Excision of 5-hydroxymethylcytosine by DEMETER family DNA glycosylases

Hosung Jang a, Hosub Shin a, Brandt F. Eichman b, Jin Hoe Huh a,*

a Department of Plant Science, Research Institute for Agriculture and Life Sciences, and Plant Genomics and Breeding Institute, Seoul National University, Seoul 151-921, Republic of Korea
b Department of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA

ABSTRACT

In plants and animals, 5-methylcytosine (5mC) serves as an epigenetic mark to repress gene expression, playing critical roles for cellular differentiation and transposon silencing. Mammals also have 5-hydroxymethylcytosine (5hmC), resulting from hydroxylation of 5mC by TET family-enzymes. 5hmC is abundant in mouse Purkinje neurons and embryonic stem cells, and regarded as an important intermediate for active DNA demethylation in mammals. However, the presence of 5hmC in plants has not been clearly demonstrated. In Arabidopsis, the DEMETER (DME) family DNA glycosylases efficiently remove 5mC, which results in DNA demethylation and transcriptional activation of target genes. Here we show that DME and ROS1 have a significant 5hmC excision activity in vitro, although we detected no 5hmC in Arabidopsis, suggesting that it is very unlikely for plants to utilize 5hmC as a DNA demethylation intermediate. Our results indicate that both plants and animals have 5mC in common but DNA demethylation systems have independently evolved with distinct mechanisms.

1. Introduction

DNA methylation is a simple but important epigenetic modification, playing an important role for diverse biological processes such as transcriptional regulation, cellular differentiation, gene imprinting, and silencing of transposable elements [1–4]. In higher eukaryotes, DNA methylation often refers to DNA methyltransferase-catalyzed methylation of the C5 position of cytosine, to generate 5-methylcytosine (5mC).

DNA methylation can be reversed by either passive or active mechanisms [5]. Passive DNA demethylation is replication-dependent and occurs gradually when maintenance DNA methylation is suppressed in dividing cells. By contrast, active DNA demethylation takes place in a replication-independent manner, and requires certain enzyme activities. Recent studies suggest that two different mechanisms may operate for active DNA demethylation. One involves a direct removal of 5mC from DNA, which is consequently replaced with unmethylated cytidine via the base excision repair (BER) pathway. The other implicates an enzymatic modification of 5mC to other bases, which do not require direct excision but nullify the silencing effect of 5mC. Recently, the latter has emerged as an important DNA demethylation pathway in mammals, in which the ten-eleven translocation (TET) family of proteins catalyze the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) [6–12]. Even though the conversion of 5mC to 5hmC may cancel the effect of DNA methylation in part, further TET-dependent oxidation of 5hmC produces 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), both of which can be excised by thymine-DNA glycosylase (TDG) initiating the BER pathway for complete DNA demethylation [7,13].

It is believed that plants and animals have evolved distinct DNA demethylation systems, particularly in that plants have enzymes that are able to remove 5mC from DNA [4,14]. The DEMETER (DME) family of proteins are DNA glycosylases that primarily recognize and excise 5mC from DNA [15–19]. DEMETER (DME) family proteins are DNA glycosylases that primarily recognize and excise 5mC from DNA [15–19]. The DEMETER (DME) family of proteins are DNA glycosylases that primarily recognize and excise 5mC from DNA [15–19]. The DEMETER (DME) family of proteins are DNA glycosylases that primarily recognize and excise 5mC from DNA [15–19]. The DEMETER (DME) family of proteins are DNA glycosylases that primarily recognize and excise 5mC from DNA [15–19].
ribose sugar, and their additional AP lyase activity induces a DNA strand break via the β-elimination process [15–19]. It is likely that BER enzymes subsequently participate in the DNA demethylation process by incorporating unmethylated cytidine in place of excised 5mC [3,4,14]. Consequently, the DME family-initiated BER pathway is regarded as a predominant DNA demethylation mechanism in plants. However, an alternative active DNA demethylation pathway involving 5mC modification as in mammals has not yet been reported in plants.

