Mutation Research xxx (2013) xxx–xxx

Contents lists available at ScienceDirect

Mutation Research/Genetic Toxicology and Environmental Mutagenesis

journal homepage: www.elsevier.com/locate/gentox
Community address: www.elsevier.com/locate/mutres

TET enzymatic oxidation of 5-methylcytosine, 5-hydroxymethylcytosine and 5-formylcytosine

Jean Cadet[a,b,*], J. Richard Wagner[b]

[a] Direction des Sciences de la Matière, Institut Nanosciences et Cryogénie, CEA/Grenoble, 38054 Grenoble, France
[b] Département de médecine nucléaire et radiobiologie, Faculté de médecine et des sciences de la santé, Université de Sherbrooke, Québec JIH 5N4, Canada

ARTICLE INFO

Article history:
Received 1 September 2013
Accepted 4 September 2013
Available online xxx

Keywords:
Epigenetic marks
Radical oxidation reactions
DNA glycosylase-mediated repair
Hydrolytic and enzymatic deamination
Biological role

ABSTRACT

5-Methylcytosine and methylated histones have been considered for a long time as stable epigenetic marks of chromatin involved in gene regulation. This concept has been recently revisited with the detection of large amounts of 5-hydroxymethylcytosine, now considered as the sixth DNA base, in mouse embryonic stem cells, Purkinje neurons and brain tissues. The dioxygenases that belong to the ten eleven translocation (TET) oxygenase family have been shown to initiate the formation of this methyl oxidation product of 5-methylcytosine that is also generated although far less efficiently by radical reactions involving hydroxyl radical and one-electron oxidants. It was found as additional striking data that iterative TET-mediated oxidation of 5-hydroxymethylcytosine gives rise to 5-formylcytosine and 5-carboxylcytosine. This survey focuses on chemical and biochemical aspects of the enzymatic oxidation reactions of 5-methylcytosine that are likely to be involved in active demethylation pathways through the implication of enzymatic deamination of 5-methylcytosine oxidation products and/or several base excision repair enzymes. The high biological relevance of the latter modified bases explains why major efforts have been devoted to the design of a broad range of assays aimed at measuring globally or at the single base resolution, 5-hydroxymethylcytosine and the two oxidation products in the DNA of cells and tissues. Another critical issue that is addressed in this review article deals with the assessment of the possible role of 5-methylcytosine oxidation products, when present in elevated amounts in cellular DNA, in terms of mutagenesis and interference with key cellular enzymes including DNA and RNA polymerases.

Abbreviations: α-KG, α-ketoglutarate; AML, acute myeloid leukemia; APOBEC, apolipoprotein B mRNA-editing enzyme complex; ARB, aldehyde reactive probe; β-GT, β-glucosyl transferase; 5-caC, 5-carboxylcytosine; 5-cadCyd, 5-carboxyl-2′-deoxycytidine; C, cytosine; C5-MTases, cytosine-5 methyltransferases; DNMT, DIP, DNA immunoprecipitation; DNA, methyltransferase; DSB4, double stranded β-helix; CAN, ceric(IV) ammonium nitrate; $CO_3^{\bullet -}$, carbonate anion radical; CE-LIF, capillary electrophoresis with laser-induced fluorescence; CpGIs, CpG islands; DKO, double-knockout; dUrd, 2′-deoxyuridine; ESI-MS, electrospray ionization-mass spectrometry; 5-fC, 5-formylcytosine; 5-fdCyd, 5-formyl-2′-deoxycytidine; 5-fU, 5-formyluracil; FTO, obesity-associated protein; 5-glu-5-hmC, glucosylated derivative of 5-hydroxymethylcytosine; HEK, human embryonic kidney; hMLH1, human mut L homolog 1; HPLC-ESI-MS/MS, high-performance liquid chromatography-electrospray ionization-tandem mass spectrometry; 5-hmC, 5-hydroxymethylcytosine; 5-hmdCyd, 5-(hydroxymethyl)-2′-deoxycytidine; 5-hpmdCyd 5-methylcytosine, 5-mC, 5-hydroxymethyluracil, 5-hmU; 5-hpmdCyd, 5-hydroperoxymethyl-2′-deoxycytidine, 5-iodo-2′-deoxycytidine, 5-IdCyd; hNTH1, human nth endonuclease III-like 1; hOGG1, human oxoguanine DNA glycosylase 1; hUNG2, human uracil DNA glycosylase 1; IDH, isocitrate dehydrogenase; LCK, leukemia-associated protein with CXXC domain; MBD4, methyl-binding domain protein 4; MQ, 2-methyl-1,4-naphthoquinone; MGMT, O^6-methylguanine-DNA-methyltransferase; MRM, multiple reaction monitoring; $Na_2S_2O_8$, sodium peroxosulfate; NP, neuronal progenitor; •NO, nitric oxide; $^1O_2$, singlet oxygen; $O_2^{\bullet -}$, superoxide anion radical; •OH, hydroxyl radical; ODN, oligodeoxynucleotide; oxBS-Seq, oxidative bisulfite sequencing; Pol II, polymerase II; ROS, reactive oxygen species; RRBS, reduced representative bisulfite sequence; SAM or AdoMet, *S*-adenosyl-L-methionine; SMUG1, single-strand specific monofunctional DNA glycosylase I; SMRT, single-molecule real-time DNA sequencing; T7H, TBDMS, tert-butyldimethylsilyl; T7H, thymine 7-hydroxylase; TDG, thymine DNA glycosylase; TET, Ten eleven translocation; UDP-glucose, uridine 5′-diphospho-D-glucose.

