Bioorganic & Medicinal Chemistry Letters 25 (2015) 1186–1191

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry Letters

journal homepage: www.elsevier.com/locate/bmcl

5-Hydroxymethylcytosine and 5-formylcytosine containing deoxyoligonucleotides: Facile syntheses and melting temperature studies

Shuguang Xuan [a], Qiong Wu [a], Liang Cui [b], Dawei Zhang [a], Fangwei Shao [a,*]

[a] Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore
[b] Interdisciplinary Research Group in Infectious Diseases, Singapore-MIT Alliance for Research & Technology (SMART), Singapore

ARTICLE INFO

Article history:
Received 11 December 2014
Revised 27 January 2015
Accepted 30 January 2015
Available online 7 February 2015

Keywords:
5-Formlycytosine
5-Hydroxymethylcytosine
Oligonucleotide synthesis
Melting temperature

ABSTRACT

An oxidation-based synthetic approach was developed for facile preparation of 5-formyl-2′-deoxycytidine and 5-hydroxymethyl-2′-deoxycytidine phosphoramidites. Upon introducing organic solvent components and copper catalysts, C5-methyl groups of 5-methyl-2′-deoxycytidine and thymidine were readily oxidized to formyl and hydroxyl functionality, respectively. Standard solid phase DNA synthesis and conventional deprotection methods were applicable to synthesize 5-formyl- or 5-hydroxymethylcytosine containing DNA oligonucleotides, which were used to study the effect of epigenetic modifications on DNA thermal dynamic stability.

5-hydroxymethylcytosine (hmC) as well as its oxidative products, such as 5-formylcytosine (fC) and 5-carboxylcytosine (CaC), was recently proposed to be a new series of epigenetic bases in mammalian genome.[1–3] The existence of these newly discovered epigenetic bases exhibit tissue specificity and show high levels in neurons and embryonic stem cells (ESCs).[4,5] In vivo genesis of hmC, fC and CaC is under the charge of a family of non-heme iron-dependent dioxygenases, ten-eleven translocation protein (TET).[2,3] 5-methylcytosine (mC) is oxidized by TETs to generate hmC and further to fC and 5-carboxylcytosine (CaC). This process is believed to be essential in active DNA demethylation.[6,7] Once converted to fC or CaC, the epigenetic cytosine becomes a substrate to thymine DNA glycosylase (TDG) and initiates the reinstallation of cytosine via base excision repair pathway.[7,8] Besides the pivotal roles in active DNA demethylation like mC, hmC is also considered as a stable epigenetic mark, which can be recognized by putative binding proteins and may have regulatory effects on downstream biological processes.[9–14] Nevertheless, the biological functions and metabolic mechanism of hmC and fC remain largely unknown. For investigation of hmC and fC associated biological processes, preparation of hmC and fC containing DNA oligonucleotides is of great importance. Herein, we present a facile synthetic route under mild aqueous conditions to prepare hmC and fC phosphoramidites from commercial available nucleosides as starting materials and to further proceed to solid phase DNA synthesis.

Though the high demands have attracted vigorous investigation on the synthesis of fC containing oligonucleotides, only two methods had been reported to prepare fC phosphoramidite for oligonucleotides synthesis. An early attempt was conducted by Karino et al., who firstly incorporated a 5-(1,2-dioxyethyl)cytosine as precursor into the synthetic oligonucleotides.[15] fC was converted in the oligonucleotide from diol functionality by post-synthetic $NaIO_4$ oxidation, which unfortunately limited the method only to the oligonucleotides without oxidation-labile components. Recently, He's and Carell's groups utilized a palladium-catalyzed CO insertion on 5-iodo-cytosine to introduce 5-formyl functionality.[16–18] However, high pressured explosive gas, complicated protocol and unconventional instrumental setup casted shadows on the method, though with the excellent yields.

