**Abstract**

Chromatin-associated enzymes are responsible for the installation, removal and reading of precise post-translation modifications on DNA and histone proteins. They are specifically recruited to the target gene by associated factors, and as a result of their activity, they contribute to modulating cell identity and differentiation. Structural and biophysical approaches are broadening our knowledge on these processes, demonstrating that DNA, histone tails and histone surfaces can each function as distinct yet functionally interconnected anchoring points promoting nucleosome binding and modification. The mechanisms underlying nucleosome recognition have been described for many histone modifiers and related readers. Here, we review the recent literature on the structural organization of these nucleosome-associated proteins, the binding properties that drive nucleosome modification and the methodological advances in their analysis. The overarching conclusion is that besides acting on the same substrate (the nucleosome), each system functions through characteristic modes of action, which bring about specific biological functions in gene expression regulation.

**Keywords** chromatin; epigenetics; molecular recognition; nucleosome; structural biology

DOI 10.15252/embj.201593377 | Received 30 October 2015 | Revised 30 November 2015 | Accepted 10 December 2015 | Published online 19 January 2016

The EMBO Journal (2016) 35: 376–388

See the Glossary for abbreviations used in this article.

**Introduction**

Chromatin dynamics and structural plasticity play a central role in defining cell function and fate. Proteins associated with chromatin, such as transcription factors, remodelers and modifiers, determine and maintain patterns of DNA and histone modifications that regulate chromatin accessibility and, therefore, cell state and differentiation. While DNA is affected by methylation and derived chemical modifications of cytosine residues, histone proteins are subject to a variety of post-translational modifications that include, but are not limited to, methylation of lysines and arginines, phosphorylation of serine and tyrosine residues, acetylation, crotonylation and butyrylation as well as ubiquitination of lysines. An extensive overview of the chemistry and effective role in chromatin regulation for these and other modifications can be found in recently published works and reviews by different research groups (Tan et al, 2011; Dai et al, 2014; Rothbart & Strahl, 2014; Tessarz & Kouzarides, 2014; Bowman & Poirier, 2014; McGinty & Tan, 2015). Despite their different nature and effect, all these modifications are installed, recognized or removed by dynamic processes, which affect and control chromatin accessibility, gene regulation and activity, as well as the repair of damaged DNA, key to maintaining genome stability.

The diverse functions of histone and DNA modifications are brought about by large protein complexes that regulate a number of biological processes, ranging from DNA replication, transcription and repair to cellular division, differentiation and apoptosis. Functional alterations in these complexes can lead to pathological conditions, such as developmental disorders and genetic syndromes, metabolic and immune system deregulation, aberrations in cell maturation and carcinogenesis (Jin et al, 2008; Lee et al, 2014; Rusconi et al, 2014; Tunicov et al, 2014; Haladyna et al, 2015; Mehdipour et al, 2015; Vallianatos & Iwase, 2015; Wang et al, 2015). Interestingly, many of the chromatin-modifying enzymes use metal (e.g. zinc, iron) or organic (e.g. NADH, acetyl-CoA, FAD) cofactors in their active sites to perform their functions. Moreover, they often use co-substrates metabolites (e.g. acetyl-CoA, S-adenosyl-methionine, ATP) and oxygen that are directly involved in biosynthetic and energy-producing pathways. In this way, histone modifications are inherently linked to energetic and oxidative metabolism, a property of chromatin biochemistry whose implications for cell physiology and disease are only now beginning to be appreciated (Forneris et al, 2005; Karytinos et al, 2009; Southall et al, 2009; Kaochar & Tu, 2012; Jeltsch, 2013; Kaelin & McKnight, 2013; Meier, 2013).

In this review, we analyse the known atomic resolution structures of nucleosome complexes and discuss the interplay between modes of nucleosome binding and the mechanisms promoting histone modification reactions [with regard to chromatin remodelling complexes and histone chaperones, the reader is referred to excellent recent reviews: Bartholomew (2014), Gerhold and Gasser (2014) and Mattioli et al (2015)]. The emerging notion is that intricate domain and subunit compositions, often involving both readers and
The heart of the matter: where and how is the nucleosome specifically recognized and modified?

