DNA methylation and control of gene expression in vertebrate development

Richard R. Meehan and Irina Stancheva

Genes and Development Group, University of Edinburgh, Department of Biomedical Sciences, Hugh Robson Building, George Square, Edinburgh EH8 9XD, Scotland, U.K.

Introduction

DNA methylation has the ability to repress transcription; in addition, the pattern of methylation can be stably inherited during successive cell divisions [1]. These two features make methylation very attractive as a potential regulator of gene expression during the development of an organism. A combination of biochemical and genetic analysis has elucidated the potential mechanisms by which methylation can inhibit transcription, and has shown that DNA methylation is necessary for normal development in many vertebrates [2–4]. Indeed a broader view would indicate that DNA methylation has been successfully utilized in many eukaryotes, including fungi, plants and animals, as a regulator of gene activity [1]. What is not clear is whether the specific role of DNA methylation is conserved between these different species or whether methylation has been adapted to regulate different aspects of gene expression in diverse species. This chapter will attempt to address this question, principally by comparing the effects of depletion of DNA methylation on the development of vertebrates: mice (Mus musculus), zebrafish (Danio rerio) and the toad (Xenopus laevis).

To whom correspondence should be addressed (e-mail: Richard.Meehan@ed.ac.uk).
DNA methylation in vertebrates

Vertebrate DNA is methylated at the fifth position of cytosine (\(5\text{mC}\)) in the dinucleotide CpG. This is carried out enzymically by DNA methyltransferases either immediately after passage of the replication fork or as a result of repair processes [1]. Approximately 70% of all methylated CpGs (\(\text{MeCpGs}\)) are found in regions of the genome that are transcriptionally inactive and late replicating. An early immuno-histochemical study with an anti-\(5\text{mC}\) antibody showed that mouse heterochromatin is rich in methylated sequences [5]. Interest in the role of DNA methylation in controlling gene expression was stimulated by the finding that many tissue-specific genes are methylated in non-expressing tissues [6]. However, the view that DNA methylation is necessary for gene inactivation was challenged when it became clear that there are many tissue-specific genes that are never methylated even in tissues where they are not expressed [7]. A general view is that DNA methylation acts as a global inhibitor of transcription due to the location of \(5\text{mC}\) throughout the genome, thus preventing background transcriptional noise or the spontaneous activation of normally silent chromosomal regions [7]. This does not rule out the possibility that specific sets of genes are regulated by DNA methylation at different points in development.

The importance of DNA methylation in mammalian development was emphasized when it was demonstrated that the targeted disruption of the maintenance DNA methyltransferase (\(Dnmt1\)) gene in mice resulted in embryonic lethality during early stages of development [2,8]. Several genes were mis-expressed in the \(Dnmt1^{-/-}\) embryos [9], probably due to loss of \(5\text{mC}\). However, tissue-specific genes were not activated in these mutants and, moreover, were not methylated during the early stages of development in wild-type mice [7]. This implies that the observed tissue-specific methylation patterns are a late developmental event and that DNA methylation may occur as a result of gene inactivation. It is not clear why the \(Dnmt1^{-/-}\) mutants die, although a strong possibility is that it is due to the mis-expression of normally dormant methylated genes and the abnormal inactivation of the X chromosome(s) in male and female mice. Indeed it has been argued on this basis that DNA methylation has been adapted for a specialized role in mammalian development, controlling the expression of imprinted genes and X-inactivation, rather than being a general regulator of gene activity during embryogenesis [7,9]. One way to examine these possibilities is to compare the role of DNA methylation in different vertebrate species and to ask whether there are any similarities or differences. However, before we do that we should explore how DNA methylation can repress gene expression.