In an attempt to develop facile methods to prepare 5-formyl-2′-deoxycytidine (fdC, **2**), we were inspired by the biological synthesis of fC via TET and considered the possibility of preparing **2** via oxidation of 5-methyl-2′-deoxycytidine (mdC, **1**). A literature survey revealed that **1** could be oxidized by $K_2S_2O_8$ to yield **2** in aqueous solution, however, only as a side product.[19,20] Nevertheless, the oxidation power and major products of $K_2S_2O_8$ can be tuned by solvent polarity and metal ions as catalysts.[21] After screening a few combinations of polar organic solvents as co-solvent with water, acetonitrile/water was found to be the optimal hosts for both organic precursor, **1**, and inorganic oxidant and catalyst in the

* Corresponding author. Fax: +65 65922511.
E-mail address: fwshao@ntu.edu.sg (F. Shao).

http://dx.doi.org/10.1016/j.bmcl.2015.01.070
0960-894X/

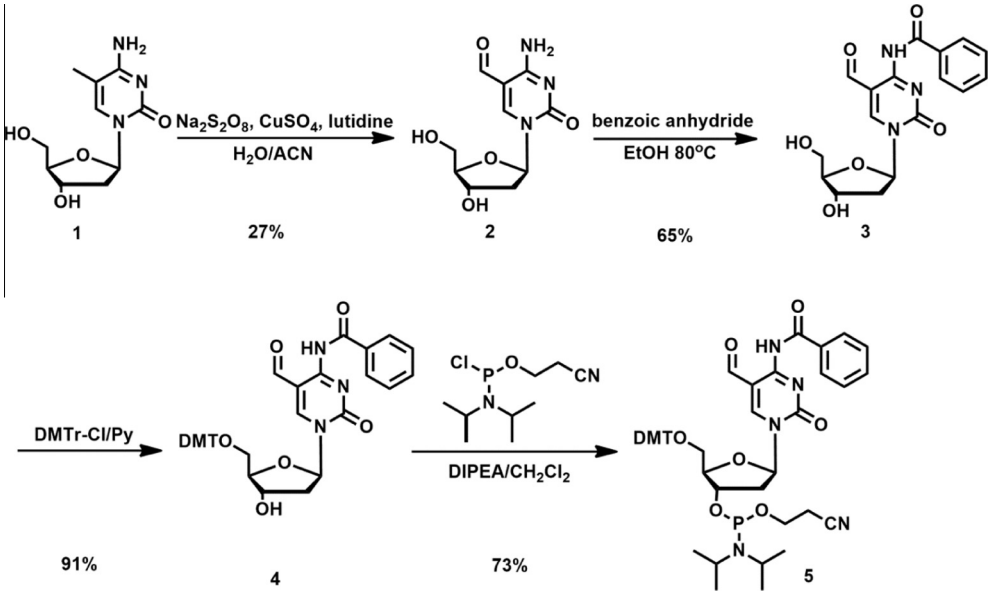
reaction mixture (Table S1). In the presence of catalytic amount of divalent copper ion, the major oxidative product of mdC shifted to fdC and generation of 5-hydroxymethyl-2′-deoxycytidine (hmdC) were not detectable. No significant improvement in yields was observed by using 3′,5′-proteced nucleosides, which indicated that hydroxyl groups on deoxyribose would neither decompose nor interfere the oxidation.[22] $Na_2S_2O_8$–Cu^{2+} can be directly applied to oxidize mC nucleoside and extra synthetic steps of introducing and removing protection groups can be avoided. The reaction successfully proceeded to completion either under microwave assisted heating (80 °C) or under conventional heating (65 °C) within regular RBF open to air. Reaction time was significantly reduced from 1 h to 5 min per reaction under microwave-assisted heating. At the end of reaction, **1** was completely consumed by radical oxidants and **2** was afforded as the only product with detectable UV absorption either on TLC or by HPLC, which made purification by silica column chromatography relatively simple. The synthetic pathway for fC phosphoramidite was developed as shown in Scheme 1. Since aldehyde can readily survive phosphoramidite based chemistry of oligonucleotide (ODN) synthesis, formyl group was not protected in the final fC phosphoramidite. We firstly, selectively protect the exocyclic amino group of fdC by reacting with benzoic anhydride in hot ethanol and generate **3** in moderate yield 65%. Compared to acetyl protecting group, benzoyl functionality is chemically more stable during solid phase synthesis of oligonucleotides, so that ultra-mild synthetic condition required by N^4-acetyl-fC phosphoramidite can be avoided.[17] In addition, **3** precipitated from reaction solvent and could be simply isolated from reaction mixture by centrifugation. Furthermore, N^4-protected fdC can also avoid the by-reactions with dimethoxytrityl chloride (DMT-Cl) if DMT was introduced to bare nucleoside. The subsequent DMT protection and 3′-phosphitylation were accomplished with decent yields of 91% and 73%, respectively. This synthetic route entails merely four steps to prepare fC phosphoramidite **5** and is more convenient in the aspect of experiment setup. fC phosphoramidite was then incorporated into oligonucleotide **ODN1** by standard solid phase DNA synthesis with extended coupling time (5 min) to ensure high synthetic yield. Decent coupling yield of fC phosphoramidite (78%) was achieved, according to the trityl graph of solid phase synthesis (data not shown). **ODN1** was cleaved from the resin by incubation in concentrated NH_4OH at 75 °C for 24 h. Simultaneously deprotection of natural nucleobases was accomplished during the incubation. Since fC phosphoramidite **5** uses the common benzoyl protection as regular cytosine phosphoramidites, no additional deprotection procedures are required for fC containing oligonucleotide. After HPLC purification before and after detritylation by 80% acetic acid at room temperature, the desired oligonucleotide **ODN1** (sequence in Table 1) was obtained and the mass of **ODN1** was verified by ESI-MS (Fig. S2). For further validation of successful incorporation of fC, **ODN1** was submitted to enzymatic digestion to yield free nucleosides and the resultant mixture was analysed by LC–MS. Similar as literature, 5-formyl-2′-deoxycytidine was eluted at the same retention time as thymidine from C18 column, and thus was not distinguishable by HPLC analysis (Fig. 1B).[16] To confirm the coexisting of fdC and dT in HPLC elutes at 16.6 min, the co-elutes were submitted to LC–MS spectroscopy for both baseline separation and mass characterization of fC and thymidine. Well resolved mass spectra of 5-formyl-2′-deoxycytidine (Fig. 1C, $[M+H]^+$ m/z for $C_{10}H_{14}N_3O_5^+$ 256.0912 and nucleobase fragment $C_6H_6N_3O_2^+$ 140.0445) and thymidine (Fig. S1, $[M+H]^+$ m/z for $C_{10}H_{15}N_2O_5^+$ 243.0916 and nucleobase fragment $C_5H_7N_2O_2^+$ 127.0498) were observed.

