ORGANIC LETTERS

XXXX
Vol. XX, No. XX
000–000

Selective Detection of 5-Formyl-2′-deoxycytidine in DNA Using a Fluorogenic Hydroxylamine Reagent

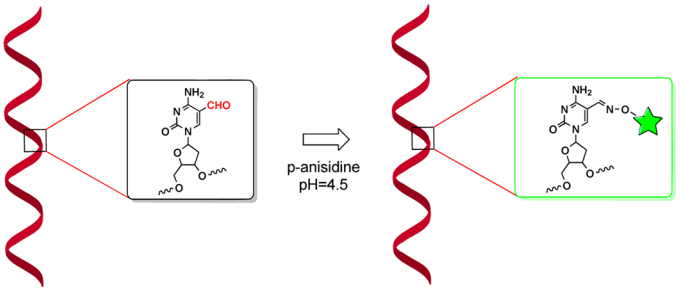
Pu Guo,[†] Shengyong Yan,[†] Jianlin Hu,[†] Xiwen Xing,[†] Changcheng Wang,[†] Xiaowei Xu,[†] Xiaoyu Qiu,[†] Wen Ma,[†] Chunjiang Lu,[†] Xiaocheng Weng,[†] and Xiang Zhou*[,†,‡]

College of Chemistry and Molecular Sciences, Key Laboratory of Biomedical Polymers of Ministry of Education, Wuhan University, Hubei, Wuhan, 430072, P. R. of China, and State Key Laboratory of Natural and Biomimetic Drugs, Beijing University, Beijing, 100191, P. R. China

xzhou@whu.edu.cn

Received May 9, 2013; Revised Manuscript Received June 11, 2013

ABSTRACT



Fluorogenic hydroxylamine reagents were used for detecting 5-fC through a labeling pathway. Chemical synthesis, HPLC, denaturing PAGE, and DNA MS were applied to testify that the probe reacted with 5-fC with oligodeoxynucleotide selectivity to achieve 5-fC detection conveniently and quantificationally with the method of fluorescence. The feasibility of fluorescently detecting 5-fC in a genome was also investigated.

The epigenetic base 5-methylcytosine (5-mC) in mammalian genomes plays essential roles in maintaining cellular function and genomic stability and is involved in biological processes including chromosome inactivation, genomic imprinting, and transposon silencing.[1] Recently, studies[2] on an active oxidative demethylation process for cytidine derivatives were reported. The demethylation mechanism involves the iterative oxidation of 5-methylcytosine (5-mC) to 5-hydroxymethylcytosine (5-hmC), 5-formylcytosine (5-fC), and 5-carboxycytosine (5-caC) by the 10-11 translocation (TET) enzymes, followed by base excision repair by thymine DNA glycosylase (TDG).[3] As the main oxidative product, 5-hmC has been implicated in stem cell biology[4] and cancer diagnosis using several detection methods.[5] It has been reported that 5-formylcytosine (5-fC) is present in mouse embryonic stem (ES) cells and the mouse cerebral cortex through LC-MS and deep sequencing.[3b,6] It has been confirmed that Activation-induced deaminase (AID)/APOBEC family cytosine deaminases cannot convert 5-hmC into (5-hydroxymethyluracil) 5-hmU to complete the DNA demethylation process,[7] whereas 5-fC, as an oxidation

[†] Wuhan University.
[‡] Beijing University.

(1) (a) Deaton, A. M.; Bird, A. *Genes Dev*. **2011**, *25*, 1010–1022. (b) Jones, P. A. *Nat. Rev. Genet*. **2012**, *13*, 484–492.