A carefully tuned equilibrium of deposition, decoding and removal of post-translational modifications is crucial for the control of cell functions and differentiation. Consequently, chromatin-associated proteins and enzymes need to be specifically targeted to the designated genomic loci and have to correctly interpret the information harboured by DNA and histones. Although the past decades have elucidated the chemical enzymology of histone-modifying reactions, the question how nucleosomes are recognized is much less well understood. What is the role and contribution of DNA and histone proteins in nucleosome recognition? How are nucleosomes individually and selectively recognized based on their patterns of epigenetic modifications? How is the balance of competing interactions among the many different players fine-tuned within cells? And how are nucleosome-modifying enzymes targeted to specific DNA sequences?

These and other specific mechanistic questions can now be addressed thanks to recent methodological advances that allow an increased understanding of the molecular processes underlying nucleosome recognition and histone modification. A major step forward has been made with crystallographic and cryo-EM analyses of the nucleosome, tetranucleosomes and chromatin fibres (Luger et al., 1997; Schalch et al., 2005; Song et al., 2014). Another important advance is the development of techniques to probe the transient and dynamic nature of protein–protein and protein–DNA interactions by various types of fluorescence-based methods, mass spectrometry and NMR. As an additional tool, the availability of recombinant nucleosome libraries has facilitated the screening of different sets of post-translational modifications in a high-throughput manner employing either barcoded DNA or peptides clipped to tailless-reconstituted nucleosomal particles (Angelov et al., 2001; Nguyen et al., 2014). Box 1 and Table 1 provide an overview of some of the current methods available for the analysis of nucleosome molecular biology.

Surfing the nucleosome: a mosaic of protein and DNA

To fully understand how nucleosomes are recognized, it is important to first describe this rather complex object. As outlined in Fig 1,
Mechanisms of nucleosome recognition

The EMBO Journal

Valentina Speranzini et al

Box 1: Methodological tools employed for deciphering nucleosome recognition mechanisms

Current research on nucleosome recognition mechanisms relies on a broad range of constantly developing methodologies. Biochemistry and biophysics are combined with structural biology to determine molecular interactions at the basis of these mechanisms and to unravel the architecture of chromatin-associated complexes.

1. Structural studies on nucleosomal particles and associated complexes

Structural investigations on nucleosome-associated complexes are based on X-ray small-angle scattering and diffraction in both solution and crystals, together with NMR, and, increasingly, cryo-electron microscopy, thanks to the improved resolution that can be achieved (Kuhlbrandt, 2014; Bai et al, 2015). With recent developments in both instrumentation and methodology, these techniques represent a pool of versatile spectroscopic and diffraction applications for the analysis of macromolecules, enabling the characterization of molecules and complexes in a wide range of size and architecture.

2. Methodologies for the characterization of protein–protein and protein–DNA interactions

In parallel to structural analyses, many highly sensitive techniques to detect protein–protein and protein–DNA interactions have been developed with the aim of interrogating elusive transient binding. Many of these techniques make use of fluorescence, which overcomes limitations of dynamic range posed by other techniques, such as isothermal calorimetry, that is however successfully employed for the analysis of nucleosome-derived peptides. Among these fluorescence-based techniques, we find fluorescence polarization (FP), a versatile in-solution method that allows quantitative and rapid quantification of protein interactions is also exemplified by the HI-FI system, which was developed with the specific aim of analysing molecular interactions between chromatin components and associated protein complexes from a quantitative viewpoint (Hieb et al, 2012; Winkler et al, 2012). This assay allows the measurement of binding affinities, interaction specificity and complex stoichiometry in microplate formats, followed by direct transfer of the samples onto native gels for further analyses (Kalashnikova et al, 2013; McGinty et al, 2014; Zheng et al, 2014; Kim et al, 2015).

3. Analytical methodologies

Many of the above-presented strategies integrate with more classical yet fundamental analytical methodologies such as mass spectrometry, pull-down assays, electrophoretic mobility assays and analytical ultracentrifugation. Combined with methodologies that determine molar mass and average size of molecular species in solution using light scattering (MALLS and DLS), these analytical techniques are crucial to the identification of exact composition and modifications of the partners engaged in the formed complexes.

