Sulfinate Based Selective Labeling of 5-Hydroxymethylcytosine: Application to Biotin Pull Down Assay

Wu, Qiong; Amrutkar, Suyog Madhav; Shao, Fangwei

2018

Wu, Q., Amrutkar, S. M., & Shao, F. (2018). Sulfinate Based Selective Labeling of 5-Hydroxymethylcytosine: Application to Biotin Pull Down Assay. *Bioconjugate Chemistry*, 29(2), 245-249.

<https://hdl.handle.net/10356/88885>

<https://doi.org/10.1021/acs.bioconjchem.7b00826>

© 2018 American Chemical Society. This is the author created version of a work that has been peer reviewed and accepted for publication by *Bioconjugate Chemistry*, American Chemical Society. It incorporates referee's comments but changes resulting from the publishing process, such as copyediting, structural formatting, may not be reflected in this document. The published version is available at:
[<http://dx.doi.org/10.1021/acs.bioconjchem.7b00826>].

Downloaded on 26 Apr 2023 20:38:56 SGT

Sulfinate Based Selective Labelling of 5-Hydroxymethylcytosine: Application to Biotin Pull Down Assay

Qiong Wu, Suyog Madhav Amrutkar, Fangwei Shao*

Division of Chemistry and Biological Chemistry, School of Physical and Mathematical Sciences,
Nanyang Technological University, 21 Nanyang Link, Singapore, 637371.

*Address correspondence to: fwshao@ntu.edu.sg

Tel: +65 6592-2511

Fax: +65 6791-1961

Abstract

We developed an enzyme-free, chemical method to selectively label the epigenetic base, 5-hydroxymethylcytosine (5-hmC) with versatile sulfinate reagents in aqueous solvent under mild reaction conditions. This method allows efficient single step conjugation of biotin to 5-hmC site in DNA for enrichment and pull down assays.

Communication

5-hydroxymethylcytosine (hmC) was recently discovered as an intermediate of Ten-eleven translocation methylcytosine dioxygenase 1 catalyzed oxidative demethylation of 5-methylcytosine (mC).^{1, 2} The biased distribution within exons and near transcriptional start sites of genes in embryonic stem cells (ESCs),^{3, 4} high abundance in neuronal cells,^{5, 6} and high depletion level in cancer cells^{7, 8} all indicate hmC could be a stable epigenetic marker in the regulation of chromatin structure, gene transcription, cellular differentiation, and tumorigenesis.⁹ hmC in human genome holds therapeutic potentials as a biomarker in the diagnosis of cancer and neurological diseases, such as Alzheimer's.^{10, 11} Moreover, in-depth insights into the biological function of epigenetic bases can be gained by mapping its genome-wide locations.¹²⁻¹⁶ Whereas, the selective labelling and enrichment of hmC containing DNA are critical steps to study the biological and medicinal functions of hmC.

Earlier attempts of labelling hmC based on DNA immuno-precipitation sequencing (hMeDIP-seq), resulted in low hmC specificity, sequence bias of antibody, batch-to-batch variations and high background noise in antibody assays. This has urged the development of chemical labelling methods.^{12, 17} There are reports of successful enrichment of hmC containing DNA with UDP-6-azide-glucose labeling (hMe-Seal),¹⁸ oxidative glucose labelling (GLIB),³ and

thiol/selenol labelling.¹⁹ However, requirements of costlier restriction enzymes and cofactors, multistep synthesis of substrates, lot-to-lot inconsistency in enzyme activity, and number of steps in labelling reactions have limited the further applications of these chemical methods.^{20, 21}