(2) (a) Miller, C. A.; Sweatt, J. D. *Neuron* **2007**, *53*, 857–869. (b) Day, J. J.; Sweatt, J. D. *Nat. Neurosci*. **2010**, *13*, 1319–1323. (c) Wu, S. C.; Zhang, Y. *Nat. Rev. Mol. Cell Biol*. **2010**, *11*, 607–620. (d) Zhu, J. K. *Annu. Rev. Genet*. **2009**, *43*, 143–166. (e) Feng, S. H.; Jacobsen, S. E.; Reik, W. *Science* **2010**, *330*, 622–627. (f) Reik, W.; Dean, W.; Walter, J. *Science* **2001**, *293*, 1089–1093. (g) Globisch, D.; Münzel, M.; Müller, M.; Michalakis, S.; Wagner, M.; Koch, S.; Brückl, T.; Biel, M.; Carell, T. *PLoS ONE* **2010**, *5*, e15367. (h) Spruijt, C. G.; Gnerlich, F.; Smits, A. H.; Pfaffeneder, T.; Jansen, P. W. T. C.; Bauer, C.; Münzel, M.; Wagner, M.; Müller, M.; Khan, F.; Eberl, H. C.; Mensinga, A.; Brinkman, A. B.; Lephikov, K.; Müller, U.; Walter, J.; Boelens, R.; Ingen, H. V.; Leonhardt, H.; Carell, T.; Vermeulen, M. *Cell* **2013**, *152*, 1146–1159.

(3) (a) Maiti, A.; Drohat, A. C. *J. Biol. Chem*. **2011**, *286*, 35334–35338. (b) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. *Science* **2011**, *333*, 1300–1303.

(4) (a) Pastor, W. A.; Pape, U. J.; Huang, Y.; Henderson, H. R.; Lister, R.; Ko, M.; McLoughlin, E. M.; Yevgeny, B.; Mahapatra, S.; Philipp, K.; Mamta, T.; Daley, G. Q.; Liu, X. S.; Ecker, J. R.; Patrice, M. M.; Agarwal, S.; Rao, A. *Nature* **2011**, *473*, 394–397.

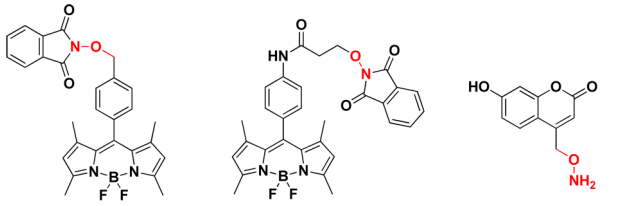
10.1021/ol401290d

intermediate of TET enzymes, remains detectable.[3b] As the major product after the TET-mediated oxidation of 5-mC, 5-fC is situated at a pivotal step in the DNA demethylation process. In addition, a recent report indicated that 5-fC plays potential roles in differentiation.[6b]

In view of the importance of 5-fC in the genome, developing a sensitive, selective, and convenient detection method for 5-fC has become a challenge. The most recent breakthrough in this field came with the development of three new methods allowing the measurement of 5-fC.[3b,6] The methods include LC-MS and deep sequencing technology. Our group just reported that a primary amino could be used to detect 5-fC; however, scant proof and interference of abasic sites limit its application in the detection of 5-fC in a genome.[6c] Herein, considering the high cost and complicated nature of the above methods, we developed a simple and effective approach to qualitatively and quantitatively detect 5-fC in DNA. Furthermore, we deduced that our method will hopefully be applied in the detection of 5-fC in a genome.

In this paper, we report a new approach for the simple detection of 5-fC in DNA using a fluorogenic hydroxylamine reagent. Hydroxylamine, an active carbonyl-reactive group, is widely used in fluorescence labeling applications. Therefore, we designed fluorescent probes for 5-fC based on the hydroxylamine group. The oxime ligation of the hydroxylamine and carbonyl groups is a popular reaction in chemical biology.[8] Oxime ligations can be significantly accelerated by using aniline and its derivatives as nucleophilic catalysts under acidic conditions.[9] Because of the high reactivity of hydroxylamine on benzyl, the reaction of DNA containing 5-fC with hydroxylamine probes could occur successfully in ammonium acetate buffer (pH = 4.5) when catalyzed by *p*-anisidine.

