CHAPTER 2

Mechanisms of DNA Methylation, Methyl-CpG Recognition, and Demethylation in Mammals

Xiaodong Cheng, Hideharu Hashimoto, John R. Horton, and Xing Zhang
Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA

INTRODUCTION

The control of transcription initiation in mammalian cells can be very broadly divided into three categories: intrinsic promoter strength and availability of core transcription machinery [1-3], the actions of promoter- or regulon-specific transcription factors (positive and negative) [4-6], and the control of DNA accessibility by altering chromatin structure [6-8]. This latter category, including posttranslational modifications to histones and postreplication modification of DNA, is the focus of recent extensive studies. Nucleosomes are the fundamental building blocks of eukaryotic chromatin, and consist of ~146 base pairs of DNA wrapped twice around a histone octamer [9]. A variety of protein-modifying enzymes (including methyltransferases, MTases) is responsible for histone modification, primarily at their flexible N-termini [10-12]. Here, we summarize the most recent structural and biochemical advances in the study of mammalian DNA MTases and their associated protein factor(s), and will touch on the functional links between histone modification and that of DNA.

In mammals and other vertebrates, DNA methylation occurs at the C5 position of cytosine (5mC), mostly within CpG dinucleotides (Fig. 2.1A), with the Dnmt enzymes using a conserved mechanism [13] that has been studied best in the bacterial 5mC MTase M.Hhal [14-18]. Briefly, this mechanism involves MTase binding to the DNA, eversion of the target nucleotide so that it projects out of the double helix ("base flipping"), covalent attack of a conserved Cys nucleophile on cytosine C6, transfer of the methyl group from S-adenosyl-l-methionine (AdoMet) to the activated cytosine C5, and the various release steps. This methylation, together with histone modifications, plays an important role in modulating chromatin structure, thus controlling gene expression and many other chromatin-dependent processes [19]. The resulting epigenetic effects maintain the various patterns of gene expression in different cell types [20]. Epigenetic processes include genomic imprinting [21], gene silencing [22,23], X chromosome inactivation [24], reprogramming in transferred nuclei [25,26], and some elements of carcinogenesis [27].
FIGURE 2.2
Schematic representation of Dnmt1 and Dnmt3. (A) Roman numerals refer to conserved motifs of DNA MTases [156]; motif IV includes the Cys nucleophile that forms a transient covalent bond to C6 of the target cytosine. (B) Maintenance vs. de novo methylation. The rectangular segments are substrate sequences (usually CpG), and the small ball shapes represent methyl groups on the cytosines. Following replication or repair, the duplex is methylated on one strand only. (C) The first domain structure of Dnmt1 (residues 350–599; PDB 3EPZ) [89] contains targeting sequence association with replication foci [88]. (Please refer to color plate section)

Both Dnmt3a and Dnmt3L C-terminal domains have the characteristic fold of Class I AdoMet-dependent MTases [56]. However, the methylation reaction product S-adenosyl-L-homocysteine (AdoHcy) was found only in Dnmt3a and not in Dnmt3L. This is consistent with Dnmt3a being the catalytic component of the complex, while Dnmt3L is inactive and unable to bind AdoMet [52,53]. The overall Dnmt3a/Dnmt3L C-terminal complex is ~16 nm long, which is greater than the diameter of a 11-nm core nucleosome (Fig. 2.3A). This complex contains two monomers of Dnmt3a and two of Dnmt3L, forming a tetramer with two 3L-3a interfaces and one 3a-3a interface (3L-3a-3a-3L). Substituting key non-catalytic residues at the Dnmt3a-3L or Dnmt3a-3a interfaces eliminates enzymatic activity, indicating that both interfaces are essential for catalysis [55].