* Corresponding author.
E-mail addresses: jean.cadet@cea.fr (J. Cadet), richard.wagner@usherbrooke.ca (J.R. Wagner).

1383-5718/$ – see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.mrgentox.2013.09.001

1. Introduction

DNA methylation and histone modifications constitute well documented inheritable changes in chromatin that participate in the epigenetic fate of cells and repression of intragenic transcription [1–4]. DNA methylation is mediated in mammals by cytosine-5 methyltransferases (C5-MTases) [5,6] and five enzymes of the DNA methyltransferase (DNMT) family [7–9] that allow the transfer of a methyl group from the ubiquitous co-factor *S*-adenosyl-L-methionine (SAM or AdoMet) to the carbon 5 of cytosine, preferentially in the context of CpG motifs while CpG islands (CGIs) are typically unmethylated [5–9]. C5 Methylation of cytosine has also been shown to occur at CpA sites in oocytes [10]. This gives rise to 5-methylcytosine (5-mC), identified as the fifth nucleobase [11,12] that was discovered more than 60 years ago [13,14] and considered until recently as a crucial and relatively stable epigenetic mark [15,16]. 5-mC has been found to impact several essential biological pathways including among others regulation of gene expression, maintenance of chromatin structure, parental imprinting, X-chromosome inactivation, control of cell development and disease pathogenesis [17–21]. 5-Hydroxymethylcytosine (5-hmC), an oxidation product of 5-mC that can be generated by either hydroxyl radical (•OH) or one-electron oxidants [22] was initially found in T-even bacteriophage [23,24] and mammalian cells [25] using either mild acid hydrolysis or enzymatic digestion of DNA combined with paper chromatography before being almost completely ignored after unsuccessful attempts to detect this oxidized base in vertebrate DNA [26]. However the concept of 5-mC being a stable and static epigenetic modification has been questioned after the rediscovery in 2009 of 5-hmC found in substantial amounts in genomic DNA of mouse embryonic stem cells [27] and in mouse Purkinje neurons and brain [28] by two independent groups. In addition, ten eleven translocation 1 (TET1), one of three enzymes of the TET family, was shown to be involved in the oxidation of 5-mC into 5-hmC [26], now defined as the sixth nucleobase [29–31]. These major findings in the field of epigenetics have provided a strong impetus toward numerous and extensive studies aimed at better assessing the ability of TET enzymes to oxidize 5-mC in various cells and tissues and determining the role of 5-hmC in the epigenome. Another major issue that rapidly emerged was the possible implication of 5-hmC and further oxidation products including 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC) (for recent review articles, see [32–34]) as short-lived intermediates of active removal pathways of 5-mC and as a more dynamic epigenetic regulation with respect to the well documented passive demethylation process involving DNA replication [35–37]. The available data reported during the last three years on the formation of 5-mC oxidation products and the proposed demethylation mechanisms, two major fast evolving research topics, are critically reviewed with emphasis on chemical and biochemical aspects. This review article starts by a critical survey of the main methods involving chromatographic, enzymatic and immuno-based analytical techniques that have been developed for globally measuring 5-mC oxidation products in the DNA of isolated cells and tissues. They are described and evaluated in terms of accuracy and sensitivity since the reliability of some of the reported assays have been questioned [38]. Similar to unprecedented efforts applied to measure oxidatively generated damage to DNA, several methods have been designed for mapping 5-hmC at the nucleotide level in order to gain further insights into the putative implication of the latter oxidized pyrimidine base in active demethylation pathways. This review is completed by presentation of the main available data on the depletion of the levels of 5-hmC in cancer tissues that could be used as a potential prognostic biomarker. In addition, recently available information is provided on the biological effects of 5-mC oxidation products including their low mutation potential. It was recently shown that N^6-methyladenine, an alkylated base present in substantial amounts in messenger RNA but not in the DNA of mammals, is also subjected to demethylation, a dynamic reversible process likely to be implicated in cellular regulation [39]. This involves, as for the demethylation of 5-mC by TET proteins [40], enzymes that belong to the iron (II)- and α-ketoglutarate (α-KG) dependent dioxygenase family proteins. Thus, the fat mass and obesity-associated protein FTO [39,41] and ALKBH5 [42] in mammals are able to demethylate RNA both *in vitro* and *in vivo*. Oxidative demethylation is expected to give rise to unstable N^6-hydroxymethyadenine intermediate exhibiting an acetal structure that is highly susceptible to hydrolysis leading to restitution of initial adenine as previously observed for other ALKB enzymes that are involved in the repair of N^1-alkyladenine [42–44]. These reactions will not be discussed further since emphasis in the present survey is placed on the oxidation products of 5-mC and their biochemical features.

