve as in vitro authentic standards for searching for 5-fC and 5-caC modifications in genomic DNA. Herein, we report the syntheses of two phosphoramidite building blocks **III** and **V** and their incorporation into DNA.

A literature search revealed that the synthesis of 5-fC-containing DNA had been accomplished by incorporating phosphoramidite **I**[7] into DNA (Figure 2). The structure of **I** does not contain 5-formyl functionality, but a postsynthetic NH_4OH treatment removes the acetyl protecting groups to yield a diol analogue, which is subsequently cleaved by treatment with $NaIO_4$ to generate the 5-formyl functionality. Although this approach can be used to prepare 5-fC in DNA, the synthesis of **I** is inconvenient. A ribo version phosphoramidite of 5-fC bearing an unprotected 5-formyl functionality (Figure 2, II) has recently been synthesized and incorporated into RNA,[8] suggesting that an unprotected 5-formyl group may be compatible with oligo synthesis. Therefore, we sought to synthesize phosphoramidite **III** (Figure 2) as the building block to prepare 5-fC in DNA.

5-Formyl-2′-deoxycytidine has been prepared by direct oxidation of 5-methyl-2′-deoxycytidine with $K_2S_2O_8$, but the yield is rather low.[9] Catalanotti et al. reported that 3′,5′-silyl-protected thymidine analogue can be oxidized to the corresponding 5-formyl-2′-deoxyuridine in a reasonable yield,[10] but its conversion to the corresponding 5-formyl-2′-deoxycytidine analogue requires multiple steps including possible protection of the 5-formyl functionality. We chose to prepare 5-formyl-2′-

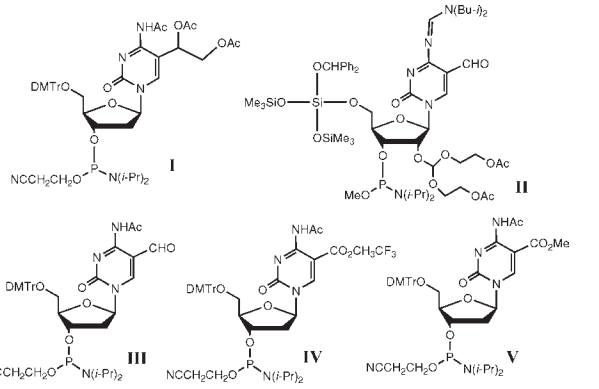
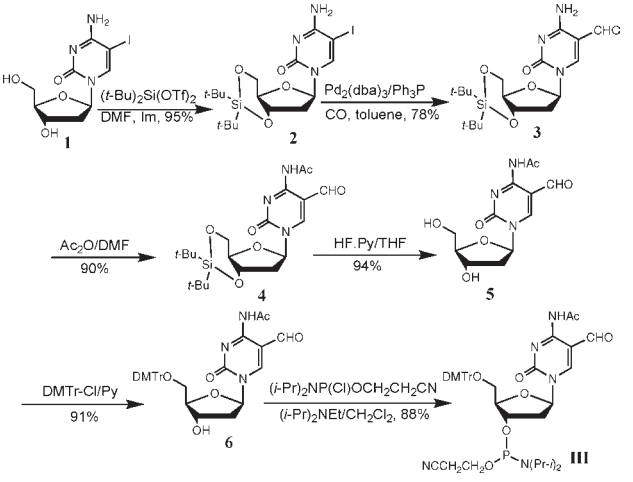


Figure 2. Structures of phosphoramidites for synthesis of DNA or RNA containing 5-fC or 5-caC modifications.

deoxycytidine using the Stille reaction to introduce the 5-formyl group.[11] Thus, the 3′- and 5′-hydroxyls of the commercial 5-iodo-2′-deoxycytidine (**1**) were first protected with di-*tert*-butylsilyl group to give **2** in 95% yield (Scheme 1).[12] Pd-catalyzed Stille reaction of **2** afforded 5-formyl-dC analogue **3** in 78% yield. Protection of the 4-amino group by treating **3** with acetic anhydride in DMF furnished **4** in 90% yield. Removal of the 3′,5′-silyl protecting group by treating **4** with hydrogen fluoride–pyridine gave **5** in 94% yield. To test the compatibility of the 5-formyl group during DNA synthesis and the subsequent deprotection conditions, we treated **5** with the deblocking, coupling, capping, and the oxidation reagents that are used for DNA synthesis, respectively. TLC showed that **5** stayed intact in all cases, confirming that the 5-formyl group could survive DNA synthesis. When we treated **5** under mild deprotection conditions (0.1 M K_2CO_3 in MeOH for 2 h at rt), quantitative conversion of **5** to 5-formyl-2′-deoxycytidine free nucleoside was observed, confirming that **5** is

Scheme 1. Synthesis of Phosphoramidite **III**



(7) Karino, N.; Ueno, Y.; Matsuda, A. *Nucleic Acids Res.* **2001**, *29*, 2456.