DIMERIC Dnmt3a SUGGESTS DE NOVO DNA METHYLATION DEPENDS ON CpG SPACING
Among known active DNA MTases, Dnmt3a and Dnmt3b have the smallest DNA binding domain (though it is absent altogether in Dnmt3L). However, dimerization via the 3a-3a interface brings two active sites together and effectively doubles the DNA-binding surface. Superimposing the Dnmt3a structure, onto that of M.HhaI complexed with a short oligonucleotide [14], yielded a model such that the two active sites are located in the DNA major groove and dimeric Dnmt3a could methylate two CpGs
FIGURE 2.3
A model of interactions between Dnmt3a-3L tetramer and a nucleosome. (A) A nucleosome is shown, docked to a Dnmt3L-3a-3s-3L tetramer (3a-C in green; 3L full length in gray). The position of a peptide derived from the sequence of the histone H3 amino terminus (purple) is shown, taken from a co-crystal structure with this peptide bound to Dnmt3L [71]. Wrapping the tetramer around the nucleosome, the two Dnmt3L molecules could bind both histone tails from one nucleosome. The amino-proximal portion of Dnmt3a is labeled as N (for N-terminal domain), PWWP domain, and CXXC domain. By analogy to Dnmt3L, the CXXC domain of Dnmt3a might interact with histones tails from neighboring nucleosomes. (B) The Dnmt3a dimer could in theory methylate two CpGs separated by one helical turn in one binding event. (C) Structure of Dnmt3L with a bound histone H3 N-terminal tail (orange) [71]. (D) The PWWP domain structure of mouse Dnmt3b, rich in basic residues [42]. (Please refer to color plate section)

A periodicity in the activity of Dnmt3a on long DNA substrates revealed a correlation of methylated CpG sites at distances of 8–10 base pairs, and the structural model of oligomeric Dnmt3a docked to DNA may explain this pattern [55]. Similar periodicity is observed for the frequency of CpG sites in the differentially-methylated regions of 12 maternally-imprinted mouse genes [55]. These results suggest a basis for the recognition and methylation of differentially-methylated regions in imprinted genes, involving detection of both CpG spacing and nucleosome modification (see next section). Zha et al. (2009) analyzed the methylation status of a large number of CpG sites (total of 580,427) of chromosome 21 and found that CpG DNA methylation patterns are correlated with the CpG periodicity of nine base pairs [57]. More recently, an 8–10 bp pair periodicity has also been evident for non-CpG methylation in embryonic stem cells [58]. Non-CG methylation disappeared upon induced differentiation of the embryonic stem cells, and was restored in induced pluripotent stem cells. Similarly, a 10-bp correlation of non-CpG DNA methylation by Arabidopsis thaliana DRM2 (which is re to mammalian Dnmt3a) has been observed [59].
**Dnmt3L CONNECTS UNMETHYLATED HISTONE H3 LYSINE 4 TO DE NOVO DNA METHYLATION**

DNA methylation and histone modifications are intricately connected with each other [60–62]. In fact, genome-scale DNA methylation profiles suggest that DNA methylation is better correlated with histone methylation patterns than with the underlying genome sequence context [62]. Specifically, DNA methylation is correlated with the absence of H3K4 methylation and the presence of H3K9 methylation. Methylation of histone H3 lysine 4 (H3K4) [63] has been suggested to protect gene promoters from de novo DNA methylation in somatic cells [64,65]. There have been reports of an inverse relationship between H3K4 methylation and allele-specific DNA methylation at differentially methylated regions [57,62,66–69]. More recently, AOIF1 (amine-oxidase flavin-containing domain 1), a homolog of histone H3 lysine 4 demethylase (LSD1), has been shown to be required for de novo DNA methylation of imprinted genes in oocytes [70], suggesting that demethylation of H3K4 is critical for establishing the DNA methylation imprints during oogenesis.

