Chromatin structure and epigenetics

A.S. Quina, M. Buschbeck, L. Di Croce

Center for Genomic Regulation, Passeig Maritim 37-49, 08003 Barcelona, Spain
ICREA and Center for Genomic Regulation, Passeig Maritim 37-49, 08003 Barcelona, Spain

1. Introduction

From the functional point of view, various chromatin structures are commonly divided into euchromatin and heterochromatin. Euchromatin corresponds, in general, to genome regions that possess actively transcribed genes (or potentially active ones), and that are decondensed during interphase. The regulatory sequences in these regions are accessible to nucleases and have, characteristically, unmethylated CpG islands and the core histones H3 and H4 are hyperacetylated on their N-terminal lysine residues. In general, euchromatic domains replicate early in S phase [1].

By contrast, heterochromatin refers to the transcriptionally inactive and highly condensed regions of the genome. Within heterochromatin, the DNA renders itself inaccessible to nucleases, it is usually methylated in the dinucleotide CpG and histones are markedly hypoacetylated. Depending on whether heterochromatin is established in every cell type or limited to (a) particular lineage(s) (and/or actively formed in certain cell types), heterochromatic domains may be further divided into constitutive or facultative heterochromatin, respectively [1,2].

Constitutive heterochromatin is in general gene-poor and forms mainly on repetitive sequences, such as satellite DNA.
centromeric and pericentromeric repeats. These regions replicate late in S-phase. Histones H3 and H4 are typically tri-methylated on lysine residues K9 and K20, respectively. In human and mouse cells, the chromatin in pericentromeric regions is enriched in histone methyltransferases and HP1 (heterochromatin protein 1) proteins, which bind specifically to tri-methylated H3-K9 [3,4].

Regions and DNA sequences that are subject to a developmentally regulated transcriptional silencing constitute the facultative heterochromatin. Large-scale heterochromatinization of the genome is frequently observed in terminally differentiated cells. Examples of sequences subject to heterochromatinization include genes silenced during cell differentiation, the inactive chromosome X in female mammal cells (cytologically observed as a dense nuclear structure referred to as the Barr body), and the inactive alleles of genes with monoaletic expression (such as those subject to imprinting). Lysine 27 of histone H3 is typically methylated within these regions, a mark set and recognized by proteins of the Polycomb group (PcG). Additionally, H3-K9 and H4-K20 are in general di- and mono-methylated, respectively.

2. The hypothesis of “histone code”

Post-translational modifications in nucleosomal histones, either at the local or the genomic level, seem to be related with, and even predict, transcriptional states. This is the hypothesis of the so-called “histone code” [5], whereby “distinct histone amino-terminal modifications can generate synergistic or antagonistic interaction affinities for chromatin-associated proteins” (in [6]), which in turn regulate access to the underlying DNA. It is proposed that the combinatorial use of histone modifications may work as a marking system that is recognized/read by regulatory proteins. On the other hand, these marks may directly affect chromatin structure. Core histone acetylation [7] for instance helps to relax the higher-order chromatin structure and transiently allows DNA accessibility. This code may then be passed on from one cell generation to the next as an epigenetic “memory” of transcriptional programs, the epigenome [8].

More than 30 residues within each of the four octameric histone partners comprising a nucleosome are described as sites that can be modified in the context of chromatin. These covalent modifications include acetylation, methylation, phosphorylation and ubiquitination. Particularly on their N-terminal tails, individual histones may acquire a series of modification marks in close proximity to each other [9]. The way these modifications interact with each other and the way they correlate with the transcriptional states are currently object of significant research efforts. The emerging view is that enzymes that catalyze histone modifications and proteins that are able to read the “code” act in a concerted and highly interdependent fashion. These “translators” or effector proteins bind to specific modifications and recruit other regulatory or remodelling factors which, in turn, will help to nucleate or maintain a particular chromatin structure, thus dictating transcriptional activity [10].

