ORGANIC LETTERS

2010
Vol. 12, No. 24
5671−5673

Efficient Synthesis of 5-Hydroxymethylcytosine Containing DNA

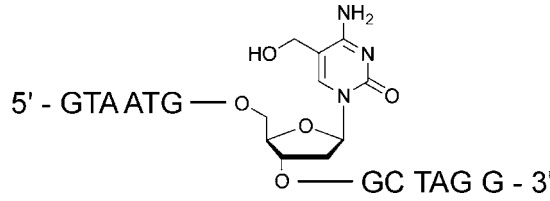
Martin Münzel, Daniel Globisch, Christian Trindler, and Thomas Carell*

Center for Integrated Protein Science, Department of Chemistry, Ludwig-Maximilians-University Munich, Butenandtstr. 5-13, 81377 Munich, Germany

thomas.carell@cup.uni-muenchen.de

Received October 6, 2010

ABSTRACT



5-Hydroxymethylcytosine ($^{5\text{-HOMe}}$dC) was recently discovered as the sixth base in the mammalian genome. The development of a new phosphoramidite building block is reported, which allows efficient synthesis of $^{5\text{-HOMe}}$dC containing DNA. Key steps of the synthesis are a palladium-catalyzed formylation and the simultaneous protection of a hydroxyl and amino group as a cyclic carbamate. DNA synthesis is possible under standard conditions, and deprotection can be carried out with dilute NaOH.

The genetic material is constructed from the four canonical bases dA, dC, dG, and dT. The dC base is furthermore subject to epigenetic modification. In eucaryotes, the dC base is often methylated at position C5 to give the base 5-methylcytosine ($^{5\text{-Me}}$dC).[1] Most recently, a new dC-based modification was discovered, which contains instead of the methyl group at position C5 a hydroxymethylene group (Figure 1).[2,3] Others and we were able to show that hydroxymethylcytosine is a widespread DNA modification in the brain and that levels vary depending on tissue type.[4,5] The function of the new "sixth" base $^{5\text{-HOMe}}$dC is currently not clear. However, it was shown that particular oxoglutarate-dependent TET (ten-to-eleven translocation) oxidases are responsible for its formation.[3,6] These enzymes specifically oxidize the 5-methyl group of $^{5\text{-Me}}$dC to give $^{5\text{-HOMe}}$dC. Recently, it was discovered

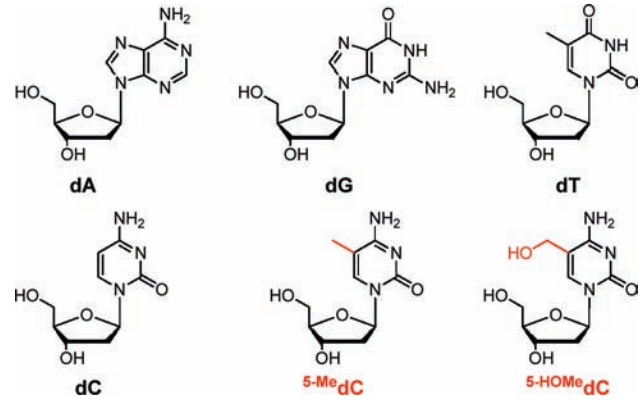


Figure 1. Nucleosides present in the mammalian genome.

that a lack of these TET enzymes yields misfunctioning stem cells providing a link between formation of the base $^{5\text{-HOMe}}$dC and cellular development.[7]

To facilitate the biochemical investigation of $^{5\text{-HOMe}}$dC-dependent biological processes, an efficient synthesis of

(1) Law, J. A.; Jacobsen, S. E. *Nat. Rev. Genet.* **2010**, *11*, 204–220.
(2) Kriaucionis, S.; Heintz, N. *Science* **2009**, *324*, 929–930.
(3) Tahiliani, M.; Koh, K. P.; Shen, Y. H.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Iyer, L. M.; Liu, D. R.; Aravind, L.; Rao, A. *Science* **2009**, *324*, 930–935.
(4) Münzel, M.; Globisch, D.; Brückl, T.; Wagner, M.; Welzmiller, V.; Michalakis, S.; Müller, M.; Biel, M.; Carell, T. *Angew. Chem., Int. Ed.* **2010**, *49*, 5375–5377.
(5) Szwagierczak, A.; Bultmann, S.; Schmidt, C. S.; Spada, F.; Leonhardt, H. *Nucleic Acids Res.* **2010**, *38*, e181.
(6) Loenarz, C.; Schofield, C. J. *Chem. Biol.* **2009**, *16*, 580–583.
(7) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129–1133.