The mammalian de novo DNA methylation Dnmt3L-Dnmt3a machinery could translate patterns of H3K4 methylation, which are not known to be themselves preserved during chromosome replication, into heritable patterns of DNA methylation that mediate transcriptional silencing of the affected sequences [71]. Dnmt3a is fully active on nucleosomal DNA in vitro [72]. Dnmt3a2 is a shorter isoform of Dnmt3a, predominant in embryonic stem cells and embryonal carcinoma cells and detectable in testis, ovary, thymus, and spleen, that is also required for genomic imprinting [73]. Dnmt3a2 and Dnmt3b, along with the four core histones, were identified as the main in vivo interaction partners of epitope-tagged Dnmt3L [71]. Peptide interaction assays showed that Dnmt3L specifically interacts with the extreme amino terminus of histone H3; this interaction was strongly inhibited by H3K4 methylation, but was insensitive to modifications at other positions [71]. Co-crystallization of Dnmt3L with the amino tail of H3 showed this tail bound to the Cys-rich 3-Zn binding domain of Dnmt3L (Fig. 2.3C), and substitution of key residues in the binding site eliminated the H3-Dnmt3L interaction. These data suggest that Dnmt3L is a probe of H3K4 methylation, and if the methylation is absent then Dnmt3L induces de novo DNA methylation by docking activated Dnmt3a2 to the nucleosome.

Mouse ES cells that lack the H3 lysine 9 (H3K9) MTases Suv39h1 and Suv39h2 show slight demethylation of satellite DNA [74]. G9a and GLP (G9a-like protein) – two related euchromatin-associated H3K9 methyltransferases [75] – have been implicated in DNA methylation at various loci, including imprinting center [76,77], retrotransposons and satellite repeats [78], a G9a/GLP target promoter [79], and a set of embryonic genes [80]. In filamentous fungi *Neurospora*, the H3K9 methyltransferase DIM-5 is required for DNA methylation [81–84], whereas in *Arabidopsis* the H3K9 methyltransferase KRYPTONITE is required for DNA methylation [85]. This suggests an evolutionarily-conserved silencing pathway in which H3K9 methylation correlates with DNA methylation. However, how H3K9 methylation contributes to DNA methylation is not clear, particularly in mammalian cells. G9a interacts directly with Dnmt1 during replication [86]. In addition, the G9a ankyrin repeat domain has been suggested to interact with Dnmt3a [80,87], a possible way for G9a to induce de novo DNA methylation [78].

**A STRUCTURAL FRAGMENT OF Dnmt1**

At the time of this writing (August, 2009), one domain structure is available (Fig. 2.2C) for part of the large 183 kDa Dnmt1 protein. The region (residues 350–599 of human Dnmt1) was initially identified as a novel targeting sequence association with replication factors [88]. This sequence has the properties expected of a targeting sequence in that it is not required for enzymatic activity, prevents proper targeting when deleted, and, when fused to
β-galactosidase, causes the fusion protein to associate with replication foci in a cell cycle-dependent manner. The domain structure, solved by the Structural Genomics Consortium at Toronto (PDB 3EPZ) [89], adopts a mainly β structure in the N-terminal half and a hel bundle in the C-terminal half (Fig. 2.2C).

Dnmt1 itself is subject to posttranslational modifications, including phosphorylation (Ser515 in mouse Dnmt1) [90,91] and methylation (Lys142 in human Dnmt1) [92]. Methylation of Dnmt1 at Lys142, mediated by Set7 (a protein lysine methyltransferase), resulted in its decreased stability [92]. Reciprocally, enhanced Dnmt1 methylation in the background of total deletion of LSD1 (a protein lysine demethylase) correlates with reduced Dnmt1 stability in vivo and progressive loss of DNA methylation [93]. Furthermore, it was hypothesized that polymers present on PARP-1 (PARylated poly(ADP-ribose) polymers) interact noncovalently with Dnmt1, preventing Dnmt1 enzymatic activity. In the absence of poly(ADP-ribose)ation of PARP-1, Dnmt1 is free to methylate DNA; if, in contrast, high levels of PARylated PARP-1 persist, Dnmt1 will be stably inhibited, preventing DNA methylation [94].