Table 1
Sequences of oligonucleotides

ODN	Sequences[a]
ODN1	5′-TTC CAC G***f*****C**G CGT TCC TGA CTG ACT C-3′
ODN2	5′-TTC CAC G***hm*****C**G CGT TCC TGA CTG ACT C-3′
ODN3	5′-TTC CAC G***C*****C**G CGT TCC TGA CTG ACT C-3′
ODN4	3′-AAG GTG C***G*****C** GCA ACC ACT GAC TGA G-5′
ODN5	3′-AAG GTG C***A*****C** GCA ACC ACT GAC TGA G-5′
ODN6	3′-AAG GTG C***C*****C** GCA ACC ACT GAC TGA G-5′
ODN7	3′-AAG GTG C***T*****C** GCA ACC ACT GAC TGA G-5′

[a] Epigenetic bases and complementary bases are highlighted in bold and italic.

The same $Na_2S_2O_8$–Cu^{2+} oxidant system were also applied to prepare hmC phosphoramidite (Scheme 2). Unlike fC phosphoramidite, hydroxymethyl group cannot survive the synthetic reactions of oligonucleotides and needs to be protected in phosphoramidite. In order to avoid the interference from N^4-exocylic amine during protection of hydroxymethyl group, we chose thymidine (dT, **6**), instead of mdC, as the starting substrate



Scheme 1. Synthetic route for fC phosphoramidite **5** building block. Synthetic yields are indicated under reaction condition of each step.

for $Na_2S_2O_8$ oxidation. As shown in Table S2, screening for reaction condition under microwave-assisted heating indicated that THF is the optimal organic solvent to prevent over oxidation of dT to 5-formyluridine and provide 5-hydroxymethyl-2′-deoxyuridine (hmdU, **7**) as the major product and with the highest conversion stoichiometry. Upon upscaling the reaction, **7** in THF/H_2O mixed solvent with $Na_2S_2O_8$/$CuSO_4$ as oxidant/catalyst mixture can achieve an apparent yield of 41% after recovering 34% of thymidine from chromatographic purification. Again, microwave assisted heating can reduce the reaction time from 4 h by conventional heating to only 5 min. The subsequent protection of C5 hydroxymethyl group was accomplished by substitution with 3-hydroxypropanenitrile to generate **8** with moderate yield 53%. Since installation of DMT protection on **8** can facilitate the amination step and generate cytosine analogue **10**, **8** was reacted with 3 equiv of 4,4′-dimethoxytrityl chloride (DMT-Cl) in pyridine and generate **9** with 81% yields. Conversion of **9** to **10** was accomplished by following a reported protocol.[23] Selective protection of the exocyclic amino group with benzoyl group was then achieved in 76% isolation yield by incubating with benzoic anhydride in anhydrous DMF for 4 days. The resultant compound **11** was 3′-phosphitylated to afford hmC phosphoramidite **12** in 59% isolation yield. The synthetic route requires six steps to prepare hmC phosphoramidite. A 25-mer hmC containing oligonucleotide, **ODN2** with the same sequence and epigenetic location as **ODN1** (Table 1), was synthesized. Incorporation of hmC in the DNA oligonucleotide was