Scheme 1. Structure of Synthetic Hydroxylamine Probes



First, we designed and synthesized the probes BODIPY-B (**1**), BODIPY-L (**2**), and Coumarin-HA (**3**) (see Scheme 1). BODIPY-B (**1**) was obtained in two steps. Amino-BODIPY was subjected to a condensation reaction and produced BODIPY-L (**2**) to increase the length of the carbon chain. In addition, Coumarin-HA (**3**) was synthesized from 4-bromomethyl-7-hydroxy-chromen-2-one through base-promoted HBr-elimination and hydrazinolysis. Reagents **1** and **2** were treated with hydrazine hydrate to release the free hydroxylamine group just before exposure to 5-fC.[8b,10]

We first tested the activity of the hydroxylamine probes using an organic reaction (see Scheme 2). Compound **1′**, derived from the hydrazinolysis of BODIPY-B (**1**), and probe **3** were added separately to MeOH containing 5-fC, followed by the addition of several drops of acetic acid. The reaction occurred in the presence of *p*-anisidine as the catalyst for 12 h at 50 °C, giving fluorescent product **4** in a 95% yield and **5** in an 83% yield.[11] The satisfactory yields support the use of hydroxylamine probes in the detection of 5-fC. Furthermore, we measured the fluorescence spectra of compound **5** to facilitate the detection of 5-fC in DNA (see Figure S1). (We added the fluorescence spectra with excitation and emission wavelengths of compound **5** in the Supporting Information (Figure S1).)

Next, we chose a 21-mer oligodeoxyribonucleotide containing 5-fC (ODN-1)[12] and incubated it with probe **3** under the optimized conditions (Figure 1a). A solution of ODN-1 (1 μM) and probe **3** (10 μM) in 10 mM ammonium acetate buffer (pH = 4.5) containing 10 mM *p*-anisidine

(5) (a) Tahiliani, M.; Koh, K. P.; Shen, Y. H.; Pastor, W. A.; Bandukwala, H.; Brudno, Y.; Agarwal, S.; Lakshminarayan, M. I.; Liu, D. R.; Aravind, L.; Rao, A. *Science* **2009**, *324*, 930–935. (b) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K. L.; Sowers, C.; Zhang, Y. *Nature* **2010**, *466*, 1129–1133. (c) Ko, M.; Huang, Y.; Jankowska, A. M.; Pape, U. J.; Tahiliani, M.; Bandukwala, H. S.; An, J.; Lamperti, E. D.; Koh, K. P.; Ganetzky, R.; Liu, X. S.; Aravind, L.; Agarwal, S.; Maciejewski, J. P.; Rao, A. *Nature* **2010**, *468*, 839–843. (d) Koh, K. P.; Yabuuchi, A.; Rao, S.; Huang, Y.; Cunniff, K.; Nardone, J.; Laiho, A.; Tahiliani, M.; Sommer, C. A.; Mostoslavsky, G.; Lahesmaa, R.; Orkin, S. H.; Rodig, S. J.; Daley, G. Q.; Rao, A. *Cell Stem Cell* **2011**, *8*, 200–213. (e) Delhommeau, F.; Dupont, S.; Della Valle, V.; Chloé, J.; Trannoy, S.; Massé, A.; Kosmider, O.; Le Couedic, J. P.; Robert, F.; Alberdi, A.; Lécluse, Y.; Plo, I.; Dreyfus, F. J.; Marzac, C.; Casadevall, N.; Lacombe, C.; Romana, S. P.; Dessen, P.; Soulier, J.; Viguié, F.; Fontenay, M.; Vainchenker, W.; A. Bernard, O. *N. Engl. J. Med.* **2009**, *360*, 2289–2301. (f) Song, C. X.; Szulwach, K. E.; Fu, Y.; Dai, Q.; Yi, C. Q.; Li, X. K.; Li, Y. J.; Chen, C. H.; Zhang, W.; Jian, X.; Wang, J.; Zhang, L.; Timothy, J. L.; Zhang, B. C.; Godley, L. A.; Hicks, L. M.; Lahn, B. T.; Jin, P.; He, C. *Nat. Biotechnol.* **2011**, *29*, 68–72.