THE SRA DOMAIN OF UHRF1 FLIPS 5-METHYLCYTOSINE OUT OF THE DNA HELIX

An accessory protein UHRF1 (ubiquitin-like, containing PHD and RING finger domains targets Dnmt1 to hemimethylated replication forks (and presumably repair sites) [95–97]. The murine ortholog of this protein is also known as NP95 (nuclear protein of 95 kDa) [98–100]; the human ortholog is called ICBP90 (inverted CCAAT binding protein of 90 kDa) [101].

The crystal structure of the SET and RING associated (SRA) domain of UHRF1 in complex with DNA containing a hemimethylated CpG site was recently determined [102–104]. It reveals that the SRA domain flips the 5-methylcytosine (5mC) completely out of the DNA helix (Fig. 2.4A,B) and is positioned in a binding pocket with planar stacking contacts, Watson-Crick polar hydrogen bonds and van der Waals interactions specific for 5mC that distinguishes 5mC from cytosine. The structure also suggests an explanation for the preference for hemimethylated sites. In the major groove side, a backbone carbonyl oxygen is close to the C5 ring carbon of the unmethylated cytosine, forming a C=O···H-C hydrogen bond. The addition of a methyl group to C5 of the unmethylated cytosine would cause steric clash between the methyl group and SRA.

BASE FLIPPING MECHANISM

Base flipping is a conserved mechanism that is widely used by nucleotide modifying enzymes, including DNA MTases [13,14], DNA repair enzymes [105–108], and RNA modification enzymes [109]. This mechanism, first discovered in the bacterial 5mC MT M.Hhal [14], involves enzyme binding to the DNA and eversion of the target nucleotidethat it projects out of the double helix and into the active-site pocket. The SRA domain first-discovered non-enzymatic sequence-specific DNA binding protein domain that use base flipping mechanism in its interaction with DNA.

There is no apparent sequence or structural similarity between the SRA and the DNA M domain (or of DNA repair enzymes). However, the phosphodiester backbone pinching [110] due to extensive protein-phosphate contacts surrounding the flipped nucleotide, the use of two loops to approach DNA from the major and minor grooves simultaneously, and the binding of the flipped base in a concave pocket are analogous to the DNA MTases (Fig. 2.5A,B) [111]. Furthermore, enzymes use base flipping to gain access to a DNA base to perform chemistry on it, but the SRA domain probably uses base flipping to increase protein-DNA interface and to prevent the SRA domain from linear diffusion away from
FIGURE 2.4

UHRF1 – a multi-domain protein. (A) Schematic representation of UHRF1 and its homolog UHRF2. (B) Structure of SRA-DNA complex. The 5mC flips out and is bound in a cage-like pocket. (C) Five domain structures are currently available. (Please refer to color plate section)

site on the DNA. This may be particularly important for the SRA domain, as its recognition sequence is only two base pairs. The surface area buried at the SRA-DNA interface is increased approximately 70% from what is buried at the MBD1-DNA interface [104] that does not involve base flipping [112,113] (Fig. 2.5C).

The 5mC base flipping by the SRA domain might also provide a more general mechanism to distinguish the methylated parental strand from the unmethylated daughter strand, an ability particularly important for mismatch repair if an error occurs during DNA replication. Supporting this hypothesis, the expression of ICBP90 (the human ortholog of UHRF1) is deregulated in cancer cells [114], and mouse UHRF1-null cells are more sensitive to DNA damaging agents and DNA replication arrest [99]. We therefore suggest that the SRA-DNA interaction (through recognition and flipping of the 5mC) serves as an anchor to keep UHRF1 at hemimethylated CpG site where it recruits Dnmt1 for maintenance methylation, and perhaps other proteins such as DNA repair enzymes for mismatch repair.

UHRF1-HISTONE INTERACTIONS

Besides the SRA domain, the Structural Genomics Consortium at Toronto has solved three additional domain structures by X-ray crystallography (Fig. 2.4C): the N-terminal ubiquitin-like domain (PDB 2FAZ [115]) and the tandem tudor domain with and without bound histone H3K9me3 (PDB 3DB3 and 3DB4 [116] for human UHRF1), and the C-terminal RING domain (PDB 1Z6U [117] of human UHRF2), whereas the RIKEN Structural Genomics/Proteomics Initiative at Japan solved an NMR structure for the PHD domain of human UHRF2 (PDB 2E6S [118]). Individual domain functions have been
suggested for the PHD that may be involved in histone H3 tail binding [98,101], and for the RING domain that may confer E3 ubiquitin ligase activity on histones [98]. The new structure of the tandem tudor domain bound with histone H3 tail contains three structurally distinct H3 residues, H3R8-H3K9me3-H3S10 (PDB 3DB3).