In this study, we investigated base excision activity of DME and ROS1 for 5mC oxidation derivatives such as 5hmC, 5fC, and 5caC. Despite being less preferred than 5mC, the 5hmC base was found to be significantly excised by DME and ROS1 in vitro, suggesting the possibility of alternative route of DNA demethylation in plants. Additionally, we explored the presence of 5hmC in the Arabidopsis genome in order to understand whether 5hmC excision activity of DME/ROS1 is biologically relevant and associated with DNA demethylation dynamics [22], or simply reflects a broad substrate specificity range of this family of proteins [23].

2. Materials and methods

2.1. Cloning of DME and ROS1

The DMEAN677A/IDR1::lnk fragment (hereafter called “DMEA”, [21]) in which both N-terminal 677 amino acids and IDR1 were removed from DME, was cloned into the Bam HI and Sal I sites of the pLM302 vector (Center for Structural Biology, Vanderbilt University). The ROS1AN509 fragment (hereafter called “ROS1A”) was PCR-amplified from full-length ROS1 cDNA and then introduced into the pLM302 vector at the sites of Eco RI and Sal I. Both DMEA and ROS1A were fused with 6xHis and maltose binding protein (MBP) tags at the N-terminus.

2.2. Protein expression and purification

DMEA and ROS1A fused with 6xHis + MBP tags were expressed in Escherichia coli Rosetta2 (DE3) strains (Novagen). Protein purification steps were essentially the same as described by Mok et al. [21]. Briefly, 6xHis-MBP-DME or -ROS1A was sequentially purified through a HiTrap FF column (GE Healthcare) and a HiTrap Heparin HP column (GE Healthcare), and gel filtration was performed on a HiLoad 16/60 Superdex 200-pg column (GE Healthcare) (Suppl. Fig. 1).

2.3. In vitro base excision assay for 5mC derivatives

Thirty-five-mer oligonucleotides containing cytosine, 5mC, 5hmC, 5fC, and 5caC were purchased from Midlent Certified (TX, USA) (Suppl. Table 1). Forty pmol of each oligonucleotide was radiolabeled with [γ-32P]ATP (6000 Ci/mmol, Perkin Elmer Life Sciences) using T4 polynucleotide kinase (Takara) and then annealed to a complementary oligonucleotide to produce double-stranded DNA substrate. In vitro base excision assay was performed as previously described [21]. Briefly, 25 nM of each radiolabeled oligonucleotide substrate was incubated with 100 nM of MBP-DME or 85 nM of MBP-ROS1A in the glycosylase reaction buffer (10 mM Tris–HCl, pH 7.4, 50 mM NaCl, 0.5 mM dithiothreitol, 200 µg/ml BSA) at 37 °C for 1 h. Reactions were terminated by adding an equal volume of stop solution (98% formamide, 10 mM EDTA, 0.2% xylene cyanol FF, 0.2% bromophenol blue) and heat-denaturing at 95 °C for 10 min. Reaction products were separated on a 15% denaturing polyacrylamide gel containing 7.5 M urea and 1× TBE.

2.4. Kinetics analysis

Twenty-five nanomolar of oligonucleotide substrate containing 5mC or 5hmC was incubated with excess amount of DMEA at a time course manner (0.5, 1, 2, 3, 5, 10, 25, and 60 min). Reaction was terminated by adding 100 mM NaOH and boiling for 10 min. Reactions were denatured and separated on a 15% denaturing polyacrylamide gel. The gel was exposed to a phosphorimager screen (Fujifilm) and the radioactivity was measured using the Fujifilm BAS-5000 phosphorimager.

2.5. Substrate competition assay

To produce the unlabeled competitor substrates with the same concentration of radiolabeled substrates, 5mC- and 5hmC-containing oligonucleotides were purified with the QIAquick Nucleotide Removal Kit (Qiagen) and annealed with complementary oligonucleotides. One hundred nanomolar of MBP-DMEA was incubated with 25 nM of radiolabeled oligonucleotides in the presence of 5mC or 5hmC containing competitor substrates. The amount of competitor oligonucleotides varied from 0 to 20-fold (0, 125, 250, 500 nM) over radiolabeled substrates. After adding an equal volume of stop solution and boiling for 10 min, reactions were separated on a 15% denaturing polyacrylamide gel and analyzed with phosphorimager scanning as described above.