Bisulfite sequencing (BS-Seq) is well established for single base resolution sequencing of mC.²² In a typical bisulfite treatment, upon reaction with concentrated sodium bisulfite solution, cytosines (C) in genomic DNA undergo hydrolytic deamination to yield uridine (U) while mC remains unreacted.²³ hmC upon NaHSO₃ treatment, rapidly converts to cytosine-5-methylsulfonate (CMS) (Figure 1).²⁴ This reaction is initialized by the activation of pyrimidine ring due to protonation of N³ under acidic condition. Next, addition of sulfonate group at C6 leads to the formation of an *exo*-methylene intermediate at C5. This intermediate in subsequent step is reacted with a second sulfonate anion to produce a disulfonated molecule. The sequential desulfonation at C6 yields CMS adduct as final product (Figure 1).²⁴ Li *et al.* conducted sulfinate labelling reaction in the presence of excess concentration of thiols that formed a thioether adduct of hmC which enabled a biotin labelling in the second step with a 30-65% yield.²¹

Nevertheless, further application of this method could be hindered by the prolong reaction time (36 to 48 hours), multiple-step labelling protocol and complicate procedures due to the undesired smell, oxidation susceptibility and poor aqueous solubility of thiol labelling compounds.

Figure 1 Chemical modification of 5-hydroxymethylcytosine: Bisulfite converts 5-hmC to the cytosine-5-methylenesulfonate (CMS) while in the presence of BioS gives sulfinate conjugated 5-hmC adducts (SFC). Plausible reaction mechanism for sulfinate conjugated 5-hmC.

In an attempt to establish chemical method for labelling hmC under amenable conditions for bioassay set up, we have selected alkyl sulfinates as the reactive group to conjugate the desired functionalities. Alkyl sulfinates possess good solubility and chemical stability in the aqueous buffers and good nucleophilicity. We speculated that alkyl sulfinates would trap *exo*-methylene intermediate at C5 which leads to the formation of 5-sulfinate-hmC (SFC) (Figure 1).²⁵

In exploratory experiments to label hmC with alkyl sulfinates, sodium ethanesulfinate (**1**) was used as a model substrate. To our delight, the reaction of hmdC with **1** yielded a new peak (Figure 2b). Its mass from ESI-MS (+) analysis confirmed the formation of ethanesulfinate labeled hmdC (**2**) (Figure 2c). As the bisulfite reaction requires protonation of pyrimidine- N^3 as

it activates the C5-C6 double bond which in the next step allows nucleophilic addition of bisulfites.²¹ The reaction was screened in NaPO₄ buffer of acidic pH (pH 2, 3 or 4) but no considerable change in yields was observed. Hence pH 5 is used in the following attempts. Inspired by the excellent reactivity of hmC towards NaHSO₃, we added a catalytic amount of NaHSO₃ into the reaction protocol in order to activate hmdC. As expected, the yield of product **2** was increased significantly when NaHSO₃ was added (Table S1). However, applying an excess amount of NaHSO₃ should be avoided because the high concentration of NaHSO₃ may compete with sodium ethanesulfinate to form side-product, CMS. At optimized reaction conditions i.e. 1 M of sodium ethanesulfinate (pH 5, in NaPO₄ buffer) in the presence of 50 mM of NaHSO₃ (pH 5) were incubated with hmdC (100 μ M) at 42°C for 17 hours, hmdC get converted to product **2** with 80% yield which was confirmed by HPLC followed with mass analysis (Figure 2b and 2c). Next, to validate the selectivity of aforesaid labelling method, nucleosides dC, mdC, hmdC, dA, dG and dT were subjected to reaction with sodium ethanesulfinate. Under the optimized protocol conditions, nucleosides dC, mdC, dA, dG, and dT remain un-reacted (Figure S1). These results showed the excellent selectivity of alkyl sulfinate labelling for hmC.

Figure 2. a) Reaction between nucleoside hmdC and ethanesulfinate; b) HPLC chromatogram at 274 nm of hmdC before (upper panel) and after (lower panel) the treatment with 1; c) ESI-MS (+) of the adduct 2 (highlighted by arrow) $m/z_{cal}=333.10$ ($M + H^+$), m/z found 334.08 ($M + H^+$).