**REPLICATION-COUPLED CROSSTALK BETWEEN DNA METHYLATION AND HISTONE MODIFICATIONS**

The fact that UHRF1 contains modules, within the same polypeptide, recognizing both DNA silencing marks (via the SRA) and histone silencing marks (via the tudor and/or PHD domain) suggests that it may be a key component to couple the preservation of histone modification through the cell cycle with maintenance DNA methylation (Fig. 2.6). We hypothesize that UHRF1 brings the two components (histones and DNA) carrying appropriate markers (the tails of H3 and hemimethylated CpG sites) ready to be assembled into a nucleosome after replication. In this context the missing key experiment is whether these domains act independently or in a cooperative fashion. For example, does binding of UHRF1 SR domain to hemi-methylated DNA improve the binding of UHRF1 tudor and/or PHD domains to histone tail? Furthermore, the E3 ubiquitin ligase activity [98], residing in the C-terminal RING domain, may add an additional level of modification, such as H2.
ubiquitylation that is enriched with inactive genes [119]. In a separate note, a recent study suggested Polycomb proteins remain bound to chromatin and DNA during DNA replication in vitro [120]. Retention of Polycomb proteins through DNA replication may contribute to maintenance of transcriptional silencing through cell division.

DNA DEMETHYLATION VIA HYDROXYLATION?

5-hydroxymethylcytosine (5hmC) has long been noted in bacterial phage DNA [121−125], and its presence in mammalian cells [126] was believed to be a by-product of oxidative DNA damage [127]. Recently, using isolated relatively homogeneous populations of Purkinje and granule neuronal nuclei of adult mouse brains, Kriaucionis and Heintz found that significant fractions (~40%) of cytosine nucleotides correspond to 5hmCs, the amount of which inversely correlates with 5mC and nuclear heterochromatin in neurons [128]. Even more fascinating, a conserved mammalian-specific family of TET (ten-eleven translocation) proteins that converts 5mC to 5hmC (Fig. 2.1B) was identified [129]. One of these proteins, TET1, is fused to the MLL protein in a subset of acute myeloid leukemia patients. Overproduction of TET1 in human cells led to the appearance of 5hmC. A concomitant reduction in DNA 5mC indicated that 5hmC is an oxidation product of 5mC. 5hmC was detected in ES cells and this percentage decreased with RNAi knockdown of TET1. The surprising finding of a 5mC oxidation pathway raises numerous questions, such as whether oxidation of 5mC is an important epigenetic modification, either as an end product or as an intermediate in active DNA demethylation, as supported by the presence of 5hmC DNA excision repair glycosylase [130]. It is intriguing to note that bacterial 5mC MTase M.Hhal can promote the reverse reaction in vitro - the removal of formaldehyde from 5hmC to yield the unmodified cytosine [131] (Fig. 2.1C). New lines of research will likely be catalyzed by the presence of 5hmC in mammalian DNA.
DNA DEMETHYLATION VIA GLYCOSYLATION?

A recent study reported that MBD4 [132], a protein containing an N-terminal methyl-CpG-binding domain (MBD) and a C-terminal glycosylase domain [133], is phosphorylated via protein kinase C (PKC) by parathyroid hormone stimulation [134]. Phosphorylated MBD4 promotes incision of methylated DNA through glycosylase activity, and a base-excision repair process seems to complete DNA demethylation in the MBD4-bound promoter. Such parathyroid-hormone-induced MBD4 phosphorylation and subsequent DNA demethylative and transcriptional derepression are impaired in Mbd4(-/-) mice.

