JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Direct Decarboxylation of 5-Carboxylcytosine by DNA C5-Methyltransferases

Zita Liutkevičiūtė,[†] Edita Kriukienė,[†] Janina Ličytė, Milda Rudytė, Giedrė Urbanavičiūtė, and Saulius Klimašauskas*

Department of Biological DNA Modification, Institute of Biotechnology, Vilnius University, 02241 Vilnius, Lithuania

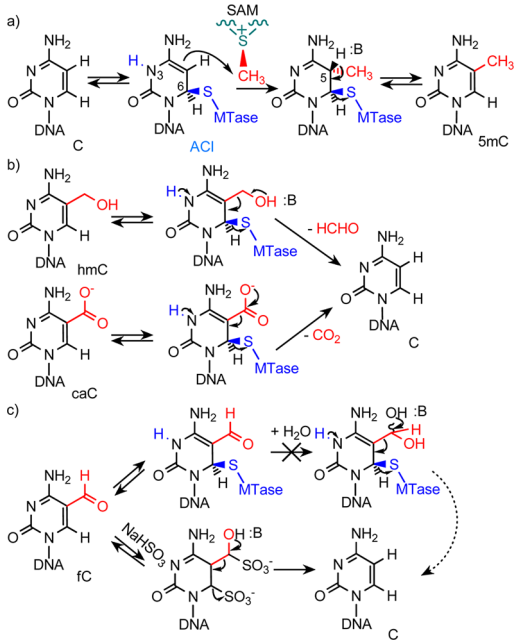
Supporting Information

ABSTRACT: *S*-Adenosylmethionine-dependent DNA methyltransferases (MTases) perform direct methylation of cytosine to yield 5-methylcytosine (5mC), which serves as part of the epigenetic regulation mechanism in vertebrates. Active demethylation of 5mC by TET oxygenases produces 5-formylcytosine (fC) and 5-carboxylcytosine (caC), which were shown to be enzymatically excised and then replaced with an unmodified nucleotide. Here we find that both bacterial and mammalian C5-MTases can catalyze the direct decarboxylation of caC yielding unmodified cytosine in DNA *in vitro* but are inert toward fC. The observed atypical enzymatic C−C bond cleavage reaction provides a plausible precedent for a direct reversal of caC to the unmodified state in DNA and offers a unique approach for sequence-specific analysis of genomic caC.

Modification of cytosine in CpG dinucleotides by *S*-adenosylmethionine-dependent DNA methyltransferases (MTases) plays important roles in the epigenetic regulation of gene function in vertebrates.[1] The enzymatic methylation reaction proceeds via direct transfer of a methyl group from *S*-adenosylmethionine (SAM) onto the C5 position of the target cytosine residue. In a reverse process termed demethylation, genomic 5-methylcytosine (5mC) is converted to 5-hydroxymethylcytosine (hmC), 5-formylcytosine (fC), and 5-carboxylcytosine (caC) by the action of the TET oxygenases.[2−5] Biological functions of the oxidized cytosines are not fully understood, although some evidence suggests that they may have certain epigenetic roles besides being intermediates of the demethylation pathway. The replacement of fC and caC into C is thought to occur via TDG-dependent base excision followed by repair of the abasic lesions in DNA strands.[6,7]

The enzymatic transfer of the sulfonium-bound methyl group from SAM onto the C5 position of the target cytosine residue requires its transient activation by nucleophilic addition of a conserved cysteine residue in the MTase enzyme to the 6-position of the ring (Scheme 1a).[8,9] In the absence of cofactor, the covalently activated intermediate can attack exogenous electrophiles, such as formaldehyde, leading to the production of hmC residues in DNA *in vitro*;[10] it was surprisingly found that the DNA C5-MTases are also capable of removing the 5-hydroxymethyl group from their target residues yielding unmodified cytosine (Scheme 1b). To understand the generality of these enzyme-promoted C−C bond cleavage reactions, we went on to

Scheme 1. Proposed Mechanisms of Covalent Catalysis by DNA Cytosine-5 Methyltransferases (MTases)



examine the activity of the bacterial C5-MTases M.HhaI and M.SssI (wild type enzymes and their engineered variants with a sterically expanded active site, eM.HhaI and eM.SssI)[11,12] with respect to cognate DNA duplexes containing caC or fC at the target position. For these studies, synthetic unlabeled 21-mer DNA duplexes were used for direct analyses using ESI-MS. In parallel, 31-mer DNA duplexes were produced to contain a uniquely ^{33}P-labeled modified cytosine nucleotide (see Supporting Table S1) at the target position of the GCGC (or CG) target site such that subsequent TLC analysis of labeled 5′-mononucleotides permitted a selective detection of covalent changes in the target residue. Both types of duplexes contained a modified target nucleotide on one strand and a 5-methylcytosine residue on the opposite strand.