4. Chemical probes and nucleosome libraries

Photochemical methodologies that facilitate studies of either small or large protein assemblies have been—and continue to be—improved and implemented in nucleosome binding studies (Bartholomew et al, 1990; Forget et al, 2009; Persinger & Bartholomew, 2009; Hota et al, 2012; Kim et al, 2015; Male et al, 2015). Deoxynucleotide analogues are used to chemically modify DNA, which thereby becomes site-specifically photoreactive. Moreover, the availability of reconstituted nucleosomes and their semi-synthetic variants bearing specific chemical modifications has enabled the generation of large nucleosome libraries for analytical experiments on binding properties (Montel et al, 2007; Simon, 2010; Maltby et al, 2012; Pichler et al, 2012; Yun et al, 2012; Lee et al, 2013; Rogge et al, 2013; Al-Aini et al, 2014b; David et al, 2015).
the nucleosome is formed by an octamer of tightly associated histone proteins (H2A-H2B)\(_2\)(H3-H4)\(_2\), and \(\sim150\) bp of DNA wrapped around the octamer to define a left-handed superhelical fragment (Luger et al, 1997), where the core of the histone proteins is well defined, while the tails mostly lack a defined structure. The presence of two copies of each of the four histone proteins generates intrinsic twofold symmetry. Variations in this basic structure can be introduced by histone variants such as CENP-A, replacing H3 in the centromeric chromatin, H2A.Z, important for transcriptional regulation, or H3.3, involved in specific developmental processes in mammalian cells (Fan et al, 2004; Falk et al, 2015; Jang et al, 2015). The histone proteins are predominantly positively charged, making them prone to associate with negatively charged surfaces, such as the nucleosomal DNA itself (Bannister & Kouzarides, 2011; Iwasaki et al, 2013). Despite their overall positive charge, the H2A and H2B cores have a number of negatively charged glutamate residues that combine to form the so-called acidic patch on the side of the nucleosomes (Figs 1 and 2A). The flexible tails of the histone proteins are a critical feature of the nucleosome, as they are the main target of epigenetic modifications. Since many combinations of modifications are possible, individual nucleosomes can organize into higher-order structures that are characterized and regulated by a large variety of combinations of post-translational modifications. As a result, the nucleosomal particles offer multiple chemically and topologically distinct binding sites that are key to the recognition processes analysed in this work: the wide protein surface on each side of the nucleosomal disc, the DNA wrapped around it and the flexible and highly charged histone tails. Different chromatin-associated proteins have adopted distinct binding strategies, which differentially and specifically exploit the available binding surfaces. In these multi-protein complexes, recognition and modification modules (enzymes, subunits or domains) can engage chromatin through coordinated interactions that involve multiple attachment points.

In this review, we will describe protein-nucleosome complexes according to the binding sites and surfaces that contribute most to nucleosome recognition (Figs 1 and 2). We first consider contacts to the histone sides, especially to the acidic patch, and then present recognition mechanisms that involve a combined DNA and protein binding, and finally, we will discuss how modifier-associated regulators can direct nucleosome recognition. In each section, we highlight how these distinct anchoring points contribute to direct the engagement of the target substrates (mainly histone tails) and the modification of epigenetic marks.

**Catching the nucleosome from its histone side**

Historically, one of the best-characterized features of the nucleosome that serves as an anchoring point for binding is the “acidic patch” (Figs 1, right panel, and 2A) (Luger et al, 1997). Mutagenesis and structural studies have demonstrated that this region is key for the interaction with many proteins. A prime example is the histone H4 tail, which specifically contacts this patch to promote inter-nucleosomal networks associated with chromatin condensation (Dorigo et al, 2003; McBryant et al, 2009). The role of the acidic patch is further illustrated by CENP-C that interacts with histone variant CENP-A during kinetochore assembly and mitosis (Milks et al, 2009). The combination of NMR and X-ray crystallography has allowed the recognition regions between these centromeric proteins to be mapped and further indicated that four specific arginines in the central domain of CENP-C bind the nucleosome by digging deeply into the acidic patch (Fig 2B) (Kato et al, 2013). This specific “arginine anchor” type of interaction is shared by other well-characterized nucleosome-associated proteins. The latency-associated nuclear antigen (LANA) interacts with its host to allow viral episome attachment to segregating chromosomes (Ballestas & Kaye, 2011). Using the N-terminal region of LANA, structural studies showed that the antigen folds into a hairpin that interacts with the glutamate residues on the H2A-H2B surface through such an “arginine anchor” (Fig 2C) (Barbera et al, 2006). A coordinated structural change between protein complex and nucleosome underlies the recognition mechanism for Sir3, a protein responsible for the formation of silent heterochromatin in *S. cerevisiae* (Buhler & Gasser, 2009; Armache et al, 2011; Arnaudo et al, 2013; Wang et al, 2013; Yang et al, 2013). In this case, binding to the acidic patch of the nucleosome involves otherwise flexible loops within the N-terminal bromo-like BAH domain of Sir3 (Fig 2D).