(6) (a) Pfaffeneder, T.; Hackner, B.; Truss, M.; Münzel, M.; Müller, M.; Deiml, C. A.; Hagemeier, C.; Carell, T. *Angew. Chem., Int. Ed.* **2011**, *50*, 7008–7012. (b) Raiber, E. A.; Beraldi, D.; Ficz, G.; Burgess, H. E.; Branco, M. R.; Murat, P.; Oxley, D.; Booth, M. J.; Reik, W.; Balasubramanian, S. *Genome Biology* **2012**, *13* (8), R69. (c) Hu, J. L.; Xing, X. W.; Xu, X. W.; Wu, F.; Guo, P.; Yan, S. Y.; Xu, Z. H.; Xu, J. H.; Weng, X. C.; Zhou, X. *Chem.—Eur. J.* **2013**, *19*, 5836.

(7) Nabel, C. S.; Jia, H. J.; Ye, Y.; Shen, L.; Goldschmidt, H. L.; Stivers, J. T.; Zhang, Y.; Kohli, R. M. *Nat. Chem. Biol.* **2012**, *8*, 751–758.

(8) (a) Wu, B.; Wang, Z. Y.; Huang, Y.; Liu, W. R. *ChemBioChem* **2012**, *13*, 1405–1408. (b) Rayo, J.; Amara, N.; Krief, P.; Meijler, M. M. *J. Am. Chem. Soc.* **2011**, *133*, 7469–7475.

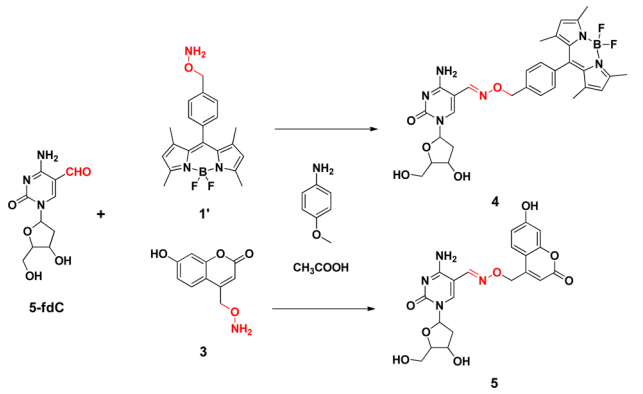
(9) (a) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7581–7584. (b) Lempens, E. H. M.; Helms, B. A.; Merkx, M.; Meijer, E. W. *ChemBioChem* **2009**, *10*, 658–662. (c) Zeng, Y.; Ramya, T. N. C.; Dirksen, A.; Dawson, P. E.; Paulson, J. C. *Nat. Methods* **2009**, *6*, 207–209.

(10) (a) Zhou, Z.; Fahrni, C. J. *J. Am. Chem. Soc.* **2004**, *126*, 8862–8863. (b) Murray, B. S.; Jackson, A. W.; Mahon, C. S.; Fulton, D. A. *Chem. Commun.* **2010**, *46*, 8651–8653. (c) Wey, S. J.; Augustyniak, M. E.; Cochran, E. D.; Ellis, J. L.; Fang, X. Q.; Garvey, D. S.; Janero, D. R.; Letts, L. G.; Martino, A. M.; Melim, T. L.; Murty, M. G.; Richardson, S. K.; Schroeder, J. D.; Selig, W. M.; Trocha, A. M.; Wexler, R. S.; Young, D. V.; Zemtseva, I. S.; Zifcak, B. M. *J. Med. Chem.* **2007**, *50*, 6367–6382. (d) Chen, F.; Song, K. S.; Wu, Y. D.; Yang, D. *J. Am. Chem. Soc.* **2008**, *130*, 743–755.

(11) Raindlová, V.; Pohl, R.; Šanda, M.; Hocek, M. *Angew. Chem., Int. Ed.* **2010**, *49*, 1064–1066.