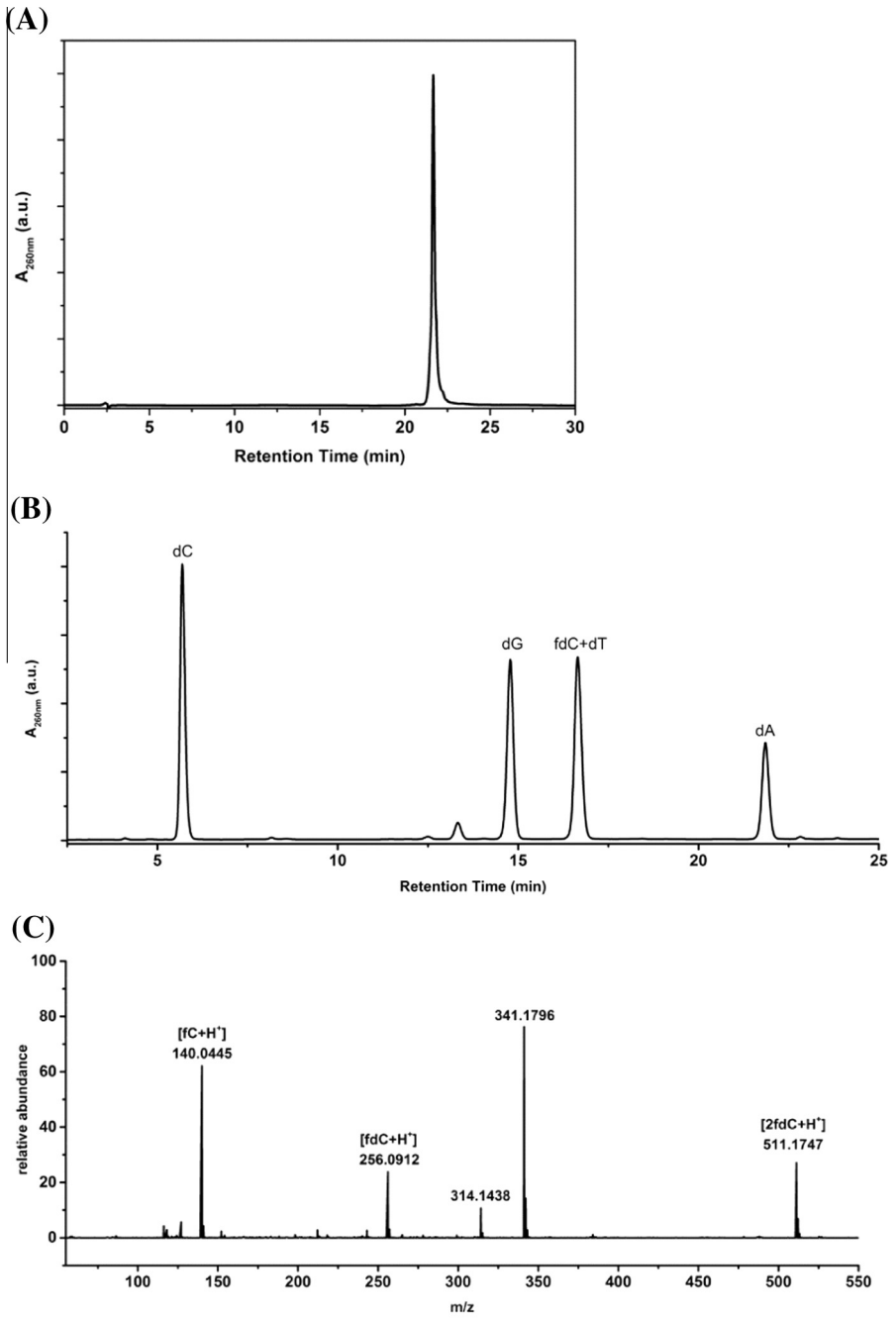
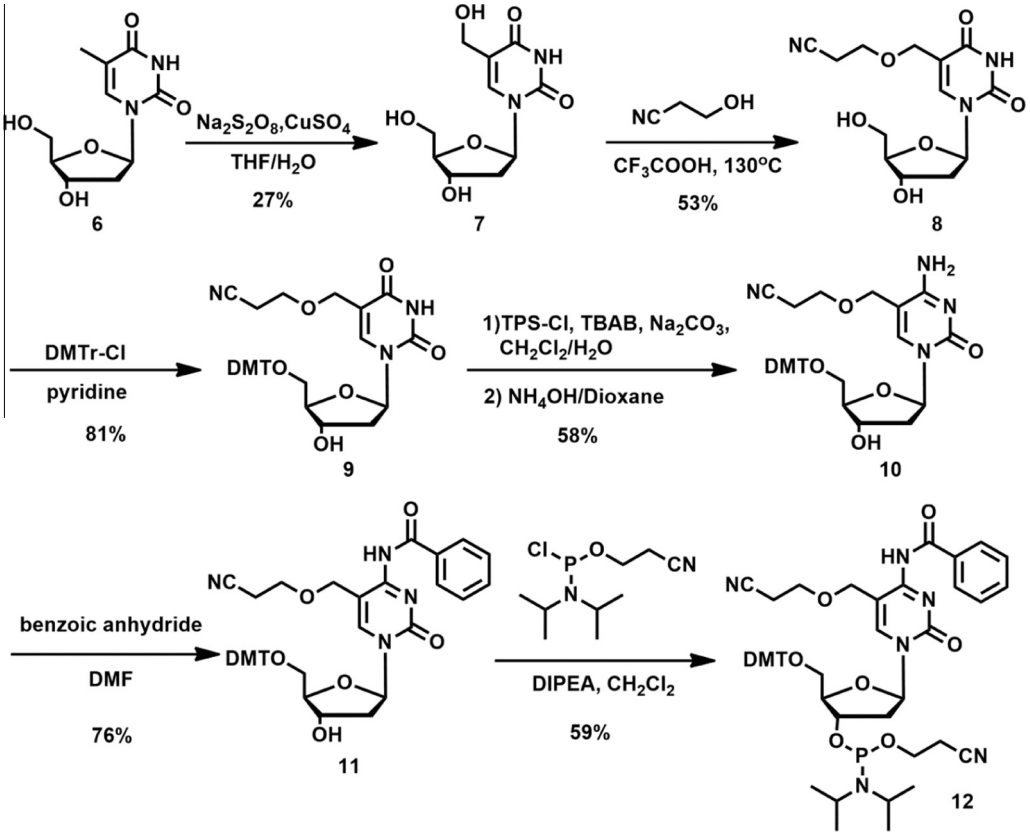


