

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

185,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Recognition of Nucleosomes by Chromatin Factors: Lessons from Data-Driven Docking-Based Structures of Nucleosome-Protein Complexes

Velten Horn and Hugo van Ingen

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/intechopen.81016>

Abstract

The function of chromatin ultimately depends on the many chromatin-associated proteins and protein complexes that regulate all DNA-templated processes such as transcription, repair and replication. As the molecular docking platform for these proteins, the nucleosome is the essential gatekeeper to the genome. As such, the nucleosome-binding activity of a myriad of proteins is essential for a healthy cell. Here, we review the molecular basis of nucleosome-protein interactions and classify the different binding modes available. The structural data needed for such studies not only come from traditional sources such as X-Ray crystallography but also increasingly from other sources. In particular, we highlight how partial interaction data, derived from for example NMR or mutagenesis, are used in data-driven docking to drive the modeling of the complex into an atomistic structure. This approach has opened up detailed insights for several nucleosome-protein complexes that were intractable or recalcitrant to traditional methods. These structures guide the formation of new hypotheses and advance our understanding of chromatin function at the molecular level.

Keywords: nucleosome, protein interactions, epigenetics, chromatin binding, acidic patch, histone tails, post-translational modifications, data-driven docking, NMR spectroscopy, XL-MS, crystallography, cryo-EM, structural models

1. Introduction

The packaging of DNA into chromatin represents one of the most fundamental layers of the biology of the cell. It provides the required structural compaction of DNA to fit in the nucleus

and plays crucial roles in controlling cell fate and protecting genome integrity. The fundamental unit of chromatin is the nucleosome in which 147 base pairs (bp) of DNA are wrapped around an octameric protein complex composed of two copies of histone proteins H2A, H2B, H3 and H4 [1–3]. Nucleosomes are arranged as beads-on-a-string forming 10 nanometer (nm) wide fiber that subsequently condense into higher order structures [4]. Nucleosomes as the basis of chromatin are responsible for its dynamics. Chromatin state and changes in DNA accessibility are determined at the nucleosome level. These changes are mediated through interactions of histone proteins and nucleosomal DNA alike with a wide range of protein complexes that control the structure of chromatin. They interpret, write and erase post-translational modifications or act as ATP-dependent nucleosome remodelers. This allows changes in the functional state of chromatin and regulation of DNA-templated processes. While promoting a large variety of effects on chromatin structure, nucleosome-interacting proteins share the molecular basis of recognizing and binding the nucleosome. Understanding the basis of chromatin dynamics therefore demands understanding the molecular basis of nucleosome-protein interactions.

In particular, insights into the molecular mechanistic basis of how histone-modifying enzymes install or remove post-translational modifications (writers and erasers, respectively) and how these modifications are recognized by effector proteins (readers) are of immense interest, especially in drug development. Deregulation of these proteins is strongly connected to pathological outcome, including cardiovascular diseases, neurological disorders, metabolic disorders and cancer [5]. So-called epigenetic drugs that target the nucleosome interaction of these chromatin factors offer new therapeutic potential [6–9]. A selection of epigenetic drugs including those currently undergoing clinical trial is described in detail elsewhere [10]. Advancement in their development requires insights into the underlying molecular mechanism of nucleosome recognition, enabling control over subsequent modification of the chromatin state.

In the following, we will review the molecular basis of nucleosome-protein interactions, focusing on the different binding epitopes presented by the nucleosome. After an overview of the nucleosome-protein structures determined by crystallography or cryo-electron microscopy (cryo-EM), we highlight several studies in which experimental data from nuclear magnetic resonance spectroscopy (NMR), cross-link-based mass spectrometry (XL-MS) or mutational analysis were used to build atomistic structural models of nucleosome complexes. Throughout, we emphasize the role of these data-driven models in deepening our understanding of nucleosome recognition.

2. Nucleosome-binding epitopes

Consisting of DNA and histone proteins, the nucleosome offers a selection of distinct interaction surfaces for binding of effector proteins with high levels of specificity (**Figure 1**).