(12) (a) Münzel, 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*, 13782–13788. (b) Münzel, M.; Globisch, D.; Trindler, C.; Carell, T. *Org. Lett.* **2010**, *12*, 5671–5673. (c) Steigenberger, B.; Schiesser, S.; Hackner, B.; Brandmayr, C.; Laube, S. K.; Steinbacher, J.; Pfaffeneder, T.; Carell, T. *Org. Lett.* **2013**, *15*, 366–369.

Scheme 2. Reaction of 5-fC with Hydroxylamine Probes



was incubated at rt for 12 h to give the ODN-1-CHA-containing compound **5** in > 97% yield. After probe **3** was added, the reaction mixture was monitored by HPLC, and the chromatograms are displayed in Figure 1b. After the reaction proceeded for 2 h, an obvious new peak was observed, and with increasing time, this new peak increased in intensity as ODN-1 disappeared. As shown in Figure 1c, the new peak in the HPLC chromatogram was collected, and its identity was confirmed by MALDI-TOF analysis (see Figures 1c, S2). The MS analysis indicated the DNA product was the desired ODN-1-CHA. Probes **1** and **2** were also used to detect 5-fC in ODN-1 (see Figures S3, S4). The products of ODN-1 that reacted with probes **1** and **2** had higher fluorescence intensities than the product of ODN-1 reacted with probe **3**. Even at the concentration of 5 μM, we could see an unambiguous green color with the naked eye under UV light (see Figure S5).

In view of the fact that there are many kinds of bases in a genome besides normal ones, and the reaction of abasic sites either apurinic or apyrimidinic (AP) with compounds including a long carbon chain primary amine has been reported,[13] experiments focused on selectivity were necessary. Because a fluorophore was introduced into ODN-1, the reaction mixture was also analyzed by denaturing PAGE followed by fluorescence detection, and the results supported our previous conclusion.

ODN-2, ODN-3, ODN-4, and ODN-5 were individually incubated with probe **3** under the same reaction conditions. The fluorescence images of the denaturing PAGE gel revealed favorable selectivity for 5-fC among the tested oligodeoxynucleotides (see Figure 2). As shown in the upper image of Figure 2a and 2b, fluorescent probe **3** reacted with ODN-1, which contains 5-fC, but not with ODN-2, ODN-3, ODN-4, or ODN-5. In view of the fact that AP has a similar level or greater compared with 5-fC in genome, ODN-9 (containing AP and 5-fC) and excess ODN-5 were used to serve as controls (see Figure 2b). 15-mer ODN-8 (5-GACTCAA5-fCAGCCGTA-3′) and

(13) (a) Fundador, E.; Rusling, J. *Anal. Bioanal. Chem.* **2007**, *387*, 1883–1890. (b) Sczepanski, J. T.; Wong, R. S.; McKnight, J. N.; Bowman, G. D.; Greenberg, M. M. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 22475–22480.