The acidic patch represents a platform for nucleosome recognition and modification also for enzyme complexes, such as histone-ubiquitylating enzymes. These factors operate through ubiquitin-charged E2 enzymes and an E3-ligase that recognizes the target and activates the transfer (Metzger et al, 2014). An E2 with a dimeric RING E3-ligase, present in the ubiquitylation module of Polycomb complex PRC1, was one of the first enzymes crystalized in complex with the nucleosome. Binding of the Ring1B/Bmi1 component of PRC1 to the central acidic patch relies on a single arginine residue (i.e. an arginine anchor) that engages the acidic patch. This allows the correct positioning of the E2 subunit, which establishes extensive contact points with the nucleosomal DNA (Table 1; Fig 2E) (Leung et al, 2014; McGinty et al, 2014). As a result of this process, Lys119 of H2A can be targeted for ubiquitylation. Another well-characterized system that uses the acidic patch as landing platform is the RING E3 ligase RNF168, which is responsible for the ubiquitylation of Lys13–Lys15 on the N-terminal tail of H2A. Intriguingly, histone modification by RNF168 requires activation of the catalytic site by the acidic patch—at a distance from the actual substrate (Mattiolo et al, 2014). The outstanding feature of this mechanism is that the patch interaction with the E3-ligase helps to position a common E2 enzyme (UbcH5C) to target the correct substrate lysine over a long-range distance, on distinct areas of the nucleosomal surface. In essence, these ubiquitylating enzyme complexes require a coordinated binding between the acidic patch as a “recognition mediator” and the actual site of ubiquitination (Fig 3A).

**Synergy between DNA and protein binding as key for recognition**

A second “recognition element” on nucleosomes is the coiled DNA. An interesting question is to what extent specific DNA sequences contribute to recognition. Structural studies on transcription factors have revealed mechanisms that explain how these regulators bind preferentially to a specific DNA sequence at their target site. The affinity of a transcription factor for a given position with a specific
sequence can be determined by patterns of non-covalent interac-
tions between amino acid side chains and the exposed surfaces
inside DNA grooves. In other cases, it is the physical property of the
DNA itself, bent in a precise conformation dictated by sequence
content, that guides specific recognition (Choi & Rajkovic, 2006;
Rohs et al., 2010; Deng et al., 2012; Ryzhikov et al., 2014). In
contrast, chromatin modifiers often do not use sequence-specific
contacts but rather find non-specific anchoring points to land on the
nucleosomal DNA.

An instructive example that combines contributions of DNA
and histone binding is observed in the H3K36me3 reader PSIP1,
an essential transcriptional coactivator subunit of the histone
methyltransferase mixed-lineage leukaemia-1 (MLL1) core complex
(Fig 2F) (Yokoyama & Cleary, 2008; Patel et al., 2009). This system
coordinates the two processes in one domain, PWPP, that binds
both H3K36me and DNA (Maurer-Stroh et al., 2003; Lukasik et al.,
2006; Laguri et al., 2008; Yokoyama & Cleary, 2008; Vermeulen
et al., 2010). Recently, the recognition process involved has been
clarified using NMR analysis: van Nuland and colleagues demon-
strated that the affinity for the nucleosome is cooperatively
increased when both DNA and methylated histone tail are engaged.
In particular, the unspecific contacts with the DNA backbone are
essential for engagement of the histone tail inside the aromatic cage
of the domain (van Nuland et al., 2013).
Figure 2. Nucleosomes are recognized through different surfaces.