Figure 1. Reverse phase HPLC chromatography of **ODN1** after purification (A) and after enzymatic digestion (B). Elute peak at 16.6 min from (B) was submitted to LC–MS characterization for fdC (C).



Scheme 2. Synthetic route for hmC phosphoramidite **12**. Synthetic yields are indicated under the reaction conditions of each step.

successful by following standard solid phase synthetic protocol, while coupling time was extended to 5 min to ensure high coupling yields. hmC phosphoramidite had comparable coupling yield (88%) with canonical bases (data not shown). Following synthesis, cleavage and deprotection of **ODN2** was accomplished simultaneously by NH_4OH treatment at 75 °C for 24 h or 65 °C for 3 days. Extra hours of incubation in basic solution are necessary to remove the 2-cyanoethyl group on hmC.[23,24] **ODN2** was obtained after HPLC purifications both before and after detritylation, and was confirmed by ESI-MS spectra (Fig. S3). To further confirm the success of hmC incorporation and completion of hmC deprotection in the oligonucleotide, **ODN2** was enzymatically digested to free nucleosides. A new peak, besides four natural nucleobases, was eluted at 6.9 min and was characterized by ESI-MS (Fig. 2B). Figure 2C shows both the pseudo-molecular ion of hmdC $[M+H]^+$ m/z for $C_{10}H_{15}N_3O_5^+$ 258.1088) and nucleobase fragment ($C_5H_8N_3O_2^+$ 142.0612) were found from LC–MS.