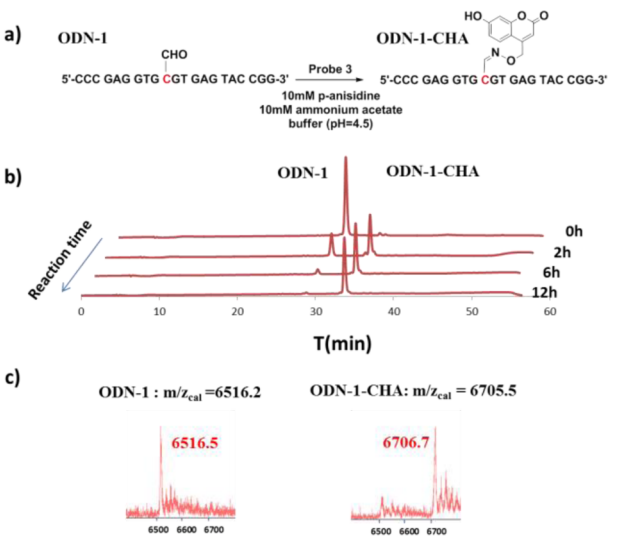


Figure 1. Conversion of ODN-1 (5′-CCCGAGGTG5-fC GTGAGTACCGG-3′) into ODN-1-CHA with probe **3**. (a) Reaction scheme with the sequences of ODN-1 and product ODN-1-CHA. (b) HPLC chromatogram of ODN-1 (starting DNA) and ODN-1-CHA (product DNA after probe **3** added at 0, 2, 6, and 12 h). (c) MALDI-TOF analysis of ODN-1 and ODN-1-CHA.

ODN-10 (5′-GACTCAA-AP-AGCCGTA-3′) were also checked as shown in Figure S6. The results still revealed excellent selectivity for 5-fC. The proposed reasons are that the steric hindrance of probe **3** and adjacent bases of AP hinder the possible reaction. The lower images of Figure 2a and 2b taken under 254 nm UV light indicate that unreacted ODN-2, -3, and -4 were present in lanes 2, 3, and 4 in Figure 2a and unreacted ODN-5 was present in lanes 2 (0.2 nmol) and 4 (0.6 nmol) in Figure 2b respectively. The relative positions are consistent with their molecular weights. Consistent with the upper image, the image of lane 1 in Figure 2a and lanes 1 and 3 in Figure 2b contained an unambiguous blue-fluorescent ODN band. We also interrupted the reaction at 3 h and performed denaturing PAGE analysis to assess the intermediate state in which ODN-1 and ODN-1-CHA coexist to observe the different locations of ODN-1 and ODN-1-CHA in the same lane (see Figure S7). To facilitate the fluorescence experiment, different concentrations of ODN-1 were incubated in the presence of equal amounts of probe **3**, and the fluorescence image of the denaturing PAGE gel indicated an apparent difference with increasing concentrations of ODN-1 (see Figure S8).

Finally, the fluorescence intensity of the reaction product was measured in ddH_2O after the removal of excess probe **3** and *p*-anisidine by extraction from the reaction mixture with ethyl acetate (AcOEt). As shown in Figure 3a, we used fluorescence-based methods to demonstrate that probe **3** can distinguish ODN containing 5-fC from ODN containing 5-hmC, 5-mC, C, and AP. It is obvious that the fluorescence intensity of ODN-1-CHA is much

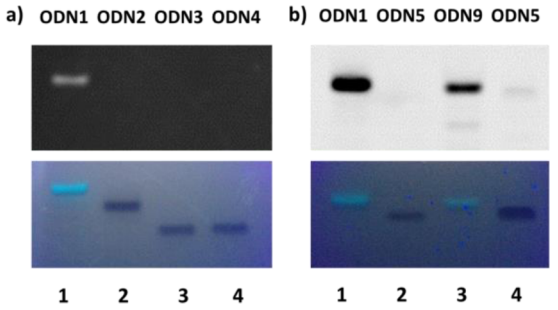


Figure 2. Denaturing PAGE probed fluorescent image analysis. (a) Lane 1: 0.2 nmol of ODN-1 (5′-CCCGAGGTG5-fCGT-GAGTACCGG-3′). Lane 2: 0.2 nmol of ODN-2 (5′-CCC-GAGGTG5-hmCGTGAGTACCGG-3′). Lane 3: 0.2 nmol of ODN-3 (5′-CCCGAGGTG5-mCGTGAGTACCGG-3′). Lane 4: 0.2 nmol of ODN-4 (5′-CCCGAGGTGdCGTGAGTAC-CGG-3′). (b) Lane 1: 0.2 nmol of ODN-1. Lane 2: 0.2 nmol of ODN-5 (5′-CCCGAGGTG-AP-GTGAGTACCGG-3′). Lane 3: 0.2 nmol of ODN-9 (5′-CCCGAGGTG5-fCGTGAGTA-AP-CGG-3′). Lane 4: 0.6 nmol of ODN-5 reacted with probe **3**.