Gene repression by DNA methylation

It is very clear that methylation of genes \textit{in vivo} and \textit{in vitro} results in transcriptional silencing, the degree of which may depend upon the location
and density of the CpGs relative to the promoter [10]. Several theories can account for how $^\text{Me}$CpG can interfere with gene expression. First, methylation may directly prevent a transcription factor from recognizing its binding sites, because the $^5\text{mC}$ changes the recognition sequence [10]. However, many transcription factors (for example Sp1) bind independently of the methylation status of their recognition sequences and the recognition sequences, for many other transcription factors do not contain a CpG. Secondly, methylation may lead to the formation of a localized chromatin structure that is incompatible with gene expression [11,12]. Time course experiments have demonstrated that repression of transcription from methylated templates is not immediate. Methylated genes are remodelled over several hours into a structure that leads to their repression [13]. The preferential condensation of methylated chromatin into a higher-order structure is an attractive idea and has received support from an initial finding that methylated DNA is preferentially located in nucleosomes containing histone H1 [14]. However, with one notable exception, CpG methylation has little or no effect on the capacity of the histone octamer to interact with DNA [15]. Finally, it is possible that nuclear factors preferentially interact with methylated DNA leading to the formation of inactive chromatin (Figure 1). Such factors have been identified in many species and have been termed methylated-CpG-binding proteins (MeCPs) [10]. The next section discusses their identification and how their repressing activity is linked to chromatin-remodelling complexes.

**Figure 1.** $^\text{Me}$CpG-binding proteins (MeCP1, MeCP2, MBD1 and MBD2) can target chromatin-remodelling complexes to methylated sequences

These complexes (NuRD and Sin3A) contain histone deacetylase (HDAC) activities which promote the establishment of an inactive chromatin conformation by removing acetyl groups from acetylated histone tails (NH-Ac-). Black dots indicate $^\text{Me}$CpGs which act as ligands for the MeCPs.
Methylated-DNA-binding proteins

The first factor that showed a binding preference for methylated DNA was identified in a filter-binding assay using human placental extracts [16]. Subsequently, band-shift and Southwestern assays were used to identify similar factors, termed MeCP1 and MeCP2, in rodent extracts [17–19]. Both factors are highly specific for symmetrically methylated CpG pairs. Hemimethylated DNA or 5mC in a non-CpG context were not substrates. When methylated templates were tested in F9 embryonal carcinoma cells, which contain low levels of MeCPs, they were found to be transcribed at high levels, whereas they were repressed in fibroblast cell lines which contained high MeCP activity. This implied that MeCPs could be transcriptional repressors. The MeCP2 protein was purified, sequenced and the corresponding cDNA isolated [19]. MeCP2 protein could be found throughout mouse chromosomes, but was especially concentrated in the heavily methylated centromeric heterochromatin. MeCP2 has a modular structure and contains a methylated DNA-binding domain (MBD) that is responsible for targeting it to mouse heterochromatin [20]. Screening of human and mouse databases indicated that there is a large family of proteins (MBD1–MBD4) that contain this motif [21,22]. With the exception of MBD3, each is capable of binding specifically to methylated DNA [19,20,23,24]. MeCP2, MBD1 and MBD2 were also found to repress transcription from methylated templates [21,24–27]. MBD4 has protein identity with bacterial DNA repair enzymes and is a glycosylase that can efficiently remove thymine or uracil from a mismatched CpG in vitro [28]. Although a truncated form of MBD2 has been reported to have a demethylase activity [29], other groups could not substantiate this observation [25,26]. It is more likely that MBD2 acts principally as a transcriptional repressor.