2. Analysis of oxidation products of 5-methylcytosine

Major efforts have been devoted during the last 3 years toward the development of methods aimed at measuring methyl oxidation products of 5-mC in cells and tissue after the rediscovery of 5-hmC in ES cells and neuronal tissues that was achieved using poor resolution and semi-quantitative restriction enzyme based assays. More accurate and quantitative analytical approaches including high-performance liquid chromatography (HPLC) based methods have become available whose benefits and limitations are critically reviewed here. In addition a strong impetus has been given to the design of several sequencing techniques aimed at specifically mapping 5-mC and the three oxidation products at the nucleotide level for gaining more insight into the functional role of the latter epigenic marks. A comparison of the different methods is presented in Table 1.

2.1. Thin-layer chromatography analysis of [^{32}P]-labelled 5-mC oxidation products

Two restriction enzyme based methods that included sequence specific digestion of DNA followed by a ^{32}P-phosphorylation step before thin-layer chromatography (TLC) analysis of [^{32}P]-radiolabelled nucleoside monophosphates were used for the discovery of 5-hmC in mouse ES and neuronal cells [27,28]. Application of an adaptation of the so-called nearest-neighbour analysis of nucleoside X at XpG sites [45] required incubation of DNA with *Fok 1* restriction enzyme and detection of the novel 3′-monophosphate of 5-hmC by 2D-TLC analysis that was later confirmed by ESI-MS analysis [28]. The second assay leading to the isolation of 5-(hydroxymethyl)-2′-deoxycytidine 5-monophosphate and characterization by ESI-MS analysis involved DNA digestion by Msps1 restriction enzyme [27]. In subsequent studies, either Taq$^{\alpha}$1 [46] or EcoNI [47] was used as the restriction enzyme to reveal in cellular DNA 5-fC and 5-caC that were shown to prevent Taq$^{\alpha}$1-mediated DNA cleavage. Global measurement of the three 5-mC oxidation products provided by the latter Taq$^{\alpha}$1 assay in CpG contexts constitutes a major advantage although it is counterbalanced by a lack of accurate quantification.

2.2. Enzymatic derivatisation 5-hmC based assays

Another global enzymatic assay aimed at measuring 5-hmC sites in DNA is based on glucosylation of the 5-hydroxymethyl group, which is achieved by T4 phage β-glucosyl transferase (β-GT) using uridine 5′-diphospho-D-glucose (UDP-glucose) as the co-factor [48]. The incorporation of [^{3}H]-labelled glucose into

Please cite this article in press as: J. Cadet, J.R. Wagner, TET enzymatic oxidation of 5-methylcytosine, 5-hydroxymethylcytosine and 5-formylcytosine, Mutat. Res.: Genet. Toxicol. Environ. Mutagen. (2013), http://dx.doi.org/10.1016/j.mrgentox.2013.09.001