Histone proteins possess a globular tertiary structure with exposed, disordered N-terminal tails. Histone tails are known to carry a wide range of covalent, post-translational side chain

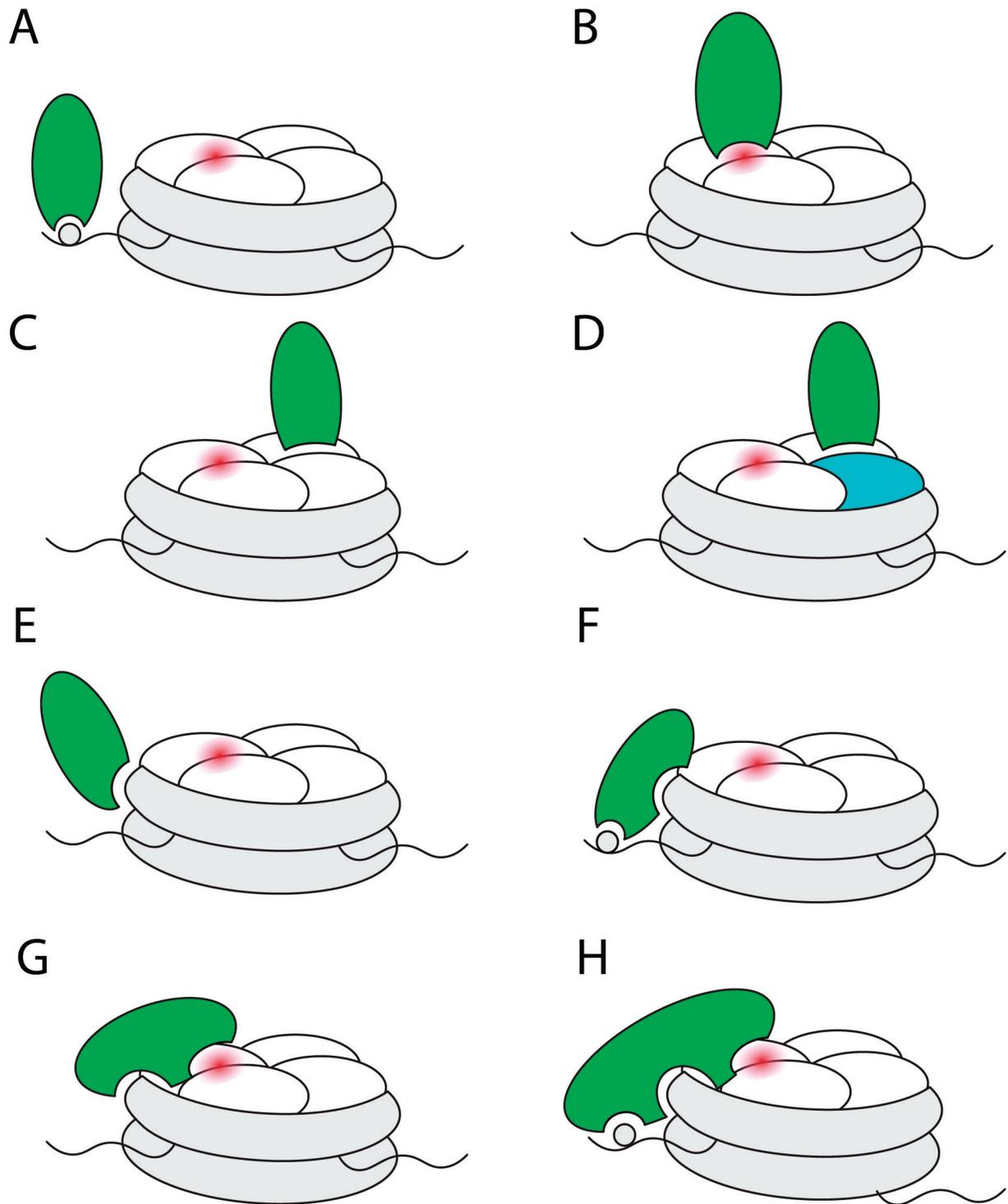


Figure 1. A schematic depiction of different modes of nucleosome recognition. Reported types of epitopes are histone tails including PTMs (A), the H2A-H2B acidic patch (B), the canonical histone surface (C), specific surface motifs formed by histone variants (D) or nucleosomal DNA (E). A manifold of synergetic combinations of binding epitopes are known, such as histone mark and DNA (F), acidic patch and DNA (G) or all three epitopes (H).

modifications (PTMs) such as, mono-, di- and trimethylation (Lys, Arg); acetylation (Lys); phosphorylation (Ser, Thr) and ubiquitination (Lys) [11, 12]. This cosmos of modifications maintains a dynamic nature through the reversibility of the covalent modifications. Modified

histones are recognized by so-called reader protein domains specific for the respective modification (**Figure 1A**). Interestingly, nucleosome-interacting proteins can possess more than one reader domain which allows cross talk between different post-translational modifications. Examples of PTM reader domains are Chromo, Tudor, PHD and MBT domains for methylated lysine residues, bromodomains for acetylated lysine residues and 14–3–3 proteins for phosphorylated serine [11, 13] (**Table 1**). The most recent addition to the list is YEATS domains that recognize crotonylated lysine [14–16]. Reader domains often have structurally conserved motifs that are able to complex a specific modification. The “Royal Family” of reader domains is in this respect a particularly instructive example. This superfamily includes the Chromo, MBT, PWWP and plant Agenet domains that bind methylated lysine (Tudor, Chromo, MBT, PWWP, plant Agenet) or arginine (Tudor) residues. Most domains of this family contain a barrel-shaped structure formed by 3–5 antiparallel β -strands that holds a cluster of aromatic residues that form the so-called aromatic cage [17]. The aromatic cage presents an electron-rich yet hydrophobic surface that is ideally suited to bind methylated lysines through cation- π interactions [18]. The structural features and similarities, as well as their substrate specificity, have been subject to literature reviews [19–21].