(8) Lusic, H.; Gustilo, E. M.; Vendeix, F. A. P.; Kaiser, R.; Delaney, M. O.; Graham, W. D.; Moye, V. A.; Cantara, W. A.; Agris, P. F.; Deiters, A. *Nucleic Acids Res.* **2008**, *36*, 6548.

(9) (a) Itahara, T.; Yoshitake, T.; Koga, S.; Nishino, A. *Bull. Chem. Soc. Jpn.* **1994**, *67*, 2257. (b) Itahara, T. *Chem. Lett.* **1991**, *9*, 1591.

(10) Catalanotti, B.; Galeone, A.; Mayol, L.; Oliviero, G.; Rigano, D.; Varra, M. *Nucleosides, Nucleotides Nucleic Acids* **2001**, *20*, 1831.

(11) Munzel, M.; Globisch, D.; Trindler, C.; Carell, T. *Org. Lett.* **2010**, *12*, 5671.

(12) Serebryany, V.; Beigelman, L. *Tetrahedron Lett.* **2002**, *43*, 1983.

also compatible with the mild deprotection conditions. Protection of 5′-OH of **5** with DMTr (91%) followed by 3′-phosphitylation afforded the phosphoramidite **III** (88%). The synthesis entails six steps in 50% overall yield starting from **1**.

Phosphoramidite **III** was then incorporated into a model sequence 5′-TCXGA (X = 5-fC) using ultramild reagents with the modified phosphoramidite **III** using double coupling to give resin-attached **ODN1**. We divided the resin into two parts and deprotected them with the following two procedures, respectively. The first part was treated with 0.1 M K_2CO_3 in MeOH/H_2O (1:1 v/v) at rt for 2 h followed by addition of AcOH to neutralize the final solution to neutral pH; the second part was treated with concentrated NH_4OH at rt for 2 h. The two samples were then subjected to reverses-phase HPLC analysis. In both cases, the desired 5mer was produced as the only major product (peak a, Figure 3A,B). The small peak b in Figure 3B is probably the imine intermediate; this peak disappeared completely after overnight standing, suggesting that the imine intermediate was cleanly hydrolyzed into aldehyde (Figure 3C).

Recently, we reported the synthesis of 5-hmC phosphoramidite with TBDMS as a 5-CH_2OH protecting group.[13] Although the synthesis is highly efficient and entailed eight steps in 32% overall yield, the generation of 5-CH_2OH was accomplished through the reduction of the corresponding 5-CHO analogue, and the resulting 5-CH_2OH was subsequently protected as a TBDMS ether. Since 5-fC phosphoramidite **III** could be synthesized more efficiently in six steps in 50% overall yield and it is completely compatible with DNA synthesis and the subsequent deprotection, we reasoned that 5-fC-containing DNA might be converted to 5-hmC-containing DNA by postsynthetic reduction so that two steps (introduction and removal of TBDMS) could be spared.

To facilitate isolation of the DNA, we chose to reduce the 5-fC-containing DNA when it was still attached with resin. Thus, resin-attached **ODN1** was treated with 0.1 M $CeCl_3 \cdot 7H_2O$ in MeOH (1 mL) followed by addition of 1 mg of $NaBH_4$. After 15 min, the resin was treated with the same two mild conditions as those used for deprotection of **ODN1** to give **ODN2**. The HPLC analysis showed that in both cases the peak of 5-fC-containing oligo **ODN1** was shifted to peak c (Figure 3D,E), which is the corresponding 5-hmC-containing oligo **ODN2**. This new method provides a more efficient and convenient approach for synthesis of 5-hmC-containing DNA.