Dnmt2, AN ENIGMATIC DNA METHYLTRANSFERASE HOMOLOG

No review on mammalian DNA methylation will be complete without mention of Dnmt2, a small protein of 391 residues in human, initially identified based on its conservation of sequence motifs to known DNA cytosine-C5 methyltransferases [135,136]. Targeted deletion of Dnmt2 gene indicated Dnmt2 is not required for global de novo or maintenance methylation of DNA in embryonic stem cells [137]. Baculovirus-expressed Dnmt2 protein failed to methylate DNA in vitro [137], whereas bacteria-expressed Dnmt2 showed no detectable activity in one study [138] or residual activity towards DNA in another study [139]. Structure of human Dnmt2 is closely related in overall structure to M.Hhal [138], a bacterial DNA methyltransferase. In 2006, Goll et al. revealed that human Dnmt2 methylates cytosine 38 in the anticodon loop of aspartic acid transfer RNA (tRNA\(^{\text{ASP}}\)) [140]. The function of tRNA\(^{\text{ASP}}\) methylation by Dnmt2 is highly conserved, as human Dnmt2 protein restored methylation in vitro to tRNA\(^{\text{ASP}}\) from Dnmt2-deficient strains of mouse, Arabidopsis thaliana, and Drosophila melanogaster [140]. Not surprisingly, human Dnmt2 methylates RNA cytosine-C5 using a D1 methyltransferase-like catalytic mechanism [141], because of high sequence and structure conservation as well as identical chemistry between DNA cytosine and RNA cytosine.

In parallel, study in Drosophila revealed that depletion of dDnmt2 had no detectable effect on embryonic development, whereas overexpression of dDnmt2 resulted in significant genomic DNA hypermethylation at CpT and CpA dinucleotides [142]. A weak but significant activity for Dnmt2 was detected in a non-CpG dinucleotide context in flies overexpressing mouse Dnmt2 [143]. More recently, Phalke et al. showed that dDnmt2 controls retrotransposon silencing in Drosophila somatic cells [144]. Loss of dDnmt2 eliminates histone H4 lysine 20 (H4K20) trimethylation (mediated by SuV4-20) at retrotransposons and impairs maintenance of retrotransposon silencing. The new study uncovers a previously unappreciated role of dDnmt2 in DNA methylation in retrotransposon silencing and telomere integrity in Drosophila [144] and will help to resolve the Dnmt2 enigma [145] as whether two substrates are better than one [146].

CONCLUSION AND PERSPECTIVES

The experimental characterizations of Dnmt3s and their associated protein factors are providing a rapidly and convergent picture of the kinetic mechanisms (activities of oligomers [147]), binding partners (UHRF1-Dnmt1 [95,96] and Dnmt3L-Dnmt3a [55]) chromatin recognition [62] (histone binding such as H3K4me0 [71] and H3K9me3 [74]) RNA-directed DNA methylation [148,149], methylation-dependent regulation (by protolysine methyltransferases G9a/GLP [78-80] and by protolysine demethylase LSD1 [93 and the discovery of 5hmC in mammalian genome [128,129]. However, understanding basis for establishing, maintaining, and disturbing DNA methylation patterns will require a much better understanding of the union between structure and function in the Dnmts and their associated protein factors than we currently possess. Without understanding the interactions and spatial relationships between their modular domains, or whether inter-domain interactions contribute to target specificity, it is not possible to construct a temporal sequence of events or causal relationships in gene silencing.
ACKNOWLEDGEMENTS

We thank most warmly our colleagues and coworkers whose hard work was responsible for much of the DNA and histone methylation work cited in this review. Work in the authors' laboratory is supported in part by grants (US National Institutes of Health GM049245, GM068680, and DK082678) and the Georgia Research Alliance. This chapter is an updated version of our previous articles “Mammalian DNA methyltransferases: A structural perspective” (originally published in Structure 16, March 2008, pp. 341–350; with permission from Elsevier) and “UHRF1, a modular multi-domain protein, regulates replication-coupled crosstalk between DNA methylation and histone modifications” (originally published in Epigenetics 4, pp. 8–14; January 2009, with permission from Landes Bioscience).