2.6. In vitro nucleotide incorporation assay

Twenty-five nanomolar of radiolabeled oligonucleotide substrate containing 5mC or 5hmC was reacted with 100 nM of MBP-DMEA and 10 U of Endonuclease IV (NEB) in the glycosylase reaction buffer at 37 °C for 1 h. Following heat-inactivation at 65 °C for 15 min, the reaction was subjected to nucleotide incorporation with 0.1 mM dCTP using 5 U of Klenow fragment (3′ to 5′ exo-) (NEB) at 25 °C for 15 min.

2.7. TLC analysis

Genomic DNA was isolated from young leaves, and floral buds of Arabidopsis thaliana Columbia ecotype using the standard CTAB method. Control DNA was prepared by PCR amplification using 5-methyl-dCTP or 5-hydroxymethyl-dCTP (Zymo Research) instead of dCTP in the PCR reaction. TLC analysis was performed using slightly modified protocols from Ito et al. [9]. Briefly, 2 µg of genomic DNA or control DNA was digested with 50 U of TaqI (NEB) and 5 U of Calf Intestinal Alkaline Phosphatase (NEB). After purification with QIAquick Nucleotide Removal Kit (Qiagen), dephosphorylated DNA was end-labeled with 20 µCi of [γ-32P]ATP (6000 Ci/mmol, Perkin Elmer Life Sciences) using T4 polynucleotide kinase (Takara). After purification, radiolabeled DNA was digested with 50 U of Benzonase (Sigma) and 4 µM of Phosphodiesterase I (Sigma) at 37 °C for 1 h. Digested DNA fragments were concentrated and spotted on a PEI-cellulose TLC plate (Merck) and separated for 16 h in the 1D-TLC buffer (isobutyric acid:NH4OH:H2O = 33:1:10). In the second dimension, TLC was performed in the 2D-TLC buffer (isopropanol:HC1:H2O = 70:15:15) for 16 h. The plate was exposed to an X-ray film or analyzed by phosphorimager scanning.

3. Results

3.1. DME and ROS1 excise both 5mC and 5hmC in vitro

The DME/ROS1 family proteins are known to primarily excise 5mC [15–19] initiating the BER pathway for DNA demethylation.
However, like many other DNA glycosylases, they also recognize and excise a broad range of DNA bases including thymine from a T/G mismatch and several modified bases paired with guanine such as 5-fluorouracil, 5-bromocytosine, 5-bromouracil, and 5-hydroxyuracil [16–18,24]. This suggests that DME/ROS1 may have excision activity for some 5mC derivatives that are chemically or enzymatically modified. In particular, a few oxidized bases (5hmC, 5fC, and 5caC; Fig. 1A) catalyzed by the mammalian TET1-family proteins immediately drew our attention because they may serve as intermediates leading to TDG-mediated base excision and DNA demethylation even in plants.

Therefore, we prepared radiolabeled-oligonucleotides containing cytosine oxidation derivatives such as 5hmC, 5fC, and 5caC along with 5mC and performed an in vitro base excision assay with purified MBP-DMEA and MBP-ROS1Δ (Fig. 1B). Because of increased protein stability, we used truncated forms of DME and ROS1 fragments that still retain all conserved domains necessary for in vitro base excision activity [21]. We observed that MBP-DMEA and MBP-ROS1Δ excised both 5mC and 5hmC, generating 3′-phosphor-α,β-unsaturated aldehyde (3′-PUA) as a major product via a β-elimination process, even though the latter was less preferentially processed (Fig. 1C). In addition, MBP-DMEA and MBP-ROS1Δ were able to excise 5caC (Fig. 1C), despite much less excision activity observed. However, we did not observe any discernable activity of MBP-DMEA and MBP-ROS1Δ for 5fC (Fig. 1C). These results suggest that both DME and ROS1 prefer 5mC to other 5mC derivatives and that they have a broad substrate specificity range even though the substituents at C5 of the cytosine ring differ. Notably, the finding that both DME and ROS1 have glycosylase activity for 5hmC suggests not only the possibility of the presence of 5hmC in the plant genome but also an alternative route of active DNA demethylation pathway to remove 5mC.