In order to show that the oligonucleotides prepared by the synthetic method here have the biochemical behaviors expected for fC and hmC containing DNA strands, duplexes, **ODN1**–**ODN4** and **ODN2**–**ODN4** were subjected to thermal denaturation analysis. Figure 3A showed that both epigenetic duplexes can undergo thermal dissociation to single strands and UV–vis absorption at 260 nm accordingly showed a sigmoid transition similar to unmodified wide type duplex, **ODN3**–**ODN4**. The transition of UV–vis absorption due to denaturation of duplex DNA completed within the similar temperature ranges, around 20 degree celsius, for both epigenetic modified and natural duplexes, which indicated that fC and hmC modified duplexes can un/fold between two major secondary structures as wild type DNA. Upon fitting the denaturing curves with sigmoid function, fC containing duplex, **ODN1**–**ODN4** showed slightly higher melting temperature (T_m = 62.3 °C) than natural duplex (62.1 °C), while hmC modified duplex, **ODN2**–**ODN4** melted at apparently lower temperature (61.0 °C). The trend is consistent with the literatures.[25,26] Furthermore, thermal stability of fC base paired to guanine versus the other three natural nucleobases was examined by melting analysis of duplexes with mismatches at fC site, **ODN1**–**ODN5/6/7** (Table 1 and Table S3). All three mismatches with fC showed much lower melting temperatures than that of well-matched duplex ($\Delta T_m = T_{m,ODN4} - T_{m,ODN5/6/7}$ is between 5 and 7 °C), though fC-A in **ODN1**–**ODN5** exhibit 2 °C higher in T_m than the rest two mismatches with pyrimidines. fC can readily distinguish not only well-matched guanine from mismatched bases, but also adenosine from pyrimidine mismatches. The thermal denaturation of fC and hmC containing oligonucleotides prepared by our protocol satisfied with the criteria of canonical dissociation transition, defined melting temperatures and sensitivity on thermal stability of base pairing, which indicates these epigenetic oligonucleotides perform the same biological functions and behaviors as those prepared by other literature protocols.

In summary, we have developed novel synthetic pathways for facile preparation of hmC and fC containing oligonucleotides. The methods rely on a peroxysulfate oxidation reaction to convert natural nucleosides, such as dT and mdC to hmdU and fdC in aqueous-organic cosolvent under mild heating. By applying reaction directly to nucleosides, repetitive de/protection steps of precursor are omitted and both epigenetic cytosine phosphoramidites can be achieved within only four to six steps. These methods are useful for preparing epigenetic DNA oligonucleotides with ready adaption to prevailing instrumental setup and would provide significant