ESI-MS experiments showed a time-dependent loss of 44 amu at the target DNA strand containing caC upon incubation of the 21-mer DNA duplexes with M.HhaI and M.SssI MTases. The methylated strand remained unaltered under these conditions

Received: February 24, 2014
Published: April 9, 2014

dx.doi.org/10.1021/ja5019223 | *J. Am. Chem. Soc.* 2014, 136, 5884−5887

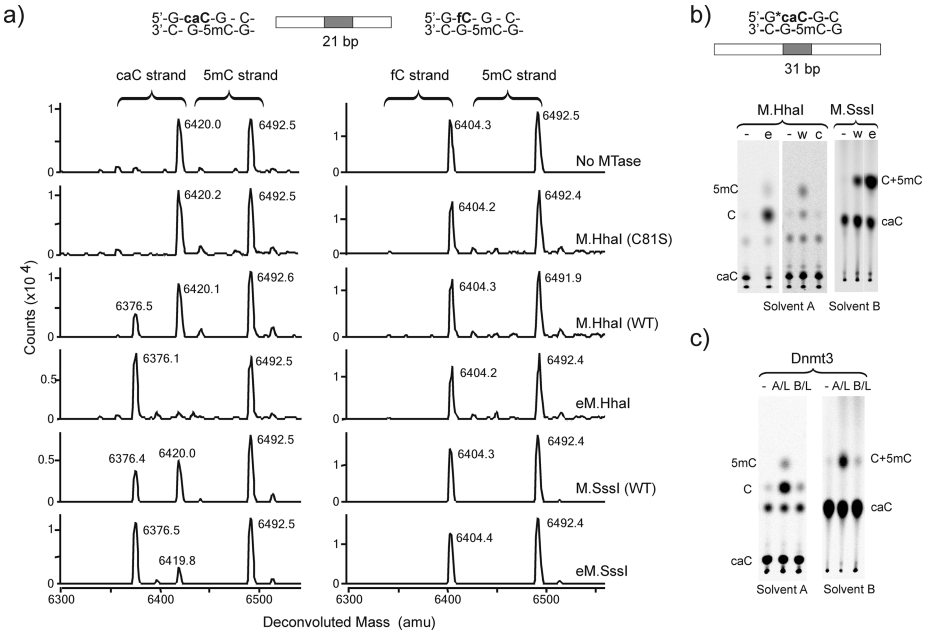
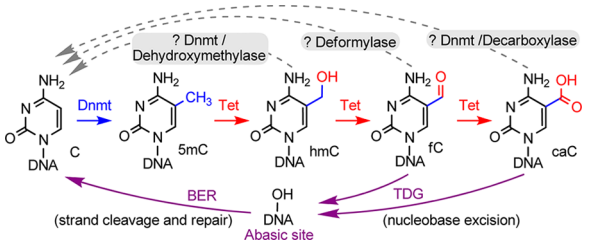


Figure 1. Decarboxylation activity of bacterial and human C5-MTases. (a) 21-mer hemimethylated DNA duplexes containing caC or fC at the target site were treated with M.HhaI or M.SssI as shown (WT, wild type; C81S, catalytic mutant) and analyzed using a Q-TOF mass spectrometer in negative ion mode. (b) TLC analysis of labeled caC nucleotide and its conversion products. A 31-mer DNA duplex containing ^{33}P-labeled caC (*caC) at the HhaI/SssI target site was treated with M.HhaI or M.SssI as shown (w, wild type variant; e, sterically engineered variant; c, C81S catalytic mutant). Modified DNA was digested to 5′-dNTPs, analyzed by TLC and autoradiography. (c) The 31-mer ^{33}P-labeled DNA duplex was treated with human Dnmt3A/3L and Dbmt3B/3L and analyzed as in (b).