To be adopted by versatile bioassays to study the biological and therapeutic roles of hmC, it is desirable that the method should afford hmC labelling in single and double stranded DNA. As a starting point, single-stranded DNA (ssCHC, table S2), comprised of one hmC in the middle of three CpG repeats with flanking random DNA sequences to build up the chemical and thermo-stability,²⁶ was reacted with **1** by following the optimal labelling protocol. HPLC analysis demonstrated the conversion of ssCHC into an adduct as a new peak with a longer retention time (Figure 3a). MALDI-TOF analysis confirms the adduct has the expected mass for

ethylsulfinate labeling of ss**CHC** (Figure 3b). To find out whether the labeling reaction occurred specifically to hmC, ss**CHC** after treated by labeling protocol was digested to nucleosides by deoxynucleases and was analyzed with HPLC (Figure 3c). Along with incubation time, a new peak with the same retention time as **2** emerged on HPLC traces. The generation of **2** was synchronized with that of single stranded adduct and with the disappearance of hmC peak in the enzymatic digestion crudes (Figure S2). The new peak was further confirmed to have the mass of **2**. To be noted that longer reaction time was not necessary to improve the labelling efficiency as shown in Table S3 and Figure S2. Maximum labelling yield at 74% can be achieved by carrying out the optimized protocol up to 24 h.

Figure 3. a) HPLC chromatograms (260 nm) of reaction between ssCHC and compound 1; b) MALDI-TOF analysis of ethanesulfinate labeled ssCHC adduct; c) HPLC chromatograms of nucleoside mixture from enzymatic digestion of DNA.

An efficient labelling of hmC in ssCHC incited us to look out the possibility of similar hmC labelling reaction with double-stranded DNA. The DNA duplex containing single hmC (dsCHC) was prepared by annealing CHC with the complementary strand in a NaPO₄ buffer of

pH 7. Then reaction pH was adjusted to 5 by H₃PO₄. ds**CHC** was incubated with **1** under protocol condition for ss**CHC**. The resulting protocol crudes were heated at 90°C to denature the duplex DNA and were followed with enzymatic digestion to yield nucleosides. HPLC analysis of protocol crudes showed 63% labelling yield of hmC adduct **2** (Figure S3). The yield was relatively lower than that for ss**CHC**. This may be due to the shielding of N³ in hmC by base pairing with G in double-stranded DNA that deactivates the pyrimidine ring. Nonetheless, the present method has shown promising potentials to label hmC in both single and double stranded DNA.

After the protocol of sulfinate labelling on hmC in DNA strands was optimized by using ethanesulfinate, the protocol was applied to introduce biotin conjugates to hmC, since the non-covalent interaction between biotin and streptavidin is among the strongest and the most specific bindings between small molecule and protein, and is widely used by variety of recognition, enrichment and purification bioassays. A sulfinate derivative of biotin (**BioS**) was designed and prepared (Figure 4a and scheme S1) as the labelling reagent.¹⁹ Synthesis of **BioS** was achieved by activating biotin carboxylic derivative with N-hydroxysuccinimide followed by linking biotin and sulfinate groups with aminocaproic acid to yield **BioS**.²⁰ To explore the biotin labelling on hmC containing DNA, ss**CHC** was incubated with 0.5 M of **BioS** under the protocol condition for 24 h. As shown in Figure 4b, a new peak emerged 1.2 mins later on HPLC trace with 67 % yield, which was further confirmed as desired biotin adduct of ssCHC (**Bio-CHC**) by MALDI-mass analysis (Figure S4). To further validate the enrichment efficiency of Bio-CHC, biotin/streptavidin pull-down assay was performed.²⁰ Single-stranded DNA, **CHC** and **Bio-CHC** were incubated with streptavidin beads and removed by centrifuge. The quantity of DNA in the supernatant was used to determine the efficiency of the pull down assay. 42 % of **Bio-CHC** was

removed from solution by binding to streptavidin beads, while free **CHC** remained in supernatant untouched (Figure 4c). This efficiency of the present method is comparable with that of the literature methods such as GLIB and thiol labeling.^{21, 27}

Figure 4: a) Synthesized biotin sulfinate probes (BioS) b) HPLC chromatograms of reaction between ssCHC and BioS; c) Efficiency of biotin/Streptavidin pull-down assay.