To prepare 5-caC in DNA, we searched for the oxidation conditions that can postsynthetically oxidize the 5-formyl group of 5-fC in DNA to the corresponding 5-carboxyl functionality. Although we have screened various oxidizing reagents, we have yet to identify the proper oxidizing conditions. As an alternative approach, we sought to synthesize the phosphoramidite that contains a 5-ester group that can be hydrolyzed to 5-carboxylic acid in postsynthetic deprotection. Nomura et al. reported the synthesis of the phosphoramidite **IV** (Figure 2) with a 5-trifluoroethyl ester functionality.[14] After incorporation into DNA, it was used to react with amines to generate amide analogues with trifluoroethoxy as the leaving group. We prepared phosphoramidite **IV** by following the literature procedure and found that, for unknown reasons, the palladium-catalyzed carbonylation reaction always led to formation of a significant amount of reduced byproduct, N^4-Ac-2′-deoxycytidine. In addition, we found that the trifluoroethyl ester could not be hydrolyzed efficiently under mild basic conditions. After phosphoramidite **IV** was incorporated into a short oligo, 5-TCXGA (**ODN3**), and deprotected with 0.1 M K_2CO_3 in 1:1 MeOH/H_2O overnight at rt, HPLC analysis showed that only 65% of the fully deprotected DNA (peak d, Figure 3F) was obtained in addition to another 35% oligo (peak e, Figure 3F) still containing trifluoroethyl ester.

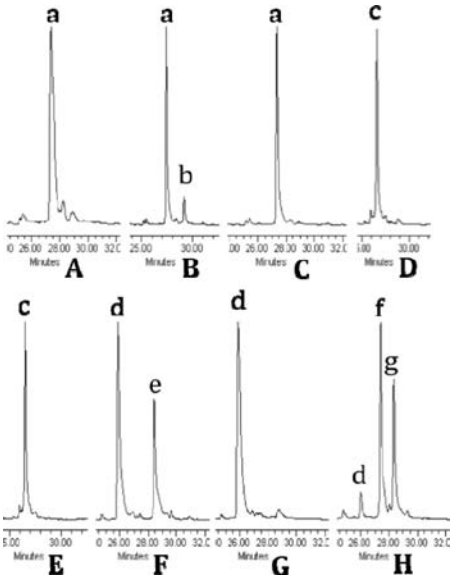


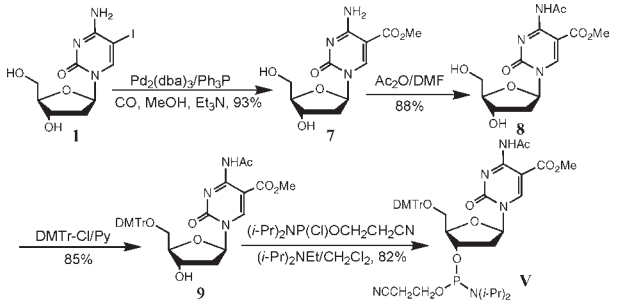
Figure 3. (A) **ODN1** was deprotected by treatment of 0.1 M K_2CO_3 in MeOH/H_2O (1:1 v/v) at rt for 2 h. Peak a is the fully deprotected 5mer, $[MH]^+ = 1491$. (B) **ODN1** was deprotected by treatment of NH_4OH at rt for 2 h; peak b is the proposed imine intermediate. (C) **ODN1** obtained in the same way as (B) was dissoved in H_2O and allowed to stand overnight at rt, after which peak b disappeared. (D) **ODN2** was deprotected by treatment of 0.1 M K_2CO_3 in MeOH/H_2O (1:1 v/v) at rt for 2 h; peak c is the fully deprotected 5mer, $[MH]^+ = 1493$. (E) **ODN2** was deprotected by NH_4OH at rt for 2 h. (F) **ODN3** made from phoshoramidite **IV** was deprotected by treatment of 0.1 M K_2CO_3 in MeOH/H_2O (1:1 v/v) at rt overnight. Peak d is the fully deprotected 5mer, $[MH]^+ = 1507$; peak e is the 5mer with 5-trifluoroethoxycarbonyl-dC modification, $[MH]^+ = 1589$. (G) **ODN3** made from phoshoramidite **V** was deprotected by treatment of 0.1 M K_2CO_3 in MeOH/H_2O (1:1 v/v) at 40 °C overnight. (H) **ODN3** made from phoshoramidite **V** was treated with NH_4OH at rt for 2 h. Peak f is the 5mer with a 5-formamide modification, $[MH]^+ = 1506$; peak g is the 5mer with a 5-methoxycarbonyl-dC modification, $[MH]^+ = 1521$.