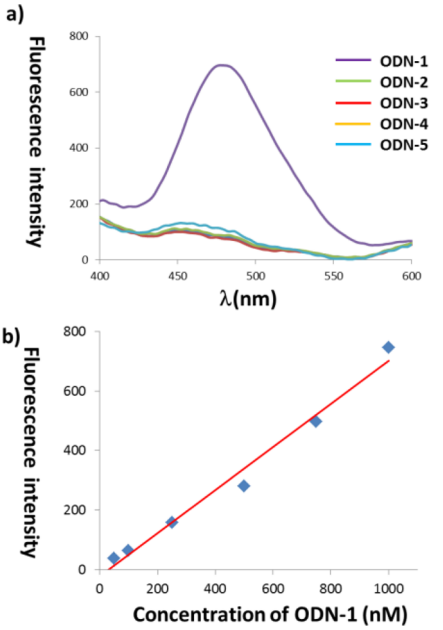


Figure 3. (a) Fluorescence intensity of 1 μM ODN-1, -2, -3, -4, -5 (containing 5-fC, 5-hmC, 5-mC, dC, AP) in H_2O reacted with 10 μM probe **3** after removal of excess compounds. (b) Fluorescence intensity (at 479 nm) of the reaction mixture containing various amounts of the ODN derived from ODN-1 in ddH_2O. Fluorescence measurements were performed at 25 °C with excitation at 325 nm. The limit of detetion is 42 nM. (Through the linearity of concentrations of DNA and fluorescence intensity, we could report the limit of detection for 5-fC is 42 nM.)

stronger than those of ODN-2, -3, -4, and -5 after incubation with probe **3**. We observed a linear correlation between the concentration of ODN-1 and the fluorescence intensity over the range 50−1000 nmol (Figure 3b). We could calculate the content of 5-fC based on the observed linear correlation with the fluorescence intensity. In order to investigate the feasibility of our method applied in the detection of 5-fC in genome, we chose DNA collected from mouse embryonic stem cells as practical samples. After reaction with probe **3** for 24 h at 37 °C, the genome was hydrolyzed with S1 nuclease, alkaline phosphatase, and snake venom phosphodiesterase I. Through collecting fractions at the corresponding retention time of compound **5** several times, we detected the presence of **5** in the enzymatic hydrolysate of the genome by LC-MS (see Figure S9). This result supported that our probes could be applied in the detection of 5-fC in genomes. (The limit of detection is a little high for being applied to the analysis in a genome. However, we could achieve this purpose through DNA enrichment or choosing a new probe with hydroxylamine. Additionally, we also proved the possibility in our manuscript (Figure S9).)

In summary, we have established a new method for the detection of 5-fC in DNA that is selective, convenient, inexpensive, and quantitative. Our method does not require complicated steps or separation processes before measuring the fluorescence. With the help of fluorescence, we are able to observe the presence of 5-fC in oligonucleotides and calculate the amount of 5-fC (this site is the same as that noted at the end of the previous paragraph). We also estimated the feasibility of our method applied in the detection of 5-fC in the genome of mouse embryonic stem cells. Finding a new probe to decrease the limit of detection and using a fluorescence-based approach for the detection of 5-fC in various genomic DNA from cancer cells will be the focus of our future work.

Acknowledgment. This work was supported by the National Basic Research Program of China (973 Program) (2012CB720600, 2012CB720603), the National Science Foundation of China (Nos. 91213302, 21272181, 21072115). Supported by Program for Changjiang Scholars and Innovative Research Team in University (IRT1030), the National Grand Program on Key Infectious Disease (2012ZX10003002-014). Supported by "the Fundamental Research Funds for the central Universities" (2012203020212).

Supporting Information Available. Full experimental details, compound characterization, HPLC, denaturing PAGE, MALDI-TOF and fluorescence data. This material is available free of charge via the Internet at http://pubs.acs.org.

The authors declare no competing financial interest.