10.1021/ol102408t © 2010 American Chemical Society
Published on Web 11/17/2010

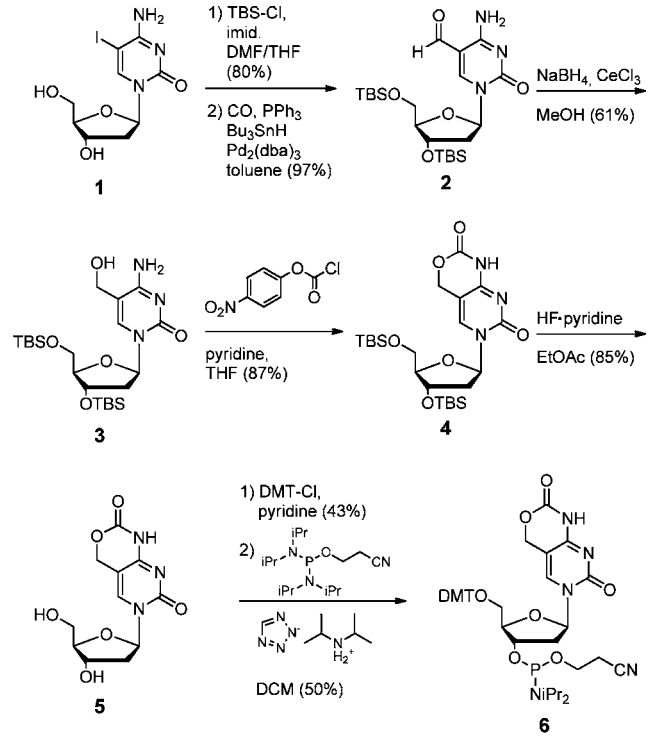
$^{5\text{-HOMe}}$dC-containing oligodeoxynucleotides (ODNs) is needed. The building block currently used for the synthesis of $^{5\text{-HOMe}}$dC containing DNA strands necessitates a rather tedious chemical synthesis via an unstable bromothymidine intermediate.[8,9] In addition, deprotection of the embedded $^{5\text{-HOMe}}$dC unit in our hands required heating of the synthesized oligonucleotides for 60 h at 60 °C with conc. ammonia,[9] which prohibits any derivatization of the oligonucleotides with fluorescence or biotin labels typically needed for many biochemical experiments.

Here we report the development of a novel $^{5\text{-HOMe}}$dC phosphoramidite building block, available in just seven steps from the stable and commercially available 5-iodo-deoxycytidine **1**. We found that the building block enables synthesis of $^{5\text{-HOMe}}$dC containing DNA strands using standard phosphoramidite chemistry with excellent coupling yield. Furthermore, deprotection of the new building block was achieved with just dilute NaOH solution over 12 h at room temperature.

We chose to protect the amino and hydroxy groups of the $^{5\text{-HOMe}}$dC base as a cyclic carbamate. This group elegantly inactivates both nucleophilic groups of $^{5\text{-HOMe}}$dC and is one of the smallest possible protective groups, therefore allowing efficient coupling in the DNA synthesizer. Furthermore, it can be easily deprotected simultaneously with cleavage of the DNA strand from the resin by simple base treatment in one step.