References


Mechanisms of DNA Methylation, Methyl-CpG Recognition, and Demethylation in Mammals


134. Kim MS, Kondo T, Takada I, Youn MY, Yamamoto Y, Takahashi S, et al. DNA demethylation in hormone
135. Vilain A, Apiou F, Dutrillaux B, Malfoy B. Assignment of candidate DNA methyltransferase gene (DNMT
136. Yoder IA, Bestor TH. A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. J
137. Okano M, Xie S, Li E. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in
DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Res
139. Hermann A, Schmitt S, Jeitsch A. The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase
140. Goll MG, Kirpekar F, Maggett KA, Yoder IA, Hsieh CL, Zhang X, et al. Methylation of tRNAAsp by the DN
141. Jurkowska RZ, Schaefer M, Assmann B, Li F, Lyko F. Comparative analysis of DNA methylation patt
143. Jeitsch A, Nellen W, Lyko F. Two substrates are better than one: dual specificities for Dnmt2 methyltransf
filaments by mammalian DNA methyltransferase Dnmt3a in complex with regulator Dnmt3L. Nucleic A
chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. Nat Ge
individual transposons is linked to de novo DNA methylation in mice. Mol Cell 2008;31:785–99.
147. Kress C, Thomasin H, Grange T. Active cytosine demethylation triggered by a nuclear receptor involves 1
148. Kaipapand M. Characterization of DNA demethylation in normal and cancerous cell lines and the regul
152. Rai K, Huggins JJ, James SR, Karpf AR, Jones DA, Cairns BR. DNA demethylation in zebrafish involves th
Mechanisms of Histone Modifications

Zdenko Herceg\textsuperscript{1} and Rabih Murr\textsuperscript{1,2}
\textsuperscript{1}Epigenetics Group, International Agency for Research on Cancer (IARC), Lyon, France
\textsuperscript{2}Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland

INTRODUCTION

The term "epigenetic" was first introduced by Conrad Waddington in 1942 to describe "The interactions of genes with their environment that bring the phenotype into being". Currently, it includes all features such as chromatin and DNA modifications that are heritable and stable over rounds of cell division, but do not alter the nucleotide sequence within the underlying DNA \cite{1}. Over the years, a wide variety of products and events have been lumped into epigenetics. These include paramutation, bookmarking, imprinting, gene silencing, X chromosome inactivation, position effect variegation, reprogramming, transvection, infection agents like prions, maternal conditioning, RNA interference, non-coding RNA, small RNAs, DNA methylation and chromatin modifications. In this chapter, we will focus on epigenetic mechanisms involving histone modifications and recent development establishing a link between chromatin modifications (with an emphasis on acetylation and methylation) and cellular processes such as transcription and DNA repair.

HISTONE MODIFICATIONS

In all eukaryotes, chromatin is a highly condensed structure that forms the scaffold of fundamental nuclear processes such as transcription, replication and DNA repair \cite{2}. Chromatin exists in at least two conceptually distinct functional forms: a condensed form during mitosis and meiosis that generally lacks DNA regulatory activity, called heterochromatin; and a looser decondensed form, which provides the environment for DNA regulatory processes, called euchromatin. Nucleosomes are the building blocks of chromatin and they represent two turns of genomic DNA (147 base pairs) wrapped around an octamer of two subunits of each of the core histones H2A, H2B, H3, and H4. The amino-terminal portion of the core histone proteins contains a flexible and highly basic tail region, which is conserved across various species and is subject to various post-translational modifications (Fig. 3.1). The structure of chromatin fulfills essential functions, not only by condensing and protecting DNA, but also in preserving genetic information and controlling gene expression \cite{3}. However, given its compacted structure, chromatin hinders several important cellular processes including, transcription, replication, and the detection/repair of DNA breaks \cite{4,5}. Therefore, chromatin