After enrichment step, recovery of the biotin-labeled DNA fragments from streptavidin is another crucial step to maintain the integrity and high yields of hmC containing DNA for further applications such as sequencing. But the strong affinity between biotin and streptavidin precludes the dissociation of biotin-labeled DNA from streptavidin under mild condition.^{28, 29}

The recovery of the biotin-labeled DNA often demands harsh denaturing conditions, such as incubation at near boiling temperature in the presence of formamide or SDS.²² Though recent attempts with adding disulphide linkage to the labelling probes or introducing Schiff base to formylated pyrimidines could allow relatively mild condition,^{30,31} the recovery of the biotin-labeled DNA often demands harsh denaturing conditions, such as incubation at near boiling temperature in the presence of formamide or SDS.^{32,33} However these harsh conditions have deleterious effects on DNA of interest. Whereas, the present labelling method offers a unique feature to cleave biotin moiety from hmC sites and to recover DNA fragments from streptavidin beads under mild reaction conditions. Different from other hmC labelling probe, the biotin adduct of hmC in present method contains a sulfonate linkage to C5-methylene at hmC site. The alkyl sulfonyl group is a good leaving group and is vulnerable towards the substitution reactions with strong nucleophiles and hence could become a potential cleavage site. Methylamine is commonly used as DNA deprotection reagent in solid phase oligonucleotides synthesis.³⁴ Altogether we speculate that the nucleophilic substitution of alkyl sulfonate group by basic nucleophile like methylamine can work under mild condition. In line with our speculation, when **Bio-CHC** was incubated with 20 % methylamine aqueous solution at 60 °C for 4 h (Figure 5), cleavage of biotin labels was observed by HPLC analysis (Figure S5) and a product (yield = 59%) with less hydrophobicity was eluted with shorter retention time. Further MALDI-TOF analysis has confirmed the successful substitution of **BioS** by methylamine on **CHC** (Figure S6). The cleavage step could also be used to introduce functionality to hmC-DNA via alkyl amine in case the sulfinate reagents of certain functionality is not readily available.

Figure 5: Methylamine mediated cleavage of Bio-CHC.

In conclusion, we have developed a novel chemical method to label hmC with sulfinate reagents. Model reaction using ethanesulfinate achieved specific conversion of hmdC nucleoside as well as hmC in single and double stranded DNA with yields up to 80%. The labelling method affords an efficient conjugation of biotin onto hmC sites in DNA via biotin-sulfinate probes in one step reaction with decent yield. Further application to biotin/streptavidin pull-down assay showed remarkable potentials of the present methods in the enrichment assays for hmC containing DNA. Moreover cleavage of DNA from biotin labels was achieved via a unique step of alkylamine incubation without harsh conditions. The mild biocompatible aqueous protocol, with good labelling yield, high selectivity for hmC and unique potentials for DNA recovery are the desirable attributes of this chemical method. Hence, the innovative method developed here provides a novel strategy for chemical labelling of hmC as nucleosides, oligonucleotides and duplexes, which could have wide applications in hmC involved biological studies, therapeutics and biotechnology, such as cellular 5-hmC quantification and epigenetic sequencing.

Supporting information

Supporting Information includes detail experimental procedures, HPLC chromatograms, MALDI-TOF analysis data, scheme for synthesis of BioS, ^1H NMR spectra. It is available free of charge on the ACS Publications website at DOI:

Acknowledgment

The authors' thank ministry of education, Singapore for MOE Tier 1 grant M4011660.

Corresponding Author

*Email: fwshao@ntu.edu.sg

Notes

The authors declare no competing financial interest.