Structural analysis of complexes between protein assemblies and nucleosomes shows that the particle is recognized by anchoring to either DNA (in white) or histone core surfaces (octamer, in cyan). (A) Schematic overview of the three major nucleosome contacts presented in (B–H). Acidic patch is highlighted in orange, whereas DNA anchoring points are indicated in green (DNA wrapped around nucleosome) and blue (DNA at entry point and/or dyad). (B) Centromere protein CENP-C recognizes histone H3 variant CENP-A by docking through hydrophobic interactions onto the nucleosome acidic surface (PDB: 4X23). (C) Kaposi’s sarcoma herpes virus LANA peptide forms a hairpin that interacts specifically with the nucleosome acidic patch formed by histones H2A-H2B (PDB: 1ZLA). (D) Silencing regulator Sir3 establishes multiple molecular interactions with the nucleosomal surface and its modifications (PDB: 3TU4). (E) Repressive complex PRC1 anchors the nucleosome through both the acidic patch and nucleosomal DNA (PDB: 4RBP), subunit Bmi1 in yellow, UbH5c in orange and Ring1B in chocolate. (F) The PWWP domain of PSIP1 H3K36me3 is positioned at the side of the nucleosome, interacting with methylated tails thanks to a cooperative binding with DNA surface (PDB: 3ZH1). (G) Histone H3 demethylase heterodimer complex LSD1–CoREST is positioned as a clamp on the nucleosomal particle, which is engaged by DNA and histone tail binding (Pilotto et al, 2015). In this model derived from SAXS data, LSD1 is depicted in firebrick, CoREST in blue. (H) Regulator of chromosome condensation protein RCC1 binds to the acidic patch in a similar mode as PRC1, CENP-C and LANA peptide (PDB: 3MVD).
The contribution of DNA has been structurally visualized also for the histone demethylation complex LSD1-CoREST (Fig 2G). The corepressor CoREST comprises a C-terminal SANT domain, reported in the literature to serve as a DNA-binding module essential for the activity of histone H3K4me2 demethylation (Aasland et al., 1996; Boyer et al., 2002; Shi et al., 2005; Forneris et al., 2007). Indeed, the LSD1-CoREST complex performs a thorough yet general scan of nucleosomal DNA as a first step for substrate recognition, which is then finalized by the capture of histone H3 for lysine modification (Fig 3B). Consistently, impaired DNA binding causes inefficient nucleosomal recognition, leading to a reduced capability of substrate modification (Pilotto et al., 2015). The central role of DNA in this process has been corroborated by Kim and co-workers. Their study showed that the linker DNA connecting nucleosomal particles enhances the functionality of the histone demethylase complex and promotes anchoring to a nucleosomal particle (Kim et al., 2015). How such an unspecific DNA recognition translates into specific recruitment on determined genomic loci emerges only when this heterodimeric protein is analysed in the context of larger assemblies. It has been demonstrated that many transcription factors, for example Snail1, Gfi1 or Insin1, function as histone mimics and specifically interact with the LSD1 active site at precise developmental stages and in distinct cellular types. In this way, LSD1-CoREST is selectively recruited to ensure the fine-tuning of transcription at desired loci to trigger specific downstream effects, such as cytokine signalling pathways, haematopoiesis, endocrine cell maturation or epithelial to mesenchymal cell transition (Saleque et al., 2007; Lin et al., 2010; Baron et al., 2011; Welcker et al., 2013). In this case, nucleosome binding is therefore a combination of transcription factor, histone and DNA recognition, where the target epigenetic marks contribute only at a later stage of the process.

Another remarkable example of nucleosome recognition involving DNA binding is given by RCC1. This guanine exchange factor is responsible for the recruitment of the Ran GTPase during nucleosome condensation (Bischoff & Ponstingl, 1991). Decades after these first studies, the details of the interaction between RCC1 and the nucleosome were revealed by X-ray crystallography (Fig 2H) (Makde et al., 2010). Looking at the structure of the complex, one loop of RCC1, located away from the interface with Ran GTPase, interacts specifically with the residues of the acidic patch on the nucleosome, strongly anchoring the unit and possibly competing out histone H4 during chromatin condensation. In addition, non-specific DNA binding to backbone phosphates contributes to the interaction. This system thus outlines a possible double anchoring scheme that makes use of both DNA and the acidic patch for association with the nucleosome, thus stabilizing the resulting complex.

In summary, DNA binding often serves the purpose of detaching histone tails from the nucleosomes. LSD1-CoREST and PSIP1 competitively displace histone H3 tails by directly binding to nucleosomal DNA to capture the substrate. Conversely, binding to DNA of RCC1 promotes docking on the particle, eventually leading to the displacement of the H4 tail from the acidic patch.