Reader domains can, in addition to the post translational modification, show specificity for a defined amino acid sequence motif around the epigenetic mark that supports complex formation. For example, the WD40 domain of the EED (embryonic ectoderm development) protein selectively reads out trimethylated lysine in a A-R-K-S sequence motif (as for H3K27me3) but not in a R-T-K-Q motif (as for H3K4me3) [37].

(sub)Domain	Modification	Protein	Function
Royal family			
Tudor	Kme1, Kme2, Kme3, Rme2	53BP1	DNA damage response [24]
		TDRD3	Transcription activation [25]
MBT	Kme1, Kme2	L3MBTL1	Transcriptional repression [26, 27]
PWWP	Kme3	PSIP1	Transcriptional co-activation, DNA repair [28, 29]
Chromo	Kme, Kme2, Kme3	CHD1	Chromatin remodeling [30, 31]
		HP1	Heterochromatin [32]
		MRG15	Splicing [33]
Plant Agenet	Kme, Kme2, Kme3	FMRP	DNA damage response [34]
Bromodomain			
	KAc	BRD2/3	Transcriptional regulation [35]
14-3-3			
	Sph	14–3–3 ζ	Transcriptional activation [36]

^aSee Refs. [21–23] for more in-depth discussion.

Table 1. Overview of selected reader domains for post-translational modifications^a.

Next to histone tails, the nucleosome also possesses intrinsic docking platforms on its histone surface. The most prominent of these is composed of histones H2A and H2B. While the histone octamer is overall highly positively charged, there is a patch on the H2A-H2B dimer surface formed by acidic residues with negative surface charge. This structural feature is named the acidic patch and engages in a manifold of interactions with specific binding domains (**Figure 1**), including the tail of histone H4 of adjacent nucleosomes that promotes chromatin compaction. A common feature observed for acidic patch-interacting proteins is a positively charged arginine residue that interacts with a triad of acidic residues on H2A (Glu61, Asp90, Glu92). This is referred to as the arginine anchor [38]. It is often supported by surrounding positively charged residues interacting with acidic H2A/H2B interface residues.

Other parts of the histone core surface may also mediate protein-nucleosome interactions (**Figure 1C**). First, a solvent exposed cleft between H4 and H2B was shown to be involved in binding interactions with Sir3 or 53BP1 [39, 40]. Interestingly, these proteins bind simultaneously to both the H4-H2B cleft and the acidic patch using one nucleosome-binding domain for each epitope. Second, incorporation of non-canonical histones in nucleosomes introduces specific interaction surfaces that allow histone variant-specific nucleosome binding (**Figure 1D**). An example hereof are CENP-N and CENP-C that recognize the incorporated histone H3 variant CENP-A [41, 42].

Finally, the nucleosomal DNA is a major protein interaction site. First, it forms the binding site of linker histone H1 [43–45] (see also Section 4.9). Second, it is often involved in additional synergistic interactions to nucleosome-binding domains (**Figure 1E**). Finally, recent studies have identified transcription factor proteins that primarily bind to nucleosomal DNA. These so-called pioneer factors bind their DNA target sites while embedded in the nucleosome [46–48]. The structural details of these are however still lacking.

Throughout the advances in studies on nucleosome binding, it has become clear that binding of effector proteins in many cases involves interactions of nucleosome-binding domains to multiple nucleosome epitopes (**Figure 1G, H**). However, due to their size and complexity as well as the stability and dynamics of complex formation, the nucleosome is a challenging system for structural biology.

3. Crystal clear: lessons from crystallography and single particles

A key role in the research of protein interactions are high-resolution three-dimensional structures of the complexes, typically obtained by crystallography and, increasingly, cryo-electron microscopy. These structures enable the identification of binding sites and intermolecular interactions, offering a guided approach to design binding-deficient mutants or competitive binders. The history of nucleosome structural biology peaked with the publication of the high-resolution crystal structure of the nucleosome in 1997 [3]. Luger *et al.* achieved crystallization of the nucleosome together with a palindromic version of human α -satellite DNA [49]. This milestone study provided the foundation to also study the structures of nucleosomes together with chromatin factors in complexes. **Table 2** lists the structures of nucleosome-protein complexes

solved to date by crystallography and cryo-electron microscopy [39, 50–58]. The most recent addition to this ever-growing list is the spectacular structures of the complex between the INO80 chromatin remodeler and the nucleosome [59, 60]. Below, we discuss a few cardinal studies to highlight the different nucleosomal binding modes of effector proteins.