MBD-containing proteins repress transcription from methylated templates

How MeCP2 can repress transcription from methylated templates was initially investigated by the use of nuclear extracts that support transcription in vitro. Under conditions where both methylated and non-methylated templates are transcribed, it was demonstrated that the addition of recombinant MeCP2 inhibited the methylated template preferentially [27]. The MBD domain of MeCP2 was required and, through the use of GAL4 fusions and a GAL4-dependent template, it was demonstrated that there is a separate transcription repression domain. In these experiments only naked DNA was used, but it could be demonstrated that recombinant MeCP2 was incorporated preferentially into methylated chromatin, displacing histone H1 in the process [27]. Subsequently it has been shown that *Xenopus* MeCP2 can associate with MeCpGs in a mononucleosome [30]. Core histones can also be modified (by methylation, phosphorylation and acetylation) and this can have functional
consequences [31]. There is a general correlation between acetylation of the N-terminal tails of the core histones and a more open chromatin structure that facilitates gene expression (Figure 1). Many transcriptional co-activators have histone acetyltransferase activity and, conversely, there are transcriptional co-repressors that have histone deacetylase activity. MeCP2 was co-purified with a histone deacetylase complex, and inhibitors of histone deacetylase relieve MeCP2-mediated repression [24,25]. In HeLa cells, MBD2 is associated with a histone deacetylase in the MeCP1 repressor complex, and MBD3 has been found in a separate Mi2/NuRD deacetylase complex that posses a nucleosome-remodelling activity [32]. These observations provide a direct link between the methylated DNA and chromatin configuration. Future work will show how distinct these complexes are and determine whether there is cross-talk between different MeCpG-binding proteins within the biochemically distinct histone deacetylase complexes. It is possible that MeCP2/MBDs target these modifying complexes to distinct regions of the genome.

### DNA methyltransferases

Four mammalian cytosine DNA methyltransferases (Dnmt) have been identified [1], all of which contain a highly conserved C-terminal catalytic domain and variable N-terminal extensions (Figure 2). With the exception of Dnmt2, they specifically methylate cytosine in CpG. Dnmt1 is the best studied and has a 500-residue C-terminal catalytic domain and a 1100-residue N-terminal extension that has a regulatory role with respect to substrate specificity and in targeting the methyltransferase to different nuclear and

![Figure 2. The family of vertebrate DNA methyltransferase enzymes](caption)

The conserved regions in the methyltransferase catalytic domains are indicated with black boxes (I, IV, V, IX and X). In Dnmt1, NLS is a nuclear localization region, RF is the region of the protein targeting the enzyme to replicating foci and Zn2+ is a zinc-binding motif. KG, lysine/glycine repeats. Both Dnmt3a and Dnmt3b have a cysteine-rich region that may bind DNA. hm, hemi-methylated DNA; um, non-methylated DNA. + and − indicate the level of activity towards its substrate. Note that Dnmt1 has different activities in vitro and in vivo.
cellular sites (Figure 2). Hemi-methylated DNA is the preferred substrate for Dnmt1, but in *vitro* it possesses some activity towards non-methylated DNA (*de novo* activity). Since genomic DNA is in the hemi-methylated state after replication, Dnmt1 is regarded as the maintenance methylase, propagating pre-existing patterns of methylation. Recent additions to the vertebrate cytosine methyltransferase family are Dnmt3a and Dnmt3b [33], which were identified by screening mouse and human databases with sequences corresponding to the catalytic domain. These enzymes have *de novo* methyltransferase activity *in vitro* and *in vivo* and are required at different periods of mouse development [34].