Regulatory proteins stimulate modification activity

The protein complexes and modifiers discussed in the previous paragraphs are mainly sub- or core complexes. Indeed, all of them form parts of larger assemblies that contribute secondary binding sites, which in turn mediate protein–protein interactions with other components of complexes associated with and working on
Table 1. Structural and methodological data on the cited chromatin-associated complexes.

<table>
<thead>
<tr>
<th>Protein/complex (PDB/EMD ID)</th>
<th>Anchoring point(s)</th>
<th>Methods for complex studies</th>
<th>Reference(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNTTIP1 (2MWL, 4DEK)</td>
<td>DNA</td>
<td>X-ray crystallography, NMR, circular dichroism, mass spectrometry</td>
<td>Itoh et al. (2015)</td>
</tr>
<tr>
<td>HP1 or SWI homologue (1EOB, 1KNE, 2RSC)</td>
<td>DNA, histone mark</td>
<td>FP, negative stain and cryo-EM, AUC, EPR spectroscopy</td>
<td>Jacobs &amp; Khorasanizadeh (2002), Ishida et al. (2012), Canzio et al. (2013)</td>
</tr>
<tr>
<td>Intasome (EMD-2992, 3OS1)</td>
<td>DNA, H2A-H2B heterodimer</td>
<td>Negative stain and cryo-EM</td>
<td>Maertens et al. (2010), Maskell et al. (2015)</td>
</tr>
<tr>
<td>LANA peptide (1ZLA)</td>
<td>Acidic patch</td>
<td>X-ray crystallography</td>
<td>Barbera et al. (2006)</td>
</tr>
<tr>
<td>LSD1-CoREST (2V1D)</td>
<td>DNA, histone mark</td>
<td>Fluorescence polarization, SAXS, MLA installation, photo-crosslinking, HI-FI</td>
<td>Forneris et al. (2007), Kim et al. (2015), Pilotto et al. (2015)</td>
</tr>
<tr>
<td>LSD2-NPAC (4GUR, 4GUS)</td>
<td>Histone mark</td>
<td>X-ray crystallography, ITC, mass spectrometry</td>
<td>Fang et al. (2013)</td>
</tr>
<tr>
<td>Nucleosome, tetranucleosomes, chromatin fibre (1AOI, 1KXS, 1ZBB, 3L2O, EMD-2600, EMD-2601, EMD-2602)</td>
<td>–</td>
<td>X-ray crystallography, cryo-EM</td>
<td>Luger et al. (1997), Davey et al. (2002), Schalch et al. (2005), Vasudevan et al. (2010), Song et al. (2014)</td>
</tr>
<tr>
<td>PR1C (4RBP)</td>
<td>Acidic patch, DNA</td>
<td>X-ray crystallography</td>
<td>McGinty et al. (2014)</td>
</tr>
<tr>
<td>PR2C (5CH1, 5CH2)</td>
<td>Histone mark</td>
<td>X-ray crystallography</td>
<td>jiao &amp; Liu (2015)</td>
</tr>
<tr>
<td>P5IP1 (3ZH1, 3ZH1)</td>
<td>DNA, histone mark</td>
<td>NMR, strip-FRAP, MLA installation</td>
<td>van Nuland et al. (2013)</td>
</tr>
<tr>
<td>RCC1 (12LM, 3MVD)</td>
<td>Acidic patch, DNA, histone mark</td>
<td>X-ray crystallography</td>
<td>Renaud et al. (2001), Makde et al. (2010)</td>
</tr>
<tr>
<td>RNF168 (B111)</td>
<td>Acidic patch, histone mark</td>
<td>Gel-shift assays, fluorescence polarization</td>
<td>Campbell et al. (2012), Mattirolli et al. (2014)</td>
</tr>
<tr>
<td>Rpd3S (Eaf3: 2K3X)</td>
<td>DNA, histone mark</td>
<td>NMR, deuterium exchange mass spectrometry, MLA installation</td>
<td>Xu et al. (2008)</td>
</tr>
<tr>
<td>Sir3 (3TU4, 4KUD, 4LD9)</td>
<td>Acidic patch</td>
<td>X-ray crystallography, mass spectrometry, ITC, AUC</td>
<td>Armaiche et al. (2011), Arnaudo et al. (2013), Yang et al. (2013)</td>
</tr>
</tbody>
</table>