The dynamic modifications of the chromatin structure are mostly observed during alterations in the transcriptional activity. But the implications of these modifications may extend to nuclear processes as diverse as cell cycle progression, and DNA replication and repair. For instance, phosphorylation of serine 10 of histone H3 is crucial for chromosomal condensation and cell cycle progression during mitosis and meiosis [11]. For the regulation of transcriptional activity it is particularly relevant the “dialogue” between methylated and acetylated marks and, indeed, the list of enzymes that potentially acetylate, deacetylate or methylate nucleosomal histones has been growing in the last years. In vitro studies suggest that histone acetyltransferases (HATs) and deacetylases (HDACs) can target different lysine residues within histones, while most histone methyltransferases (HMTs) have higher specificiﬁty for particular arginine or lysine residues [12].

Chromatin modulators possess at a set of conserved domains (including bromo and chromo domains) that catalize or recognize histone modifications [13]. These protein modules bind specifically to different lysine modifications and can thus act as starting transmission points of appropriate regulatory signals. Specifically, the bromo domain interacts selectively with acetylated lysines and is in general linked to transcriptional activity, whereas the chromo domain may work as a recognition module for methylated marks and is typically associated with gene silencing and assembly of heterochromatic domains [13]. Additionally, within lysine modifications, the protein domains may be specific depending on the position of the residue in the histone. For instance, the chromo domain from HP1 is selective for H3-K9, and only poorly binds to H3 peptides with methylated lysine K4 [4,14]. Moreover, lysine residues may be mono-, di- or tri-methylated, adding even more complexity to the signalling cues generated by this mark [15].

Not all methylated marks correlate with gene silencing, and some acetylated marks repress instead of activate transcription [16]. For instance, H3-K4 methylation seems to constitute an euchromatic mark, and methylation of arginines in histones H3 and H4 synergistically lead to transcriptional activation. By contrast, acetylation of H4-K12 seems to reinforce a silent chromatin state.

Histone modifications are interdependent and can favour or repress other modifications. In histone H3, phosphorylation of serine 10 inhibits methylation of K9 [17], and may act in a synergistic manner with acetylations of K9 and K14, or methylation of K4 [18-20]. On the other hand, deacetylation of H3-K14 facilitates the subsequent methylation of K9 [21].

3. DNA methylation as an epigenetic mechanism

In the nucleus of mammal cells, the stable silencing of a gene, i.e. maintained in a hereditary manner, is frequently correlated with DNA methylation in its promoter, along with specific modifications in the N-terminal tails of nucleosomal histones. As such, DNA methylation appears as one other important epigenetic mechanism used by the cell, for the establishment and maintenance of the correct patterns of gene expression. Indeed, alterations in the patterns of genomic methylation are strongly associated with several
human diseases, making the use of specific inhibitors of the processes involved a common practise in their treatment [22].

DNA methylation in mammals occurs in the cytosine of the CpG dinucleotide via a reaction catalysed by proteins named DNA methyltransferases (DNMT). In mammals, there are three of these proteins whose presence is crucial to embryonic development: DNMT1, DNMT3A and DNMT3B [23,24]. DNMT1 is referred to as the maintenance methyltransferase, as it possesses the capacity to reproduce the methylation pattern of a DNA sequence during replication, due to its preference to hemi-methylated substrates. The proteins DNMT3A and DNMT3B are mainly involved in de novo methylation. They are therefore important for the establishment of new methylation patterns of the genome.

The CpG islands, regions with more than 500 bp and a G + C content larger than 55%, are localized in the promoter regions of ~40% of all the genes in mammals and are normally maintained in the non-methylated form [23,25]. The stable silencing of tumour suppressor genes in several human cancers, as well as of lineage specific genes during cell differentiation, frequently involves methylation of CpG islands, but these modifications seem to be preceded by modifications of nucleosomal histones. A complex interplay between histone and DNA marks may then stabilize the repressive chromatin structure and thereby manifest to transcriptional inactivity [26,27].

Methylated cytosines can serve as binding platform for specific proteins. On the other hand, this modification can also prevent binding of proteins to DNA. An example of this last case may be observed in the mechanism of regulation in also prevent binding of proteins to DNA. An example of this specific proteins. On the other hand, this modification can [28,31].

The recognition of methylated cytosines is done by proteins that possess a specific binding domain, the so-called methyl-CpG binding domain (MBD). This set of proteins includes the methyl CpG binding proteins MeCP2, and the methyl-CpG binding domain (MBD). This set of proteins that possess a specific binding domain, the so-called family of proteins containing the so-called JmjC domain with [28].