Scheme 2. Biological DNA Methylation and Demethylation in Vertebrates[a]



[a]Cytosine (C) is converted to 5-methylcytosine (5mC) by endogenous C5-MTases of the Dnmt1 and Dnmt3 families (blue); 5mC can be consecutively converted to 5-hydroxymethylcytosine (hmC), 5-formylcytosine (fC), and 5-carboxylcytosine (caC) by the TET oxygenases (red). Known reverse pathways (magenta) involve base excision repair (BER) of fC and caC by thymine DNA glycosylase (TDG), leading to transient formation of abasic sites in DNA. Dashed arrows denote the newly discovered 5-dehydroxymethylation and 5-decarboxylation reactions performed by C5-MTases and which may be carried out *in vivo* by Dnmts or other putative enzymes (dehydroxymethylase, deformylase, decarboxylase). Adapted from ref 22.

(Figure 1a). The fC-containing duplex remained completely inert to a similar MTase treatment under a variety of reaction conditions. Control mobility shift experiments showed efficient formation of corresponding binary complexes (Figure S1), indicating that the modified substrates were properly recognized by the MTases. TLC analyses of the binary MTase-DNA complexes containing caC showed that, upon incubation in aqueous buffers, the amount of the caC nucleotide decreased, whereas the amount of cytosine increased (Figure 1b). In certain cases, the presence of trace amounts of the endogenous cofactor in the MTase preparations manifested itself in the concomitant formation of detectable labeled 5mC nucleotide (Figure 1b). Altogether, these observations clearly demonstrate that the C5-MTases promote the removal of the 5-carboxyl group from caC yielding unmodified C in DNA. Moreover, we find that the reaction requires the presence of the catalytic cysteine residue in the enzyme (Figure 1) suggesting a covalent reaction intermediate (Scheme 1b).

In comparative dehydroxymethylation vs decarboxylation studies (Figures S2−S3), the wild type MTases proved more efficient than their engineered variants in the hmC reaction, but less efficient toward caC. The engineered variants contain alanine replacements of conserved nonessential Gln and Asn residues in the active site.[11,12] With the assumption that deprotonation of the exocyclic group is required for the resolution of the covalent intermediate via C−C bond cleavage (Scheme 1), the lower acidity of the 5-hydroxymethyl group as compared to the 5-carboxyl group (pK_a ~13 and ~3, respectively)[13,14] suggests that one or both of these residues somehow assist in the deprotonation of the 5-hydroxymethyl group (by solvating a water molecule or directly interacting with the hydroxyl group). In contrast, no such assistance is required of the more acidic carboxyl group, in which case an extended steric space provided by the alanines caters for the increased bulk requirements. This might also explain the lack of detectable conversion in the case of fC, since the 5-formyl group is much less likely to be deprotonated than the latter two groups. Under conditions imitating enzymatic catalysis (high concentrations of exogenous thiol and imidazol, pH 5.0), the C−C bond cleavage occurred with all three types of oxidized cytosines in the following rate order: caC > fC > hmC.[15] The deformylation reaction most likely involves the addition of a nucleophile (water, thiol or bisulfite) followed by deprotonation of a hydroxyl group and release of formic acid (or corresponding thioester) (Scheme 1c); yet, the formation of such a hydrate may be obstructed in the active site of the enzymes.