References

- (1) Wu, X., and Zhang, Y. (2017) TET-mediated active DNA demethylation: mechanism, function and beyond. *Nat Rev Genet* 18, 517-534.
- (2) Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L. M., Liu, D. R., and Aravind, L. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *science* 324, 930-935.
- (3) Pastor, W. A., Pape, U. J., Huang, Y., Henderson, H. R., Lister, R., Ko, M., McLoughlin, E. M., Brudno, Y., Mahapatra, S., and Kapranov, P., et al. (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* 473, 394-397.
- (4) Xu, Y., Wu, F., Tan, L., Kong, L., Xiong, L., Deng, J., Barbera, A. J., Zheng, L., Zhang, H., and Huang, S., et al. (2011) Genome-wide Regulation of 5hmC, 5mC, and Gene Expression by Tet1 Hydroxylase in Mouse Embryonic Stem Cells. *Molecular Cell* 42, 451-464.

- (5) Kriaucionis, S., and Heintz, N. (2009) The Nuclear DNA Base 5-Hydroxymethylcytosine Is Present in Purkinje Neurons and the Brain. *Science* 324, 929-930.
- (6) Münzel, M., Globisch, D., Brückl, T., Wagner, M., Welzmler, V., Michalakis, S., Müller, M., Biel, M., and Carell, T. (2010) Quantification of the Sixth DNA Base Hydroxymethylcytosine in the Brain. *Angewandte Chemie International Edition* 49, 5375-5377.
- (7) Valinluck, V., and Sowers, L. C. (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* 67, 946-950.
- (8) Jin, S.-G., Jiang, Y., Qiu, R., Rauch, T. A., Wang, Y., Schackert, G., Krex, D., Lu, Q., and Pfeifer, G. P. (2011) 5-Hydroxymethylcytosine is strongly depleted in human cancers but its levels do not correlate with IDH1 mutations. *Cancer research* 71, 7360-7365.
- (9) Branco, M. R., Ficz, G., and Reik, W. (2012) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nature Reviews Genetics* 13, 7-13.
- (10) Ellison, E. M., Bradley-Whitman, M. A., and Lovell, M. A. (2017) Single-Base Resolution Mapping of 5-Hydroxymethylcytosine Modifications in Hippocampus of Alzheimer's Disease Subjects. *Journal of Molecular Neuroscience* 63, 185-197.
- (11) Haffner, M. C., Chaux, A., Meeker, A. K., Esopi, D. M., Gerber, J., Pellakuru, L. G., Toubaji, A., Argani, P., Iacobuzio-Donahue, C., and Nelson, W. G., et al. (2011) Global 5-hydroxymethylcytosine content is significantly reduced in tissue stem/progenitor cell compartments and in human cancers. *Oncotarget* 2, 627-637.

- (12) Matarese, F., Carrillo de Santa Pau, E., and Stunnenberg, H. G. (2011) 5 - Hydroxymethylcytosine: a new kid on the epigenetic block? *Molecular Systems Biology* 7.
- (13) Smagulova, F., Gregoret, I. V., Brick, K., Khil, P., Camerini-Otero, R. D., and Petukhova, G. V. (2011) Genome-wide analysis reveals novel molecular features of mouse recombination hotspots. *Nature* 472, 375-378.
- (14) Suzuki, M. M., and Bird, A. (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nature Reviews Genetics* 9, 465-476.
- (15) Hardisty, R. E., Kawasaki F., Sahakyan A. B., and Balasubramanian, S. (2015) Selective Chemical Labeling of Natural T Modifications in DNA. *Journal of the American Chemical Society* 137, 9270-9272.
- (16) Liu, C., Wang, Y., Zhang, X., Wu, F., Yang, W., Zou, G., Yao, Q., Wang, J., Chen, Y., and Wang, S., et al. (2017) Enrichment and fluorogenic labelling of 5-formyluracil in DNA. *Chem Sci.* 8, 4505–4510.
- (17) Booth, M. J., Raiber, E.-A., and Balasubramanian, S. (2014) Chemical methods for decoding cytosine modifications in DNA. *Chemical reviews* 115, 2240-2254.
- (18) Song, C.-X., Szulwach, K. E., Fu, Y., Dai, Q., Yi, C., Li, X., Li, Y., Chen, C.-H., Zhang, W., and Jian, X., et al. (2011) Selective chemical labeling reveals the genome-wide distribution of 5-hydroxymethylcytosine. *Nat Biotech* 29, 68-72.
- (19) Liutkevičiūtė, Z., Kriukienė, E., Grigaitytė, I., Masevičius, V., and Klimašauskas, S. (2011) Methyltransferase - Directed Derivatization of 5 - Hydroxymethylcytosine in DNA. *Angewandte Chemie* 123, 2138-2141.