The synthesis is depicted in Scheme 1. The starting point is 5-iododeoxycytidine **1**,[10] which was reacted with TBS-Cl to protect the hydroxyl groups. The further synthesis can alternatively be carried out without OH protection; however, the yields of the following reactions turned out be lower, and the purification is more tedious. To insert the hydroxymethylene group, we chose to utilize a Pd-catalyzed formylation reaction with CO. This reaction turned out to be extremely efficient even in the presence of the unprotected exocyclic amino group, providing **2** in yields of above 95%. We next reduced the obtained formyl group at C5 with $NaBH_4$ to obtain compound **3**. For this step application of Luche conditions is absolutely crucial.[11] Without the addition of $CeCl_3$ the hydride presumably adds to the extremely electrophilic C6 position of the base resulting in decomposition of the starting material **2**. To introduce the cyclic carbamate, compound **3** was treated with 4-nitrophenylchloroformate[12] to give the protected compound **4** in very good yield. Subsequent deprotection of the silyl groups was achieved with HF in pyridine. In ethylacetate as solvent, the diol **5** precipitates after completion of the deprotection allowing its isolation by simple centrifugation. We finally converted **5** into the $^{5\text{-HOMe}}$dC phosphoramidite building block **6** using standard procedures.[13]

Scheme 1. Synthesis of the Phosphoramidite Building Block **6** (C* = $^{5\text{-HOMe}}$dC)



ODN1: 5' - GTA ATG C*GC TAG G - 3'

DNA synthesis was performed using standard procedures. To develop the appropriate deprotection method, we prepared ODN1 (Scheme 1) using the phosphoramidite **6**. Coupling times with **6** were doubled to allow efficient incorporation into the oligonucleotide chain. Initial attempts to deprotect the strands with a standard protocol (conc. ammonia at room temperature overnight) furnished oligonucleotides containing $^{5\text{-HOMe}}$dC. However, the urea derivative **7** and the aminomethyl-dC nucleobase **8** were formed as major byproducts (Figure 2). To prevent these undesired side reactions, we

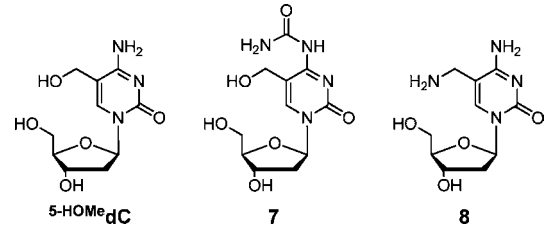


Figure 2. Nucleosides **7** and **8** were obtained after deprotection of the oligonucleotides using standard NH_3-based conditions. Deprotection with NaOH, however, yields exclusively $^{5\text{-HOMe}}$dC.

used 0.4 M NaOH in MeOH/H_2O 4:1 as the deprotection solution at room temperature overnight. Cleavage of the DNA strand from the solid support and deprotection of all

(8) Shiau, G. T.; Schinazi, R. F.; Chen, M. S.; Prusoff, W. H. *J. Med. Chem.* **1980**, *23*, 127–133.
(9) Tardy-Planechaud, S.; Fujimoto, J.; Lin, S. S.; Sowers, L. C. *Nucleic Acids Res.* **1997**, *25*, 553–558.
(10) Hwang, C. H.; Park, J. S.; Won, J. H.; Kim, J. N.; Ryu, E. K. *Arch. Pharm. Res.* **1992**, *15*, 69–72.
(11) Luche, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 2226–2227.
(12) Sammet, B. *Synlett* **2009**, 3050–3051.

bases including the cyclic carbamate were achieved under these conditions, yielding a DNA strand that exclusively contained $^{5\text{-HOMe}}$dC. Interestingly, related urea derivatives were proven to be stable during DNA synthesis and deprotection.[14]