dx.doi.org/10.1021/ja5019223 | J. Am. Chem. Soc. 2014, 136, 5884−5887

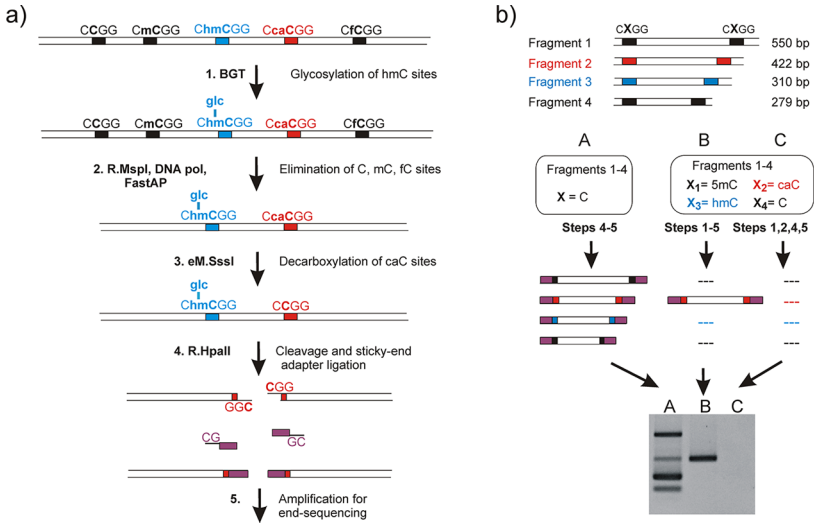


Figure 2. (a) Outline of the caC analytical procedure. Step 1, blockage of hmC sites by glucosylation with BGT; Step 2, elimination of C, 5mC, and fC sites by R.MspI cleavage followed by end-processing with T4/Klenow exo-DNA polymerases and FastAP phosphatase; Step 3, decarboxylation of caC with eM.SssI; Step 4, R.HpaII cleavage and ligation of CG-specific adaptors; Step 5, adaptor-specific PCR. (b) Validation of the caC enrichment procedure. A set of 4 DNA fragments with differently modified sites was analyzed in 3 parallel experiments (A, B C) as shown. A unique correctly sized fragment was obtained in lane B out of 4 possible products produced in control lane A (R.MspI used in Step 4). No PCR product was obtained if decarboxylation (Step 3) was omitted (lane C).

Besides the bacterial C5-MTases, mouse and human Dnmt1, Dnmt3A, and Dnmt3B were similarly examined. Under similar reaction conditions, the formation of cytosine was hardly detectable above the background (Figures 1c and S4). However, clear reactivity toward both caC (Figures 1c and S4) and hmC (Figure S5) was observed in the case of the human Dnmt3A/Dnmt3L complex. The catalytic efficiencies of the mammalian Dnmt enzymes *in vitro* are typically lower than those reported for the bacterial MTases. The DnmtL protein has no catalytic activity but is known to stimulate *de novo* methylation when bound to Dnmt3A and Dnmt3B MTases.[16,17] In our hands, two preparations of human Dnmt3B/Dnmt3L gave fairly weak but detectable amounts of the decarboxylation product (Figures 1c and S4). Given the high structural similarity in the catalytic domains of Dnmt3A and Dnmt3B, it is likely that both MTases may be active in this reaction under certain conditions. To this end, an hmC-to-C conversion has recently been reported for human Dnmt3A and Dnmt3B.[18] However, in contrast to the reported stimulation of the dehydroxymethylation activity of the individual Dnmt3MTases under oxidative conditions,[18] we observed the clear inhibition of the Dnmt-dependent 5-dehydroxymethylation, 5-decarboxylation, and methylation activities in the presence of hydrogen peroxide (Figure S6).

The described novel reactions of C5-MTases to catalyze the removal of exocyclic groups in modified cytosines provide yet another atypical example of the catalytic versatility of these cofactor-dependent enzymes.[10,19] The key feature of the catalysis is the formation of an activated covalent intermediate (ACI) promoting a variety of reactions depending on conditions and the presence of exogenous compounds (Scheme 1). Similarly, a covalent addition at C6 is operative in the thiol- or bisulfite-induced decarboxylation of caC in DNA.[6,15,20] However, known decarboxylases, including iso-orotate (5-carboxyluracyl) decarboxylase, seem to avoid covalent catalysis and typically utilize an organic cofactor or a transition metal coupled with dioxygen to activate their substrates.[21]

In mammalian DNA, caC is produced by oxidation of 5mC residues by the TET oxygenases as part of a DNA demethylation process (Scheme 2). caC was shown to be actively excised by thymidine DNA glycosylase (TDG), and the general base excision repair (BER) pathway is thought to restore the unmodified target sites in DNA.[6,7] Although this pathway is clearly important in many scenarios, it may not account for all situations. For instance, a concerted demethylation of high 5mCpG density regions via the BER pathway is expected to create multiple and closely spaced single- and double-strand cleavage events in DNA, presumably leading to the compromised genetic integrity of the locus. Our findings offer a demonstration of a direct enzymatic caC-to-cytosine conversion *in vitro*. In the context of an oxidative DNA demethylation mechanism *in vivo*, this provides a plausible chemical precedent for a possible direct reversal of the oxidized state to the unmodified state by Dnmts, their complexes with other proteins, or some yet unknown specialized enzymes (Scheme 2). In contrast to bacterial MTase M.SssI (Figure S3B), human Dnmt3A/L showed a substantial reduction of the decarboxylation activity at physiological SAM concentrations (Figure S4), suggesting that it is unlikely to act in the present form *in vivo*. Notably, other studies have found detectable decarboxylation activity in mouse cell extracts;[23] MS-based proteomics of mouse embryonic stem cells identified a strong genomic colocalization of Dnmt1 with caC, whereas binding of DNA-repair-associated proteins was most pronounced for fC.[24] This may point to different processing mechanisms involving caC and fC *in vivo*,[14] in line with the observed behavior of C5-MTases *in vitro*.