### 3.2. Reconstitution of DNA demethylation pathway following 5mC or 5hmC excision

We expect that following 5mC excision by the DME/ROS1 family of proteins, subsequent BER enzymes participate in the DNA demethylation process by incorporating unmethylated cytidine in place of excised 5mC. We reconstituted the in vitro BER-mediated DNA demethylation pathway by showing the replacement of 5mC or 5hmC with unmethylated cytidine after DME base excision. When the radiolabeled oligonucleotide containing 5mC was reacted with MBP-DMEA for 1 h, 3′-PUA and 3′-phosphate were produced via β- and δ-elimination processes, respectively (lane 2 in the left panel of Fig. 1D). A subsequent reaction with AP endonuclease converted both products to 3′-OH (lane 3 in the left panel of Fig. 1D). Finally, the dCTP incorporation at the site of base excision was achieved by Klenow DNA polymerase (lane 4 in the left panel of Fig. 1D). Similar results were obtained when 5hmC-containing oligonucleotide was subjected to the same series of reactions (right panel of Fig. 1D). These results demonstrate that DME has the capacity to initiate active DNA demethylation after 5mC or 5hmC excision via the BER pathway.

### 3.3. DME has a lower 5hmC excision rate over 5mC

To compare base excision efficiencies of DME for 5mC and 5hmC, a time-dependent kinetics analysis was performed (Fig. 2). In order to measure the formation of single base excision products, reactions were terminated at various time points and treated with strong base (100 mM NaOH) to obtain homogeneous abasic sites, we obtained kcat values under single-turnover conditions with enzyme in excess of substrate (see Section 2 for experimental details). The single turnover rate...
constants ($k_{cat-st}$) for 5mC and 5hmC were 0.049 min$^{-1}$ and 0.019 min$^{-1}$, respectively, indicating that DME has a 2.6-fold higher rate of base excision for 5mC over 5hmC (Fig. 2).

### 3.3.1. 5mC is more preferred substrate for DME than 5hmC

In order to assess differential DME substrate preference between 5mC and 5hmC, we performed a substrate competition assay (Fig. 3A). When the radiolabeled oligonucleotide substrates containing 5mC or 5hmC were reacted with MBP-DME in the presence of varying amounts of unlabeled oligonucleotide competitors, the product formation decreased as the amount of cold competitor increased (Fig. 3A). Notably, the product formation for 5mC was more affected by the competitors containing 5mC than 5hmC (left panels of Fig. 3A and B). Similar results were obtained when radiolabeled 5hmC oligonucleotides were competed with 5mC- or 5hmC-containing cold oligonucleotides (right panels of Fig. 3A and B). These results indicate that 5mC is a more preferred substrate for DME than 5hmC.

### 3.4. No 5hmC is detected in plant DNA

The 5hmC excision activity of DME/ROS1 raises the possibility of the presence of 5hmC in the plant genome that can be utilized as another DNA demethylation substrate. In order to explore the existence of 5hmC in plant DNA, we extracted genomic DNA from several tissues in Arabidopsis including floral buds where DME is primarily expressed [20]. Isolated genomic DNA was treated with TaqI and calf intestine phosphatase and radiolabeled with polynucleotide kinase (Suppl. Fig. 2). After digestion with Benzonase and Phosphodiesterase I, the nucleotides were separated on a TLC plate. In the 1D TLC analysis, compared to the control experiment (Fig. 4A), no spots corresponding to 5hmC were detected in any tissues examined (Fig. 4B). Even in the 2D TLC analysis, a small amount of 5hmC (0.2% of 5hmC approximately equal to its content in mammalian genome [9]) was identifiable (Fig. 4C), but no 5hmC signal was detected in plant tissues (Fig. 4D and Suppl. Fig. 3). This implies that plants have no detectable amounts of 5hmC in DNA or very little, if any. This also suggests that 5hmC may not be used as a DNA demethylation substrate in plants by the DME/ROS1 family of proteins even though they have 5hmC excision activity.
4. Discussion

