Biochimica et Biophysica Acta 1819 (2012) 757–762

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbagrm

Review

The ghosts in the machine: DNA methylation and the mystery of differentiation ☆

Victorino Briones, Kathrin Muegge *

Laboratory of Cancer Prevention, SAIC-Frederick, Inc., NCI-Frederick, Frederick, MD, USA

ARTICLE INFO

Article history:
Received 16 November 2011
Received in revised form 1 February 2012
Accepted 9 February 2012
Available online 22 February 2012

Keywords:
DNA methylation
SNF2
Embryogenesis

ABSTRACT

Methylation regulates DNA by altering chromatin and limiting accessibility of transcription factors and RNA polymerase. In this way, DNA methylation controls gene expression and plays a role in ES cell regulation, tissue differentiation and the development of the organism. In abnormal circumstances methylation can also induce diseases and promote cancer progression. Chromatin remodeling proteins such as the SNF2 family member Lsh regulates genome-wide cytosine methylation patterns during mammalian development. Lsh promotes methylation by targeting and repressing repeat sequences that are imbedded in heterochromatin. Lsh also regulates cytosine methylation at unique loci. Alterations in histone modifications (such as H3K4me3, histone acetylation, H3K27me3 and H2Aub) can be associated with DNA methylation changes making Lsh-mediated cytosine methylation part of a larger epigenetic network defining gene expression and cellular differentiation during development. This article is part of a Special Issue entitled: Chromatin in time and space.

Published by Elsevier B.V.

1. What is cytosine methylation?

Cytosine methylation involves a covalent modification at the carbon 5 position of the cytosine base [1,2]. In somatic mammalian cells cytosine methylation occurs preferentially in the CpG context while in ES cells, interestingly, it is also present at non-CpG sites [3]. Most eukaryotic genomes contain cytosine methylation but the distribution greatly varies between organisms [4,5]. For example, in honey bees cytosine methylation is almost exclusively concentrated at genes. In contrast, in mammalian cells genes as well as intergenic regions are highly methylated with the exception of so-called CpG islands, these are short CG rich regions often located around transcriptional start sites.

There are several important waves of global cytosine methylation changes in mammals [4,6,7]. First, genome-wide erasure of cytosine methylation occurs in primordial germ cells between days 10.5 and 13.5 during murine gestation. This is followed by re-methylation of the genome and establishment of a gender specific methylation pattern at imprinted sites (DMR = differentially methylated regions). A second wave of genome-wide reduction of methylation happens briefly after fertilization, and is based, in part, on an active process of de-methylation of the male genome. After implantation, re-methylation occurs and is associated with cellular differentiation suggesting that specific patterns are established in distinct tissues. In addition, large scale changes in cytosine methylation are observed after reprogramming and generation of iPS cells using overexpression of the four "Yamanaka" factors Oct4, Sox2, Klf4 and c-myc [8]. Thus methylation patterns in iPS cells resemble closely those in ES cells but differ from somatic tissues suggesting that specific cytosine methylation patterns mark pluripotency [3,8,9].

2. What are the functional consequences of DNA methylation?

DNA methylation plays a role in genomic imprinting (or parental allele specific expression), it regulates X inactivation and contributes to tissue specific gene expression patterns [1,2,7]. Deletions of the major DNA methyltransferases, Dnmt1 and Dnmt3b, lead to early lethality during embryogenesis [10,11]. In addition, enzymes that are involved in de-methylation, including iterative oxidation of methylated cytosine and subsequent base excision by repair enzymes, are crucial and their targeted deletion in mice also results in embryonic lethality [12,13]. This suggests that cytosine methylation plays an important role in development and the findings are consistent with a model of DNA methylation as part of the epigenetic memory.