- (20) Song, C.-X., Yi, C., and He, C. (2012) Mapping recently identified nucleotide variants in the genome and transcriptome. *30*, 1107.
- (21) Li, W. W., Gong, L., and Bayley, H. (2013) Single - Molecule Detection of 5 - Hydroxymethylcytosine in DNA through Chemical Modification and Nanopore Analysis. *Angewandte Chemie International Edition* *52*, 4350-4355.
- (22) Susan, J. C., Harrison, J., Paul, C. L., and Frommer, M. (1994) High sensitivity mapping of methylated cytosines. *Nucleic Acids Research* *22*, 2990-2997.
- (23) Shapiro, R., Servis, R. E., and Welcher, M. (1970) Reactions of Uracil and Cytosine Derivatives with Sodium Bisulfite. *Journal of the American Chemical Society* *92*, 422-424.
- (24) Huang, Y., Pastor, W. A., Shen, Y., Tahiliani, M., Liu, D. R., and Rao, A. (2010) The Behaviour of 5-Hydroxymethylcytosine in Bisulfite Sequencing. *PLoS ONE* *5*, e8888.
- (25) Brownbridge, P., and Jowett, I. C. (1988) 'One-Pot' Synthesis of Sulphinic Esters from Disulphides. *Synthesis* *1988*, 252-254.
- (26) Wu, Q., Wong, J. R., Qing Yeo, P. L., Zhang, D., and Shao, F. (2016) Methylation on CpG repeats modulates hydroxymethylcytosine induced duplex destabilization. *RSC Advances* *6*, 48858-48862.
- (27) Song, C.-X., Clark, T. A., Lu, X.-Y., Kislyuk, A., Dai, Q., Turner, S. W., He, C., and Korlach, J. (2012) Sensitive and specific single-molecule sequencing of 5-hydroxymethylcytosine. *Nat Meth* *9*, 75-77.
- (28) Leblond - Francillard, M., Dreyfus, M., and Rougeon, F. (1987) Isolation of DNA - protein complexes based on streptavidin and biotin interaction. *The FEBS Journal* *166*, 351-355.

- (29) Livnah, O., Bayer, E. A., Wilchek, M., and Sussman, J. L. (1993) Three-dimensional structures of avidin and the avidin-biotin complex. *Proc Natl Acad Sci U S A* 90, 5076-5080.
- (30) Kawasaki, F., Beraldi, D., Hardisty, R. E., McInroy, G. R., Delft, P., and Balasubramanian, S. (2017) Genome-wide mapping of 5-hydroxymethyluracil in the eukaryote parasite *Leishmania*. *GENOME BIOL* 18, 23.
- (31) Li, F., Zhang, Y., Bai, J., Greenberg, M. M., Xi, Z., and Zhou C. (2017) 5-Formylcytosine yields DNA–protein cross-links in nucleosome core particles. *Journal of the American Chemical Society* 139, 10617-10620.
- (32) A. Holmberg, A. Blomstergren, O. Nord, M. Lukacs, J. Lundeberg and M. Uhlén. (2005) The biotin-streptavidin interaction can be reversibly broken using water at elevated temperatures *Electrophoresis* 26, 501-510;
- (33) M. Wilchek, E. A. Bayer and O. Livnah. (2006) Essentials of biorecognition: The (strept)avidin–biotin system as a model for protein–protein and protein–ligand interaction *Immunology letters*, 103, 27-32.
- (34) Reddy, M., Hanna, N., and Farooqui, F. (1994) Fast cleavage and deprotection of oligonucleotides. *Tetrahedron letters* 35, 4311-4314.

Table of Content

A novel, biocompatible and selective chemical method to label hmC by using biotin sulfinate probes which allows versatile recovery reaction via a facile methylamine.