Here we report that the DME/ROS1 5mC DNA glycosylases excise 5mC oxidation derivatives such as 5hmC and 5caC. After 5mC or 5hmC excision, DME and ROS1 produce both 3′-PUA and 3′-phosphate via β- and δ-elimination processes, respectively (Fig. 1C). The 3′-end structures of these excision intermediates are converted to 3′-OH by AP endonuclease, allowing downstream DNA polymerase to incorporate unmethylated cytidine for the replacement of 5mC with cytosine (Fig. 1D). DME excises a significant amount of 5hmC although its excision rate is approximately 2.6-fold lower than that of 5mC (Fig. 2). The substrate competition assay also demonstrates that DME clearly has a preference for 5mC over 5hmC (Fig. 3). Despite our finding that DME/ROS1 proteins effectively remove 5hmC, we could not detect 5hmC from the plant DNA from TLC analysis (Fig. 4), implying that plants have no or very little 5hmC in the genome, if any, and may not utilize 5hmC as a primary 5mC-oxidation intermediate for DNA demethylation as in mammals.

Therefore, the 5hmC excision activity of DME/ROS1 may simply reflect a broad range of substrate specificity of this family of proteins, which is a common property of most DNA glycosylases [23,25,26]. The substrate specificity is determined by the recognition pocket of DNA glycosylase. The size of the electron cloud at C5 of 5hmC is larger than that of 5mC [27]. Therefore, 5hmC would need a larger recognition pocket than 5mC, which may explain why DME and ROS1 prefer 5mC to 5hmC. It was previously reported that ROS1 can excise several bases including 5-hydroxypuracil [24]. However, the finding that 5-hydroxymethyluracil was hardly excised by ROS1 [24] suggests the importance of the amine group at C6 of 5mC, possibly for thermodynamic stability of the target base via hydrogen bonding with amino acids inside the recognition pocket during base excision.

A recent study reported that a low but detectable amount of 5hmC was present in the Arabidopsis genome, which was demonstrated by immunoblotting methods using anti-5hmC antibody [28]. This raises the possibility of 5hmC serving as a DNA demethylation intermediate in plants, which can be processed by 5hmC-specific DNA glycosylases including DME/ROS1 family proteins. However, the approach that Yao et al. [28] used might not be sensitive enough to reflect the real base composition of plant genome due to non-specificity of antibody-based methods. In addition, no functional counterparts of TET-family enzymes or hydroxylating enzymes acting on 5mC have yet been identified in plants. Alternatively, 5hmC can be spontaneously produced by oxidative damages resulting from reactive oxygen species such as superoxide anions (O2•−), hydroxyl radicals (•OH), and hydrogen peroxide (H2O2) [29]. Therefore, a trace amount of 5hmC can be present in the plant genome with no aid of coordinating enzyme activities. It is still an intriguing question to be answered how plants and animals have evolved distinct DNA demethylation systems, in particular, the evolution of 5mC-specific DNA glycosylases in plants and the use of 5mC-oxidation derivative(s) intermediates mediating DNA demethylation in mammals.

To the best of our knowledge, DME/ROS1 proteins are the first reported DNA glycosylases that recognize and excise 5hmC from DNA. Even though its biological relevance in plants is still under question due to a lack of clear evidence for the existence of 5hmC in the plant genome, DME/ROS1 proteins have a promising potential for epigenome editing [30]. In particular, active DNA demethylation at specific targets combined with a transcription activator-like effector (TALE) or CRISPR system may allow transcriptional activation of the genes that are silenced by DNA methylation. A recent report demonstrated that TALE-TET1 fusion proteins successfully induced targeted DNA demethylation and gene expression [31]. Considering the fact that DME/ROS1 proteins remove both 5mC and 5hmC, the approach using these proteins can be used for epigenome editing in plants and mammals because 5mC is a universal silencing mark utilized in both systems.

Acknowledgments

This work was supported by a Grant from the Next-Generation BioGreen 21 Program (No. PJ009092), Rural Development Administration, and the Global Ph.D. Fellowship (NRF-2013H1A2A1034385, H.J.), National Research Foundation, Republic of Korea.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2014.03.060.

References