**Patterns of DNA methylation during vertebrate development**

A big question is ‘How are DNA methyltransferases recruited to, or excluded from, particular regions of the genome during development?’ In mice there are complex changes in DNA methylation levels during embryogenesis and cell differentiation [35]. This involves a global demethylation during cleavage followed by a wave of *de novo* methylation in the growing embryo (Figure 3). These alterations in methylation appear to be intimately associated with the phenomenon of imprinting in mammals, whereby the expression pattern of a gene can be influenced by whether it is paternally or maternally inherited. The establishment of imprinted epigenetic marks takes place during gametogenesis, a process that is related to differential methylation of paternal and maternal alleles of some genes that are essential for embryo growth [9]. Initially the germ line cells are undermethylated. In fully mature gametes, sperm DNA has higher levels of methylation than oocyte DNA. Immunohistochemical experiments with a monoclonal antibody against 5mC, and bisulphate sequencing of maternal and paternal alleles of imprinted genes, have demonstrated that the male pronucleus is rapidly demethylated after fertilization by an activity present in the oocyte cytoplasm [36]. The maternal genome remains more methylated than the paternal DNA up to the 8-cell stage, when they become equivalent. The level of methylation decreases to 15% in the early blastocysts and returns to higher levels during implantation. Meanwhile the maternal oocyte form of Dnmt1 is excluded from the nucleus, and its relocation there is coincident with the remethylation of the mouse genome [37]. It has been argued that there is a highly regulated process of reprogramming in the developing mouse embryo that involves erasing of epigenetic modifications present in the zygote followed by subsequent *de novo* methylation necessary for resetting the developmental patterns of gene expression in differentiating cell lineages [35].

In contrast to mammals, neither *X. laevis* nor zebrafish have imprinted genes or identifiable sex chromosomes. Compared with the dynamic changes of 5mC levels in mammals, the early embryonic patterns of methylation in
these animals are set up in a different manner. Dnmt1 is asymmetrically localized in the oocyte and egg cytoplasm of toads and fish [3,4]. The animal pole of the egg, which gives rise to mesoderm and ectodermal tissues, contains high levels of Dnmt1 RNA and protein, while the vegetal hemisphere, that later differentiates into ectoderm, contains considerably lower levels of the enzyme. Such localization perhaps has a functional relevance, since up to 3-fold differences in the amount of $^5$mC can be detected around midblastula in DNA derived from *Xenopus* animal and vegetal cells. As in the mouse, DNA

Figure 3. Relative levels of methylation and localization of Dnmt1 enzyme during mouse, *Xenopus* and zebrafish development
(a) Compared with the dynamic alterations in $^5$mC content during early mouse development, methylation changes are much less pronounced in amphibia and absent in fish. The methylation patterns of mammalian somatic cell lineages are established during gastrulation. In lower vertebrates, early embryonic cells are partially fated to differentiate before gastrulation. (b) In *X. laevis* and zebrafish, Dnmt1 (blue) is asymmetrically localized in the oocyte cytoplasm and blastula stage embryos. Consequently, DNA (black) of *Xenopus* animal pole cells is more methylated than that of the vegetal hemisphere cells (grey). (c) In mouse, sperm DNA (black) is actively demethylated in the oocyte cytoplasm after fertilization, while the female pronucleus remains unchanged. The $^5$mC content of the zygote decreases progressively and is very low in the blastocyst. Somatic methylation patterns are established in the gastrula. Dnmt1 (blue) migrates from cytoplasm to the nucleus at the 8-cell stage, but is excluded later during the demethylation of blastocyst cells.
methylation decreases during the cleavages, but overall changes in 5mC content are not so dramatic [4]. Low levels of methylation coincide with the initiation of zygotic transcription at midblastula and, in contrast to mouse, remain relatively low during gastrulation. At the other extreme, there are no detectable changes in methylation levels during zebrafish development [38]. Most likely, amphibia and fish do not need a genome-wide wave of demethylation and remethylation to reset the initial methylation patterns, since they might be established already in sub-populations of cells before gastrulation. This observation is supported by the finding that Xenopus blastomeres are already fated to differentiate to different lineages in 8-cell and 16-cell embryos [39], while embryonic stem cells derived from mouse blastocysts are uniformly methylated and totipotent. Methylation patterns of differentiated somatic cells in lower vertebrates and mammals are also probably achieved by tissue-specific expression of de novo and maintenance methyltransferase enzymes.