(13) Dai, Q.; Song, C.-X.; Pan, T.; He, C. *J. Org. Chem.* **2011**, *76*, 4182.

(14) Nomura, Y.; Haginoya, N.; Ueno, Y.; Matsuda, A. *Bioorg. Med. Chem. Lett.* **1996**, *6*, 2811.

Our past experience that methyl ester could be readily hydrolyzed under mild basic conditions prompted us to synthesize the corresponding phosphoramidite **V** (Figure 2). Our initial attempt to introduce the 5-methoxycarbonyl group into dC by treating 5-iodo-N^4-2′-deoxycytidine with CO in the presence of Et_3N, $(PhCN)_2PdCl_2$, and MeOH in DMF at 50 °C, similar to the procedure used previously for preparation of the corresponding CF_3CH_2 ester,[14] gave only a very low yield of the desired methyl ester. The byproduct, N^4-acetyl-2′-deoxyctidine, was always isolated in high yields. However, we found that when 5-iodo-dC (**1**) was used as substrate and $Pd_2(dba)_3$ as catalyst, **1** was quantitatively converted into the corresponding methyl ester **7** and no reduction byproduct was detected as indicated by TLC (93% isolated yield, Scheme 2). Protection of the 4-amino group with Ac gave **8** in 88% yield. Protection of 5′-hydroxyl of **8** with DMTr (85%) followed by 3′-phosphitylation produced the phosphoramidite **V** (82%). The synthesis entails four steps from **1** in 57% overall yield.

Scheme 2. Synthesis of Phosphoramidite **V**



Phosphoramidite **V** was then incorporated into the sequence 5′-TCXGA (X = 5-methoxycarbonyl-dC) to give the polymer-attached **ODN3**. When the resin was treated under mild basic conditions (0.1 M K_2CO_3 in $MeOH/H_2O$ overnight at 40 °C), HPLC showed that the methyl ester was cleanly hydrolyzed to afford 5-caC-containing DNA as the major product (peak d, Figure 3G). In contrast, when polymer-attached **ODN3** was treated under the other mild conditions (concentrated NH_4OH at rt for 2 h), HPLC showed that the major product is an amide (peak f, Figure 3H) with only a small amount of desired 5-caC in DNA formed (peak d, Figure 3H), suggesting that NH_4OH treatment is not suitable in this case.

In order to further confirm the application of phosphoramidites **III** and **V**, we incorporated them into 20mer oligos (5′-TTTCAGCTCXGGTCACGCTC). Using the same deprotection conditions, we purified the desired modified DNA oligos by HPLC and characterized them by MALDI-Tof MS (X = 5-fC, $[MK]^+$ = 6080; X = 5-caC, $[MK]^+$ = 6096).

In summary, we have achieved an efficient synthesis of the phosphoramidite of 5-formyl-2′-deoxycytidine **III** without protection of the 5-formyl group. The synthesis entails six steps in 50% overall yield. The 5-formyl group can survive all transformations during the phosphoramidite and subsequent DNA syntheses and two different mild deprotection conditions to give 5-fC-containing DNA **ODN1**. We also demonstrate that an additional treatment of the resin-attached 5-fC-containing DNA with $NaBH_4$ could postsynthetically convert the 5-fC-containing DNA into the corresponding 5hmC-containing DNA quantitatively, which provides a more convenient alternative synthesis of 5-hmC-containing DNA. We have also developed an efficient synthesis of the 5-caC phosphoramidite **V** with methyl ester as the protecting group in four steps in 57% overall yield and demonstrated that **V** is compatible with subsequent DNA synthesis. We believe these synthetic procedures will provide sufficient materials for future investigation of 5-hmC-containing DNA and 5-fC- and 5-caC-containg DNA oligos in biological investigations.

Acknowledgment. This work is supported by the NIH (GM071440 and GM088599). We thank C. Song and S. F. Reichard at the University of Chicago for helping with oligo synthesis and editing the manuscript, respectively.

Supporting Information Available. Analytical data for compounds **2**–**9**, **III**, and **V** including their ^1H and ^{13}C NMR spectra. This material is available free of charge via the Internet at http://pubs.acs.org.