However, several questions remain unresolved. Although tissue specific methylation is in part associated with gene expression, the cause and consequences of DNA methylation in the process of transcription remain undetermined. Although CG methylation is generally thought to result in gene silencing, particularly in cancer cells at tumor suppressor genes [14], methylation of the gene body is observed throughout most of the animal kingdom [4,5] and does not correlate with gene expression in somatic cells [3,8]. The loss of Dnmt1 in cultured cells results in both the up- and downregulation of many genes [15]. Genes that are directly targeted by Dnmt3a can

☆ This article is part of a Special Issue entitled: Chromatin in time and space.

* Corresponding author at: Laboratory of Cancer Prevention, National Cancer Institute, BLDG 560, Room 12-91, Frederick, MD 21701. USA. Tel.: +1 301 846 1386; fax: +1 301 846 7077.
E-mail address: Kathrin.Muegge@nih.gov (K. Muegge).

1874-9399/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.bbagrm.2012.02.013

be up- or down-regulated upon loss of Dnmt3a bringing cytosine methylation in a more complex transcriptional context [16]. Recent analysis in ES cells suggests a positive correlation between non-CG methylation and gene expression in ES cells [3]. Although, there is currently no hypothesis linking non-CpG methylation to gene expression, several mechanisms have been described that connect CG methylation to transcriptional repression [17]. Binding of transcription factors may be modulated by cytosine methylation and thus alter transcriptional initiation. Recognition of methyl-cytosine or unmethylated CG sites can result in histone modifications that modulate transcription. For example, interaction of DNA with methyl DNA binding proteins can lead to HDAC recruitment and hypoacetylated chromatin is associated with repression. Specific methyl–DNA binding proteins (MecP2 and CTCF) can also affect splicing and Pol II stalling, which then compromises Pol II elongation [18,19]. Finally, cytosine methylation may contribute to nucleosomal positioning [20] and ultimately to changes in chromatin structure and nuclear architecture.

3. The significance of cytosine methylation

The enzymatic machinery that maintains cytosine methylation patterns at the replication fork consists of Dnmt1 and the hemi-methylation binding protein Uhrf1 [2]. Patterns of cytosine methylation appear stable, making it an attractive mechanism to participate in the epigenetic memory (although at specific genomic sites cyclic methylation and de-methylation have been reported [21]). Large scale cytosine methylation changes (up to several Mb in size) as well as site specific changes including the promoter region of pluripotency genes such as Oct4 or Nanog, accompany the transition from pluripotent cells to somatic cells or in the reverse process from somatic to iPS cells (refer to Fig. 1 for a general summary of transitional epigenetics events occurring in ES cells, somatic cells and iPS cells) [3,8,9,22]. This suggests a functional role for cytosine methylation during cellular differentiation, reprogramming and possibly regenerative biology. In addition, frequent observations of genomic hypomethylation and site specific hypermethylation at tumor suppressor genes in cancer suggest a role in tumorigenesis [14]. Moreover, the possibility of transgenerational epigenetic inheritance [23] suggests another avenue to explore aside from genetic inheritance of familial diseases. In this manner, identifying factors that play a critical role in cytosine methylation raise the prospect of controlled modulation of the epigenetic memory for therapeutic purposes.

4. The physiologic role of Lsh

Murine Lsh was first cloned using a degenerative PCR technique to amplify novel helicase super family members in T cell precursors [24]. The gene is a member of the SNF2 subfamily of helicases, which largely consist of chromatin remodeling proteins (Fig. 2). Because of the prominent expression profile of murine Lsh in proliferating T or B cells, it was termed Lsh (lymphoid specific helicase) [24–26], although, Lsh mRNA has been detected at low levels in many tissues. The human gene has been cloned from human leukemic cells and is also known as PASG (proliferation associated gene) [27], other names for Lsh are HELLS (helicase, lymphoid specific) or SMARCA6 (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 6).