**DNA methylation is essential for vertebrate development**

In all species studied to date (mouse, zebrafish and Xenopus), Dnmt1 is expressed as an abundant oocyte form and later shows varying tissue-specific levels [3,4,37]. Loss of Dnmt1 or inhibition by 5-azacytidine (an analogue of cytidine) are injurious to embryo development in all three species [2–4]. Due to the phenotypic complexity and early lethality of Dnmt1−/− mutants in mice, the question as to whether methylation is essential for regulation of gene expression at the onset of gastrulation in the developing embryo has been difficult to answer [2,8]. Studies have shown that loss of methylation affects mesoderm formation during gastrulation [3,4] of Dnmt1-deficient Xenopus and zebrafish embryos grown in the presence of 5-azacytidine. Dnmt1−/− mice exhibit development delay and asynchrony, which may also be indicative of gastrulation abnormalities [2,8]. In Xenopus DNA, methylation is involved in the maintenance of genome-wide transcriptional silencing that occurs between fertilization and midblastula transition [4]. It is also important in the regulation of appropriate expression patterns of developmentally essential genes, among which are transcription factors (Xbra and Otx2) and signalling molecules (Cerberus). Transient depletion of maternally expressed Dnmt1 by anti-sense RNA in Xenopus leads to premature transcription of these genes at least two cell cycles earlier than normal. Loss of the initial methylation patterns and inappropriate gene activation lead to changes in the developmental competence of animal pole cells. Similar defects were observed in zebrafish embryos grown in the presence of 5-azacytidine. One generalization from the analysis of Dnmt1 depletion or inhibition in all three species is that it is not essential for the survival of embryonic cells during early cleavage stages. The effect of disrupting pre-existing methylation patterns only becomes apparent during and after gastrulation. The common features of methylation-depleted embryos in all three species include a failure to organize
neural and muscle tissues and a high degree of apoptosis (Figure 4). The accumulated evidence argues that methylation-mediated transcriptional repression is important for the normal development of all vertebrates.

**Future perspectives**

Nuclear transplantation experiments in mammals and amphibia have shown that, in rare cases, somatic nuclei are able to support normal embryonic development [40,41]. Presumably, the somatic patterns of methylation have to be erased to allow the full potential of the transferred nucleus to be realized. This suggests that epigenetic mechanisms may have a role in potentiating embryonic cells to differentiate by reinforcing, through DNA methylation and the modification of chromatin structure, particular expression states. The

![Figure 4. Loss or inhibition of Dnmt1 results in severe developmental defects](image)

recent findings that methyl-binding proteins are involved in chromatin remodelling by targeting of large protein complexes that involve Sin3A/histone deacetylase complex or NuRD to methylated DNA will ultimately answer the question as to how the expression of specific genes is regulated by methylation [31].

From a developmental point of view it is essential to know how DNA methylation patterns and the subsequent gene repression are established. It is tempting to speculate that DNA methylation is not simply a consequence of gene inactivation, but instead depends upon active targeting of de novo DNA methyltransferases to specific sequences. It is also possible that there is developmentally controlled feedback between the maintenance and de novo methyltransferases in the establishment of differential tissue-specific methylation patterns.

Another question is whether DNA methylation in vertebrates is involved in the regulation of the similar sets of genes (apart from the imprinted loci in mammals) during embryogenesis. Unfortunately this question has not been answered in mice because of the complexity of Dnmt1 phenotypes, which are greatly dominated by the negative effects of mis-expression of imprinted genes and X-chromosome inactivation. Since most of the components of the methylation-mediated gene repression machinery are well conserved between Xenopus, mouse and human, this raises the possibility that methylation plays a very similar function in the development of all vertebrates.

Summary

- MeCpGs act as ligands for nuclear factors (repressors) that are components of chromatin modification and remodelling activities.
- The DNA-methylation-mediated repression system (Dnmt1s, MeCPs and MBDs) is highly conserved in vertebrates.
- DNA methylation is essential for normal vertebrate development.
- It is possible (but remains unproven) that the role of DNA methylation in regulating development is highly conserved in vertebrates.
- In mammals, DNA methylation has an additional role in regulating the expression of imprinted genes and in controlling X-inactivation.

References