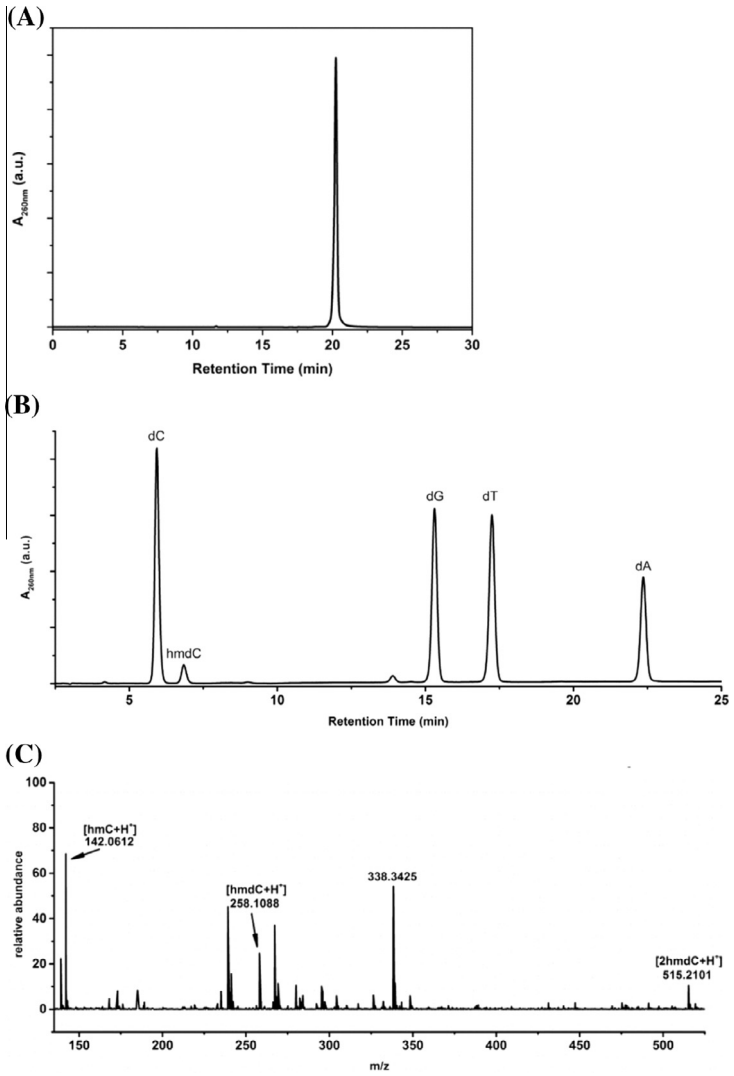


Figure 2. Reverse phase HPLC chromatography of **ODN2** after purification (A) and after enzymatic digestion (B). Elute peak at 6.83 min from (B) was submitted to LC–MS characterization for hmdC (C).

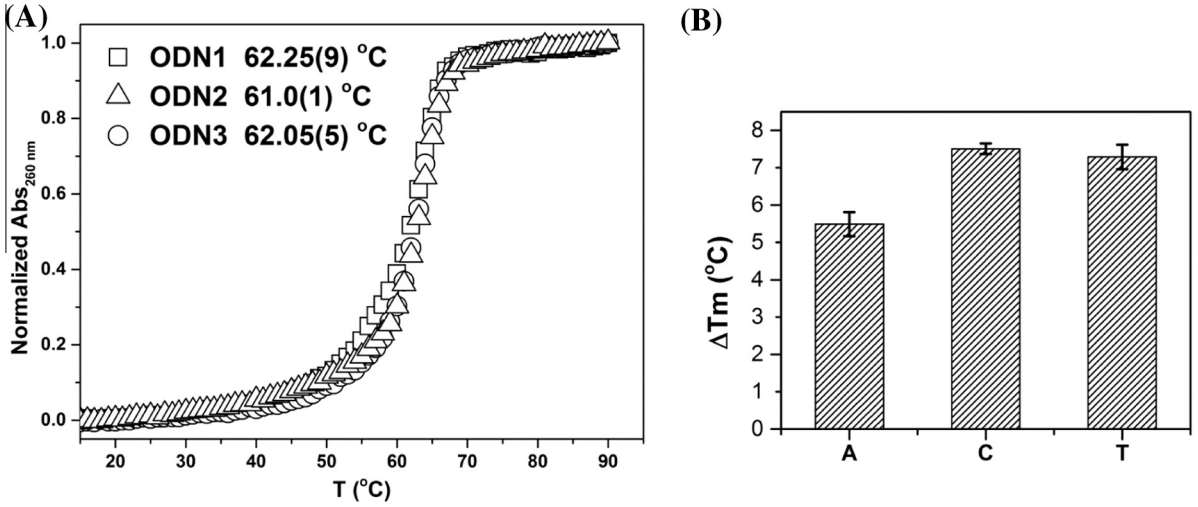


Figure 3. Melting analysis of epigenetic modified duplexes. (A) UV–vis absorption denaturing profiles and melting temperatures (standard deviations on last decimal are in parentheses) of duplexes **ODN1**–**3**/**ODN4**; (B) differences of melting temperatures between well-matched fdC containing duplex (**ODN1**/**ODN4**) and mismatched duplexes (**ODN1**/**ODN5**–**7**).

convenience for the investigation of the biological functions of the newly discovered epigenetic bases.

Acknowledgments

The financial support for this research work by Ministry of Education of Singapore (M4011040, M4020163) and Nanyang Technological University (M4080531) is greatly appreciated.

Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bmcl.2015.01.070.