chromatin. These large entities comprise accessory proteins that regulate the activity of their partner enzymes by mediating and enhancing the catalytic activities related to nucleosome binding, readout or modification (Fig 3C). This notion is nicely portrayed by H3K4 methyltransferases, which share structural domain composition. The very recent crystal structure of Polycomb repressive complex 2 (PRC2) provides hints about H3K27 trimethylation involving multiple enzymes surfaces for histone tail recognition (Jiao & Liu, 2015). In turn, H3K27me3 serves as a recruitment signal for canonical PRC1 complexes with a Polycomb (CBX) subunit that for unrelated non-catalytic domains in distinct protein complexes (Fang et al., 2010; Mosammaparast et al., 2013; Zhang et al., 2013; Saez et al., 2015). More specifically, LSD2 has been found to bind NPAC, a putative oxidoreductase (Fang et al., 2013). Although the precise physiological role of NPAC is still not understood, it has been shown both in vitro and in vivo that this protein functions as a cofactor that enhances the demethylase activity of H3K4 of LSD2. Structural analysis of the ternary complex deposition of methyl groups. Acting together with the other modules RbBP5 and Ash2L, WDR5 regulates the product specificity of the MLL1 core complex. Complex formation leads to the organization of a secondary catalytic module that can accommodate the product of SET monomethylation and redirect overall activity towards multiple methylation (Patel et al., 2009, 2014).

Another example of partner/enzyme interactions required for chromatin modification is represented by LSD2. This enzyme belongs to the same family of flavin-dependent histone demethylases as LSD1, but it functions through a different mechanism for nucleosome recognition as it does not interact with CoREST (Karytinos et al., 2009). LSD1 and LSD2 represent a prime example of how the same catalytic module (a flavin-dependent amine oxidase) can be associated with unrelated non-catalytic domains in distinct protein complexes (Fang et al., 2010; Mosammaparast et al., 2013; Zhang et al., 2013; Saez et al., 2015).
LSD2-NPAC-H3 showed that the oxidoreductase interacts with the demethylase through a deep hydrophobic pocket located close to the active site, in direct contact with the histone tail substrate. In this way, NPAC is proposed to directly contribute to substrate binding by LSD2 but without inducing conformational changes as in the case of SET methyltransferases.

In other cases, the chromatin itself can induce a conformational change in a protein subunit that is subsequently transmitted to the enzymatic components of the protein complex. Rpd3S is a histone deacetylase involved in the stabilization of chromatin during transcription (Keogh et al., 2005). To achieve nucleosome modification, the Rpd3S complex combines the activity of two modules, Eaf3 and Rco1, responsible for DNA and histone recognition. Rpd3S normally rests in an auto-inhibited state, in which Eaf3 cannot access histone H3. Upon contact with chromatin, this closed state is released due to a conformational change, and the SID region within Rco1 stimulates specific binding of the other subunit, Eaf3, to H3K36me (Ruan et al., 2015). Therefore, when not bound to the target, these enzymatic complexes rest in a closed state preventing recognition. This state is stimulated by contact with chromatin to allosterically activate modification of histone substrates (Fig 3D). Likewise, HP1 uses a mechanism based on a resting, auto-inhibited conformation that becomes active upon binding to chromatin, in a step that is directly responsible for the spreading of heterochromatin for genetic repression (Canzio et al., 2013).

On a larger scale, the MiDAC complex comprises important enzymatic activities for chromatin modification, such as histone deacetylases HDAC1/2. MiDAC also encompasses MEDES, a co-repressor protein bearing ELM-SANT domains for protein–protein and protein–DNA interactions. Itoh and collaborators characterized DNTTIP1 as a DNA-binding subunit that is fundamental for MiDAC functionality (Itoh et al., 2015). In this case, the non-specific anchoring to DNA implemented by DNTTIP1 induces a conformational change in a protein subunit that is subsequently transmitted to the enzymatic components, making a major contribution to nucleosome engagement for subsequent modification.

The idea emerging from these studies is that ancillary partners can either directly or through long-range conformational changes alter the catalytic and binding sites of the enzymatic components responsible for histone modification, effectively controlling the productivity (turnover and/or substrate affinity) of the enzymatic step.