The MTase-directed decarboxylation reaction occurs under mild conditions and retains the cognate sequence- and nucleotide-specificity of the enzyme. The C5-MTases are also similarly active toward hmC, but not toward 5mC and fC. We therefore examined if this base selectivity can be exploited for analysis of caC residues in genomic DNA. Although several chemoenzymatic methods have been described for analysis of hmC and fC,[25−29] few are available for caC.[30] We first assessed the sensitivity of the CpG specific restriction endonucleases MspI and HpaII to the cytosine modifications using synthetic fluorescently labeled 89-mer DNA duplexes (Figure S7). The

dx.doi.org/10.1021/ja5019223 | J. Am. Chem. Soc. 2014, 136, 5884−5887

cleavage of the target site was quantitated by qPCR in 0.2 kb DNA fragments containing a single CCGG site with hemi- or symmetrical cytosine modifications (Figure S8A). Both endonucleases were fully blocked by caC and showed expected[7] differential sensitivity with respect to 5mC, hmC, and fC (Figure S8B). Treatment of the model fragments with eM.SssI converted the originally immune caC sites to R.HpaII-sensitive unmodified sites with >90% efficiency (Figure S8B). The R.HpaII cleavage produces termini with 5′-CG dinucleotide overhangs; this can be exploited for selective attachment of end-specific adapters followed by sequencing and genomic mapping of the original caC sites provided that CCGG sites containing other modifications are excluded. We therefore went on to develop a new method in which hmC sites are blocked by treatment of DNA with T4-β-glucosyltransferase (BGT), whereas C, 5mC, and fC sites are selectively eliminated by R.MspI cleavage and subsequent end-processing to preclude their ligation to CG-specific adaptors (Figure 2a). As a proof of principle, we carried out control experiments in which a mixture of four model fragments each containing differently modified cytosines at the CCGG target sites was analyzed (Figure 2b). Remarkably, we found a single amplicon originating from the caC-containing fragment, which comprised 10% of the input DNA sample (Figure 2b, lane B). Omission of the eM.SssI treatment gave no detectable amplification products under identical conditions (lane C), attesting efficient eM.SssI-directed decarboxylation of caC. Altogether, these findings indicate the high potential of the proposed approach for genomic CCGG-specific mapping of caC.

■ ASSOCIATED CONTENT

Supporting Information

Supporting Table S1, Figures S1−S8, and Experimental Procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

saulius.klimasauskas@bti.vu.lt

Author Contributions

†These authors contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to Shoji Tajima and Isao Suetake for mouse Dnmt1, Arvydas Lubys for M.SssI, Guoliang Xu and Bangan Wang for mDnmt3 expression plasmids, Artūras Petronis for FAM-oligonucleotides, and Indrė Grigaitytė for technical assistance. This work was supported by the European Social Fund under the Global Grant measure [Grant VP1-3.1-ŠMM-07-K-01-105 to S.K.].

■ REFERENCES

(1) Jurkowska, R. Z.; Jurkowski, T. P.; Jeltsch, A. *ChemBioChem* **2011**, *12*, 206.

(2) Kriaucionis, S.; Heintz, N. *Science* **2009**, *324*, 929.

(3) 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.

(4) Ito, S.; D'Alessio, A. C.; Taranova, O. V.; Hong, K.; Sowers, L. C.; Zhang, Y. *Nature* **2010**, *466*, 1129.

(5) Globisch, D.; Müzel, M.; Müller, M.; Michalakis, S.; Wagner, M.; Koch, S.; Brückl, T.; Biel, M.; Carell, T. *PLoS One* **2010**, *5*, e15367.

(6) He, Y.-F.; Li, B.-Z.; Li, Z.; Liu, P.; Wang, Y.; Tang, Q.; Ding, J.; Jia, Y.; Chen, Z.; Li, L.; Sun, Y.; Li, X.; Dai, Q.; Song, C.-X.; Zhang, K.; He, C.; Xu, G.-L. *Science* **2011**, *333*, 1303.

(7) Ito, S.; Shen, L.; Dai, Q.; Wu, S. C.; Collins, L. B.; Swenberg, J. A.; He, C.; Zhang, Y. *Science* **2011**, *333*, 1300.

(8) Wu, J. C.; Santi, D. V. *J. Biol. Chem.* **1987**, *262*, 4778.

(9) Klimašauskas, S.; Kumar, S.; Roberts, R. J.; Cheng, X. *Cell* **1994**, *76*, 357.