Lsh−/− mice are embryonic lethal [28,29], pointing to an important role during development which is not shared by all SNF2 homologues. For example, neither mice with a targeted deletion of the SNF2 homologues Brm or Rad54 are lethal [30–32]. Lsh−/− mice have multiple developmental defects, kidney necrosis, reduced embryonal growth, aberrant gene expression of various Hox genes, signs of premature aging, and early senescence of fibroblasts [28,29,33,34]. In addition, a defect in the generation of stem cells has been observed in multiple tissues: neither male nor female germ cells thrive [35,36]; there is a delay or incomplete differentiation of ES cell differentiation in *in vitro* cultures [37] and an impaired lymphoid development and defects in hematopoiesis [26,38]. In short, Lsh plays a unique role in murine development, and is a non-redundant SNF2 family member.

Members of the SNF2 family disrupt histone–DNA interactions and perform chromatin remodeling in part via nucleosomal sliding and by altering the accessibility to nucleosomal DNA. In line with this role as a SNF2 family member, Lsh is found exclusively in the nuclear compartment and associates with chromatin [39]. It localizes at heterochromatic regions and deletion of Lsh alters chromatin structure at heterochromatic sites [39,40]. In particular, Lsh controls cytosine methylation, and deletion of Lsh shows a 50% reduction of cytosine methylation as measured, for example, by HPLC [41–43]. This property of Lsh (to affect cytosine methylation) is shared with DDM1, the Lsh homologue in *A. thaliana*. Indeed, DDM1 has been identified based on the occurrence of DNA hypomethylation in mutants (decrease in DNA methylation 1) [44]. In addition, Lsh deletion alters H3K4me3 level [40], which is again a phenotype shared with DDM1 mutants in *A. thaliana* [45]. However, it is not yet known whether the increase in H3K4me3 is due to DNA hypomethylation, or if DNA hypomethylation follows H3K4me3 increases. For example, Cfp1 is a DNA binding protein recognizing unmethylated CpG islands and recruiting the H3K4me3 methyltransferase Setd, thus connecting hypomethylated DNA to a rise in H3K4me3 [46]. On the other hand, DNMTs associate preferentially with H3 histone tails devoid of H3K4me3 modification and thus linking a decrease of H3K4me with methylated DNA [47,48]. In addition to cytosine methylation, changes in H3K27me3 and H2AK116 ubiquitylation have been observed at specific loci in Lsh−/− cells [34]. This would suggest, at least in part, functional interaction of Lsh with other epigenetic pathways such as the Polycomb silencing pathway.

5. Maintenance of methylation versus *de novo* cytosine methylation

The distinction between both pathways is important since for therapeutic purposes one would like to interfere mostly with *de novo* methylation, e.g. by blocking aberrant *de novo* methylation at tumor suppressor genes or by controlling site specific *de novo* methylation during cellular differentiation for use in regenerative medicine. There may be a partial overlap of the two pathways. For example, Dnmt1 controls maintenance, since it resides at the replication fork, and it efficiently methylates hemi-methylated DNA and is supported by the hemi-methylation binding protein Uhrf1 [2,49]. On the other hand, it may not be exclusively involved in maintenance since many reports have shown recruitment of Dnmt1 to genomic sites via interaction with specific transcription factors suggesting a role in *de novo* methylation [50,51]. Conversely, the *de novo* methyltransferases Dnmt3a and Dnmt3b, may contribute to maintenance since deletion in cell lines results in a moderate loss of cytosine methylation at some repetitive sequences [52]. Several observations suggest that Lsh primarily aids in *de novo* methylation. For example, Lsh is required for *de novo* methylation of retroviral sequences introduced into cell lines and is not obligatory for maintenance of *in vitro* pre-methylated episomal DNA [53]. Also, Lsh does not co-localize with Dnmt1 or does not localize at sites of replication in early S-phase which would be expected for a role in maintenance [39]. Moreover, during *in vitro* culture Lsh has been shown to be required for complete establishment of cytosine methylation at pluripotency genes such as Oct4 or Nanog [37]. Partial depletion of Lsh in ES cell cultures compromises silencing of pluripotency genes and delays their expression during embryogenesis [37]. Furthermore, Lsh is not in general required for genomic imprints as would be expected for a functional role in maintenance [54]. Finally, global analysis of cytosine methylation in Lsh−/− MEF cell lines demonstrates discrete