The mechanism of methyltransferase recruitment to specific regions of the genome is not entirely known, but it apparently involves interactions with chromatin proteins, transcriptional factors (see below) or even RNAs [29]. Indeed, the synthesis of an antisense RNA covering a CpG island may induce the methylation of the island [30], similar to that observed in the normal imprinting mechanism of somatic loci (e.g. Igf2r in mouse), and during X-chromosome inactivation [28,31].

4. The dynamics of histone and DNA modifications

The dynamic nature of histone modifications determines the stability (or instability) of a given chromatin structure. The concerted relationship between histone acetyltransferases and deacetylases determines the level and the turnover of acetylation (within minutes) [32]. By contrast, the methylated marks “enjoy” a larger half-life, frequently in a time-scale of hours, and are therefore considered as stable modifications that might contribute to epigenetic “memory” at long course [3]. Histone methyltransferases of specific lysine and arginine residues have been known for some time [16]. Methylation of H3-K9 is mostly associated with the assembly of heterochromatin and to the stable silencing of genes. By contrast, methylation of H3-K4 and of some arginines in histones H3 and H4 are related with transcriptional activation. In these cases, for the dynamic regulation of gene expression, the methylated mark has to be actively and dynamically removed. Only recently two different classes of enzymes were described that are capable of removing the methylation of lysines via an oxidative reaction [33], or of antagonising arginine methylation by conversion into citrulline [34,35]. Recently, a new family of proteins containing the so-called JmjC domain with histone demethylases activity has been described [36–38].

One of the first evidences for a functional relationship between DNA and histone methylation came from work in Neurospora crassa, where it was shown that mutations in the HMT that methylates H3-K9 severely compromises genomic DNA methylation [39]. In mouse, regional heterochromatinization of pericentromeric regions also involves the initial methylation of H3-K9 that then drives DNA methylation to the same regions [40], and, in human cells, an interaction between DNMTs and chromatin proteins known to be associated with gene repression and heterochromatin (such as HP1 and H3-K9 HMT) has been reported [41]. A recently described functional link between proteins of the Polycomb group and DNA methyltransferases further suggest that heritable patterns of gene silencing may be established and sustained by the interconnection of these major silencing pathways [42]. This interaction is achieved through a mechanism that involves the direct recruitment of DNMTs (DNMT1, 3A and/or 3B) to regulatory regions of PcG-repressed genes by the H3-K27 methyltransferase EZH2, and the resulting methylation of local CpG dinucleotides, thus implying for the presence of a self-reinforced set of chromatin modifications working in concert to establish and propagate the structure of a silenced state, potentially all through different cell generations. This might work as a platform for the deposition of linker histones and for the binding of additional epigenetic factors, such as methyl-binding proteins, components of the polycomb repressive complex 1 (PRC1), and HDACs and H3-K9 HMTs [43–45].

5. Chromatin structure and human diseases

Chromatin structure affects gene expression as well as replication, recombination and DNA repair. Several human diseases are linked to or are even based on defects in the machinery maintaining and/or modifying chromatin struc-
ture. For instance, DNA methylation patterns are severely altered in tumors with a bias for overall hypomethylation of the genome and hypermethylation of specific CpG rich regions [46,47]. Locally restricted hypermethylation can lead to gene repression, while genomic hypo-methylation particularly of repetitive sequences was suggested to relate to increased genomic instability [48]. The mechanism that drives hypomethylation of large regions of the genome remains unclear. So far hypo-methylation is known to coincide with loss of acetylation at K16 and trimethylation at K20 of histone H4 [49], strengthening the notion of a cooperation between DNA and histone modifications in the establishment of both normal and abnormal patterns of chromatin structures.

Research of the recent years, however, have shed some light on the mechanism leading to locally restricted hypermethylation of some promoters and the consequential silencing of their genes. Such a hypermethylation can result from the abnormal recruitment of DNA methyltransferases by transcription factors which confer locus specificity. Deregulated transcription factors are considered to play a dominant role in the development of leukemias. This idea is supported by analysis of gene-knockout mice, which uncovered crucial roles of several transcription factors in normal hematopoiesis [50], and of individuals with leukemia, in whom transcription factors are frequently miss-expressed or mutated [51]. Several chromosomal translocations, which are associated with specific forms of leukemia, generate fusion genes that encode altered transcription factors [52].