(10) Liutkevičiūtė, Z.; Lukinavičius, G.; Masevičius, V.; Daujotytė, D.; Klimašauskas, S. *Nat. Chem. Biol.* **2009**, *5*, 400.

(11) Lukinavičius, G.; Lapinaitė, A.; Urbanavičiūtė, G.; Gerasimaitė, R.; Klimašauskas, S. *Nucleic Acids Res.* **2012**, *40*, 11594.

(12) Kriukienė, E.; Labrie, V.; Khare, T.; Urbanavičiūtė, G.; Lapinaitė, A.; Koncevičius, K.; Li, D.; Wang, T.; Pai, S.; Ptak, C.; Gordevičius, J.; Wang, S.-C.; Petronis, A.; Klimašauskas, S. *Nat. Commun.* **2013**, *4*, 2190.

(13) Hall, R. H. *The modified nucleosides in nucleic acids*; Columbia University Press: New York, 1971.

(14) Maiti, A.; Michelson, A. Z.; Armwood, C. J.; Lee, J. K.; Drohat, A. C. *J. Am. Chem. Soc.* **2013**, *135*, 15813.

(15) Schiesser, S.; Pfaffeneder, T.; Sadeghian, K.; Hackner, B.; Steigenberger, B.; Schröder, A. S.; Steinbacher, J.; Kashiwazaki, G.; Höfner, G.; Wanner, K. T.; Ochsenfeld, C.; Carell, T. *J. Am. Chem. Soc.* **2013**, *135*, 14593.

(16) Suetake, I.; Shinozaki, F.; Miyagawa, J.; Takeshima, H.; Tajima, S. *J. Biol. Chem.* **2004**, *279*, 27816.

(17) Kareta, M. S.; Botello, Z. M.; Ennis, J. J.; Chou, C.; Chédin, F. *J. Biol. Chem.* **2006**, *281*, 25893.

(18) Chen, C.-C.; Wang, K.-Y.; Shen, C.-K. J. *J. Biol. Chem.* **2012**, *287*, 33116.

(19) Liutkevičiūtė, Z.; Kriukienė, E.; Grigaitytė, I.; Masevičius, V.; Klimašauskas, S. *Angew. Chem., Int. Ed.* **2011**, *50*, 2090.

(20) Isono, K.; Asahi, K.; Suzuki, S. *J. Am. Chem. Soc.* **1969**, *91*, 7490.

(21) Liu, A.; Zhang, H. *Biochemistry* **2006**, *45*, 10407.

(22) Kriukienė, E.; Liutkevičiūtė, Z.; Klimašauskas, S. *Chem. Soc. Rev.* **2012**, *41*, 6916.

(23) Schiesser, S.; Hackner, B.; Pfaffeneder, T.; Müller, M.; Hagemeier, C.; Truss, M.; Carell, T. *Angew. Chem., Int. Ed.* **2012**, *51*, 6516.

(24) 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. r.; Boelens, R.; van Ingen, H.; Leonhardt, H.; Carell, T.; Vermeulen, M. *Cell* **2013**, *152*, 1146.

(25) Pastor, W. A.; Pape, U. J.; Huang, Y.; Henderson, H. R.; Lister, R.; Ko, M.; McLoughlin, E. M.; Brudno, Y.; Mahapatra, S.; Kapranov, P.; Tahiliani, M.; Daley, G. Q.; Liu, X. S.; Ecker, J. R.; Milos, P. M.; Agarwal, S.; Rao, A. *Nature* **2011**, *473*, 394.

(26) 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.

(27) Booth, M. J.; Branco, M. R.; Ficz, G.; Oxley, D.; Krueger, F.; Reik, W.; Balasubramanian, S. *Science* **2012**, *336*, 934.

(28) Khare, T.; Pai, S.; Koncevicius, K.; Pal, M.; Kriukiene, E.; Liutkeviciute, Z.; Irimia, M.; Jia, P.; Ptak, C.; Xia, M.; Tice, R.; Tochigi, M.; Moréra, S.; Nazarians, A.; Belsham, D.; Wong, A. H. C.; Blencowe, B. J.; Wang, S. C.; Kapranov, P.; Kustra, R.; Labrie, V.; Klimasauskas, S.; Petronis, A. *Nat. Struct. Mol. Biol.* **2012**, *19*, 1037.

(29) 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.

(30) Lu, X.; Song, C.-X.; Szulwach, K.; Wang, Z.; Weidenbacher, P.; Jin, P.; He, C. *J. Am. Chem. Soc.* **2013**, *135*, 9315.