References and notes

1. Kriaucionis, S.; Heintz, N. *Science* **2009**, *324*, 929.
2. 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.
3. Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. *Science* **2011**, *333*, 1300.
4. Globisch, D.; Munzel, M.; Muller, M.; Michalakis, S.; Wagner, M.; Koch, S.; Bruckl, T.; Biel, M.; Carell, T. *PloS one* **2010**, *5*, e15367.
5. Kinney, S. M.; Chin, H. G.; Vaisvila, R.; Bitinaite, J.; Zheng, Y.; Esteve, P. O.; Feng, S.; Stroud, H.; Jacobsen, S. E.; Pradhan, S. *J. Biol. Chem.* **2011**, *286*, 24685.
6. Klug, M.; Schmidhofer, S.; Gebhard, C.; Andreesen, R.; Rehli, M. *Genome Biol.* **2013**, 14.
7. Hu, X.; Zhang, L.; Mao, S. Q.; Li, Z.; Chen, J. K.; Zhang, R. R.; Wu, H. P.; Gao, J.; Guo, F.; Liu, W.; Xu, G. F.; Dai, H. Q.; Shi, Y. G.; Li, X. L.; Hu, B. Q.; Tang, F. C.; Pei, D. Q.; Xu, G. L. *Cell Stem Cell* **2014**, *14*, 512.
8. Maiti, A.; Drohat, A. C. *J. Biol. Chem.* **2011**, *286*, 35334.
9. Wu, H.; D'Alessio, A. C.; Ito, S.; Wang, Z. B.; Cui, K. R.; Zhao, K. J.; Sun, Y. E.; Zhang, Y. *Gene Dev.* **2011**, *25*, 679.
10. Williams, K.; Christensen, J.; Pedersen, M. T.; Johansen, J. V.; Cloos, P. A.; Rappsilber, J.; Helin, K. *Nature* **2011**, *473*, 343.
11. Wu, H.; D'Alessio, A. C.; Ito, S.; Xia, K.; Wang, Z.; Cui, K.; Zhao, K.; Sun, Y. E.; Zhang, Y. *Nature* **2011**, *473*, 389.
12. Shen, L.; Zhang, Y. *Curr. Opin. Cell Biol.* **2013**, *25*, 289.
13. Song, C. X.; Szulwach, K. E.; Dai, Q.; Fu, Y.; Mao, S. Q.; Lin, L.; Street, C.; Li, Y.; Poidevin, M.; Wu, H.; Gao, J.; Liu, P.; Li, L.; Xu, G. L.; Jin, P.; He, C. *Cell* **2013**, *153*, 678.
14. Tan, L.; Shi, Y. G. *Development* **1895**, *2012*, 139.
15. Karino, N.; Ueno, Y.; Matsuda, A. *Nucleic Acids Res.* **2001**, *29*, 2456.
16. Munzel, M.; Lischke, U.; Stathis, D.; Pfaffeneder, T.; Gnerlich, F. A.; Deiml, C. A.; Koch, S. C.; Karaghiosoff, K.; Carell, T. *Chem. Eur. J.* **2011**, *17*, 13782.
17. Dai, Q.; He, C. A. *Org. Lett.* **2011**, *13*, 3446.
18. Schroder, A. S.; Steinbacher, J.; Steigenberger, B.; Gnerlich, F. A.; Schiesser, S.; Pfaffeneder, T.; Carell, T. *Angew. Chem., Int. Ed.* **2014**, *53*, 315.
19. Itahara, T.; Yoshitake, T.; Koga, S.; Nishino, A. *Bull. Chem. Soc. Jpn* **1994**, *67*, 2257.
20. Itahara, T. *Chem. Lett.* **1991**, 1591.
21. Catalanotti, B.; Galeone, A.; Mayol, L.; Oliviero, G.; Rigano, D.; Varra, N. *Nucleos. Nucleot. Nucl.* **1831**, *2001*, 20.
22. Catalanotti, B.; Galeone, A.; Mayol, L.; Oliviero, G.; Rigano, D.; Varra, M. *Nucleos. Nucleot. Nucl.* **2002**, *21*, 107.
23. Hansen, A. S.; Thalhammer, A.; El-Sagheer, A. H.; Brown, T.; Schofield, C. J. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 1181.
24. Tardy-Planechaud, S.; Fujimoto, J.; Lin, S. S.; Sowers, L. C. *Nucleic Acids Res.* **1997**, *25*, 553.
25. Lopez, C. M. R.; Lloyd, A. J.; Leonard, K.; Wilkinson, M. J. *Anal. Chem.* **2012**, *84*, 7336.
26. Thalhammer, A.; Hansen, A. S.; El-Sagheer, A. H.; Brown, T.; Schofield, C. J. *Chem. Commun.* **2011**, 5325.