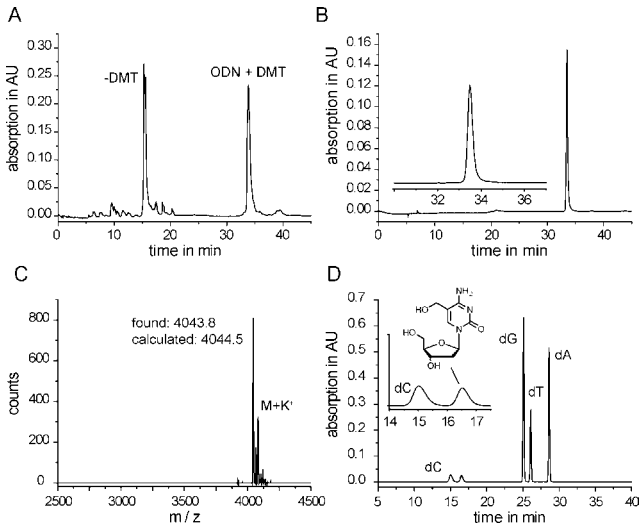


Figure 3. (A) Reversed-phase HPLC chromatogram directly after cleavage from the resin (0−50% buffer B in 45 min). (B) Reversed-phase HPLC chromatogram after cleavage of the DMT group and purification (0−20% buffer B in 45 min). (C) MALDI spectrum of the purified strand ODN1. (D) Digest of purified DNA strand ODN1 (for buffers see the Supporting Information).

Figure 3A depicts the raw HPLC chromatogram obtained directly after DNA cleavage and deprotection. The spectrum shows that the building block **6** indeed couples with high efficiency during DNA assembly in the synthesizer. Figure 3B shows the reversed-phase HPLC chromatogram of the purified $^{5\text{-HOMe}}$dC containing oligonucleotide together with the MALDI-TOF mass spectrum (Figure 3C) proving the correct incorporation of $^{5\text{-HOMe}}$dC into the DNA strand. This is noteworthy because we observed S_N2-type reactions at the pseudobenzylic position of $^{5\text{-HOMe}}$dC especially under acidic conditions or when the oxygen atom was derivatized with an electron-withdrawing group. The unusually high reactivity of the primary OH group initially hampered our attempts to protect $^{5\text{-HOMe}}$dC as a bisacetate. This reactivity also explains the formation of the byproduct **8** in the reaction with NH_4OH.

To gain further evidence for the exclusive formation of $^{5\text{-HOMe}}$dC, we conducted an enzymatic digestion study. To this end, we treated the obtained oligonucleotide ODN1 first with nuclease S1 for 3 h at 37 °C followed by incubation with antarctic phosphatase and snake venom phosphordiesterase for an additional 3 h at 37 °C. The obtained nucleoside mixture was analyzed by HPLC-ESI-MS. The chromatogram is depicted in Figure 3D and shows besides the four canonical bases dA, dC, dG, and dT an additional signal, which features the correct molecular weight for $^{5\text{-HOMe}}$dC. The high-resolution MS data support the molecular formula $C_{10}H_{15}N_3O_5$ expected for the target compound (see Supporting Information).

In summary, we report a short and efficient synthesis of a novel $^{5\text{-HOMe}}$dC phosphoramidite building block. The key step in the synthesis is a Pd(0)-catalyzed formylation and the simultaneous protection of the primary hydroxyl group together with the exocyclic amino group at the heterocycle as a cyclic carbamate. Deprotection of this unit is conveniently achieved with NaOH solution at mild conditions now enabling the synthesis of $^{5\text{-HOMe}}$dC oligonucleotides containing additional modifications such as fluorophores or biotin labels. For these purposes, the here reported chemistry in combination with the new ability to perform Cu(I)-catalyzed click modification on DNA[15,16] should be particularly applicable.

Acknowledgment. We thank Jessica Steinbacher (LMU Munich) for preparative work. Financial support was obtained from the DFG (CA 275/8-4). M.M. thanks the Fonds of the Chemical Industry for a Kekulé fellowship.

Supporting Information Available: Information concerning the preparation of **1**−**6**, oligonucleotide synthesis and purification, enzymatic digestion, and detailed LC-MS data. This material is available free of charge via the Internet at http://pubs.acs.org.

OL102408T

(13) Caruthers, M. H. *Acc. Chem. Res.* **1991**, *24*, 278–284.
(14) Miyata, K.; Tamamushi, R.; Ohkubo, A.; Taguchi, H.; Seio, K.; Santa, T.; Sekine, M. *Org. Lett.* **2006**, *8*, 1545–1548.
(15) Gierlich, J.; Burley, G. A.; Gramlich, P. M. E.; Hammond, D. M.; Carell, T. *Org. Lett.* **2006**, *8*, 3639–3642.
(16) Gramlich, P. M. E.; Wirges, C. T.; Manetto, A.; Carell, T. *Angew. Chem., Int. Ed.* **2008**, *47*, 8350–8358.