Acute promyelocytic leukemia (APL), for instance, is genetically characterized the 15;17 chromosome translocation fusing the PML gene to the gene of the transcription factor retinoic acid receptor α (RAR α) [53,54]. The expression of the consequential PML-RARα fusion protein in hematopoietic precursor cells blocks differentiation and promotes leukemia development [55–58]. The oncogenic potential of PML-RARα is based on the aberrant silencing of genes including several tumor suppressor genes. PML-RARα, like the wild type form of RARα, represses transcription of target genes through binding to so-called retinoic acid responsive elements (RARE) and subsequent recruitment of co-repressors such as histone deacetylases [59,60] (Fig. 1). Oligomerization mediated by the PML moiety allows PML-RARα to recruit transcriptional corepressors more efficiently and with higher stoichiometry to target promoters [61]. In contrast to wild-type RARα, the fusion protein PML-RARα is thus rendered insensitive to physiological concentrations of retinoic acid (10⁻⁹ to 10⁻⁶ M) that would usually trigger transcriptional activation, and therefore functions as a constitutive and potent transcriptional repressor of RARE-containing promoters.

Specifically, oligomerized PML-RARα was shown to assemble a multi-protein complex containing several enzymatic activities on target promoters [62]. The recruitment of DNMT1 and DNMT3 results in hypermethylation of CpG islands [63] and thereby generates binding sites for the repressive methylated DNA binding protein MBD1 [64]. In addition to DNA methylation, the presence of HDACs and Suv39H1

Fig. 1 – Schematic representation of PML-RARα-mediated epigenetic silencing. The oncoprotein binds to a well-defined DNA sequence and recruits NCoR co-repressor protein, which serves as a binding platform for HDACs/DNMTs/HMTs. The activity of these epigenetic modifier enzymes leads to modifications of histone tails, DNA methylation, and transcriptional silencing. Methylated CpGs and histone tails are in turn potential docking sites for MBDs and HP1 proteins, respectively. Administration of RA – alone or in combination with epigenetic drugs – induces release of the co-repressor complex, and promotes recruiting of the co-activators containing histone acetyltransferases (HAT) and ATP-dependent chromatin remodeling activity.
further contribute to gene repression by inducing histone hypo-acetylation and histone H3 methylation [61,65]. Importantly, these different chromatin modifying activities were shown to collaboratively contribute to PML-RARα mediated gene silencing and block of differentiation. The recent observation that oligomerization of the RARα fusion protein is required for the induction of leukemia in mice [66] moreover substantiates the pathophysiological relevance of this mechanism in vivo.

Recent work from other groups has further expanded these original findings to other leukemic fusion proteins, such as AML1-ETO and PLZF-RARα [67] and to the oncoprotein Myc [68], thus suggesting a more general role for oncoprotein-directed chromatin-alterations in gene silencing and tumor development (DNA methylation being only one of them). By contrast, inappropriate activation of genes may be observed when enzymes involved in epigenetic modifications are mis-targeted and/or function as constitutive transcriptional activators, as it may be the case in leukemic cells expressing MLL-fusion proteins [62].

Epigenetic alterations can occur faster and more frequently than genetic mutations and might account for the fast evolution of many hyper-proliferative diseases. On the other hand, epigenetic changes are potentially reversible, and thus they make attractive targets for therapeutic intervention. For instance, dissociation of co-repressor complexes from the fusion protein PML-RARα is achieved with pharmacological doses of retinoic acid (10⁻⁶ M). Yet, clinical evidence indicates that, in these conditions, retinoic acid is insufficient to cure APL. Most probably, epigenetic modifications are still persistent at PML-RARα target promoters. Combinatorial treatment of retinoic acid and epigenetic drugs (such as HDAC/HMTi's inhibitors and/or DNA demethylating agents) are more likely able to reverse these chromatin modifications thus regaining transcription of the silenced gene.

Acknowledgments

The work from our laboratory described in this review was supported by grants from the Spanish “Ministerio de Educación y Ciencia”-MEC (BFU2004-03862/BMC) and the Fundación “La Caixa”. ASQ is supported by a fellowship from Fundaçao Calouste Gulbenkian; MB by an EMBO Fellowship.

REFERENCES