DNA deformation in nucleosome recognition

The malleability of DNA can underline selective nucleosome recognition, a concept that is beautifully demonstrated in recently reported work on the mechanism of retroviral DNA integration into the genome (Maskell et al., 2015). It is known that retroviruses use their integral transposases to severely bend and distort target DNA and disrupt backbone phosphodiester bonds for integration (Kvaratskhelia et al., 2014). Within the nucleosome, DNA is wrapped in a quite compact fashion that apparently would not leave enough room for retroviruses to integrate. However, the structural insights obtained by electron microscopy show that in the process of integration, gene activation would lead to a transient relaxation of the nucleosomes. This enables the viral machinery to opportunistically bind and integrate in such genomic locus. The relative assembly of the integrase machinery and the nucleosomal particle displays a broad network of interactions involving three of the integrase subunits, viral and nucleosomal DNA, and one H2A–H2B heterodimer. Upon interaction and binding, viral DNA strongly interacts with and deforms nucleosomal DNA, which is almost displaced from the core surface in a bent and unprotected conformation, allowing the subsequent integration of the external DNA in the desired locus.

Concluding remarks

Protein complexes that bind and modify chromatin select their targets through a variety of mechanisms, many of which rely on association with other protein partners. These processes cooperatively lead to specific engagement of nucleosomal particles. Each enzymatic system evolved its own specific approach to achieve this goal, depending on the type and localization of the modification that has to be introduced and/or read and associated with specific functions in chromatin regulation. Clearly, it is often the presence of non-catalytic domains within chromatin-associated enzymes or accessory subunits of macromolecular complexes that generates specificity and context dependence in nucleosome engagement. Along this line, domain reshuffling and combinatorial assembly of single proteins into larger complexes have contributed to the evolution and cell-type specificity of these systems. Chromatin itself functions as an active player in these regulatory mechanisms, serving not only as a substrate but also as an “inducer” of recognition. The regulation of chromatin recognition and access is thus the result of competitive interactions and interplay among different players: histone modifications, histone surfaces, nucleosomal DNA and all the subunits found in small and large molecular complexes associated with nucleosomes. The latest progress also demonstrates that in many cases, it is not the DNA sequence per se that determines the exact anchoring point to chromatin (Makde et al., 2010; van Nuland et al., 2013; Patel et al., 2014; Ahmad et al., 2015; Clark et al., 2015). Both nucleosomal and linker DNA can be strongly bound by modifying enzymes and complexes, which must be first targeted to DNA by specific transcription factors in a selective fashion. In the future, it will be interesting to see how widespread the role of DNA “malleability” is, a characteristic suggested to be crucial by the investigation on intasome complex (Maskell et al., 2015) and, very recently, by the structural studies on RNA polymerase stalled on the nucleosome (Gaykalova et al., 2015).

Recent advances in both imaging and cell biochemistry provide the ability to investigate recognition and modification processes in the context of chromatin fibres and chromosome assembly within eukaryotic cells. It will be interesting to see how these mechanistic chromatin studies can promote the progression of drug discovery and design. There are huge ongoing efforts (including clinical trials) aimed at the study and development of specific compounds able to interfere with epigenetic processes, with a focus on oncological applications (Kooistra & Helin, 2012; Falkenberg & Johnstone, 2014; Finley & Copeland, 2014; Mai, 2014). The challenge ahead will be to see whether non-catalytic surfaces involved in nucleosome recognition can be exploited for targeting by small molecules. 
that selectively interfere with specific disease-related epigenetic mechanisms.

Acknowledgements
We thank Federico Fornieri for critical reading of the manuscript and assistance in figure preparation. We acknowledge Francesca Mattiroli for suggestions on the manuscript. This work was supported by AIRC (IG-15208) and MIUR (Progetto Bandiera Epigenomica - EPIGEN).

Conflict of interest
The authors declare that they have no conflict of interest.

References


Mechanisms of nucleosome recognition

Valentina Speranzini et al

Published online: January 19, 2016

The EMBO Journal


Lee JH, Friso S, Choi SW (2014) Epigenetic mechanisms underlying the link between non-alcoholic fatty liver diseases and nutrition. Nutrients 6: 3303–3325

RING18/BMI1-dependent H2AX and H2A ubiquitination and DNA damage signaling. PloS Genet 10: e1004178


Mechanisms of nucleosome recognition  
Valentina Speranzini et al