3.1. The first crystal structure of a nucleosome complex (LANA)

The first high-resolution structure of a nucleosome-protein complex was the crystal structure of a peptide model of Kaposi's sarcoma-associated herpesvirus LANA N-terminal region bound to the nucleosome [61]. The binding site identified in this study was the acidic patch. The atomistic resolution allowed to identify intermolecular side chain interactions including the arginine anchor bound to the acidic triad. Ever since, the LANA-nucleosome has become a golden standard for comparisons with other acidic patch interactions [50, 55]. Importantly, LANA is used to investigate the acidic patch binding ability of other proteins by competitive binding [62–64]. Interestingly, this exact epitope happened to be the binding interface also for the first full protein domain that was crystalized in its nucleosome-bound state.

Name	PDB-id	Role	Year	Technique	Reference	Resolution [Å]
Proteins						
RCC1	3MVD	Ran recruitment	2010	X-Ray	[50]	2.9
Sir3 BAH	4JJN, 3TU4, 4LD9, 4KUD	Chromatin compaction	2011, 2011, 2013, 2013	X-Ray	[39, 51–53]	3.0 – 3.3
CENP-C	4X23	H3 variant binding	2013	X-Ray	[54]	3.5
Ring1B	4R8P	E3 ligase	2014	X-Ray	[55]	3.3
53BP1	5KGF	Reader	2016	EM	[40]	4.5
SAGA/DUB	4ZUX	Eraser	2016	X-Ray	[56]	3.8
Set8	5HQ2	Writer	2016	X-Ray	[57]	4.5
Chd1	5O9G	Remodeler	2017	X-Ray	[74]	4.8
Snf2	5X0X, 5X0Y	Remodeler	2017	EM	[71]	4.0
CENP-N	6BUZ, 6C0W	H3 variant binder	2017, 2018	EM	[72, 73]	3.9/4.0
H1	4QLC, 5NL0	Linker histone	2015, 2017	X-Ray	[45, 75]	3.5
INO80	6FML, 6ETX	Remodeling complex	2018, 2018	EM	[59, 60]	4.4/4.8
Peptides						
LANA	1ZLA, 5GTC	Viral protein	2006, 2017	X-Ray	[61, 76]	2.9/2.7
IE1	5E5A	Viral protein	2016	X-Ray	[77]	2.8
GAG	5MLU	Synthetic acetylation system	2017	X-Ray	[78]	2.8

Table 2. Structures of nucleosome-protein or nucleosome-peptide complexes deposited in the RCSB protein databank PDB.

3.2. The first crystal structure of a nucleosome-bound protein domain (RCC1)

The first structure of a protein bound to the nucleosome was the RCC1-nucleosome complex published by the Tan lab in 2010. RCC1 (regulator of chromosome condensation) is essential during mitosis by recruiting Ran GTPase, which plays a role in nucleus reorganization, to the nucleosome [65, 66]. A comparison with LANA highlighted the crucial and conserved interaction of arginine residues with the acidic patch triad [50]. Strikingly, RCC1 binds to the acidic patch using the canonical arginine anchor, here contained in a loop, and also binds the nucleosomal DNA through its N-terminal tail. Such synergetic interactions have been observed later in many other nucleosome-binding proteins [50, 55, 67–70]. This study was the first to show such complexity of nucleosomes as interaction platforms. It also highlights the importance of properly defining the boundaries of binding domains to capture all binding epitopes in order to reveal possible synergetic interactions and fully understand complex formation and subsequent effects on chromatin structure.

3.3. Specificity of effector protein orientation in nucleosome complex formation (PRC1)

Besides determining the binding mode, synergetic interactions can also provide the structural basis for specificity of effector protein activity. This was shown in the crystal structure, also from the Tan lab, of the polycomb repressive complex 1 (PRC1) that ubiquitinates H2A K119 in a highly specific manner [55]. On its surface, the nucleosome displays various lysine residues that can be ubiquitinated by the respective writer proteins. However, the downstream response wildly differs depending on the position of the ubiquitinated lysine. Thus, target specificity is of high importance for ubiquitin writer proteins. In case of PRC1, this is based on two distinct binding processes. For one, there is the interaction between acidic patch and the arginine anchor of the Ring1B/Bmi1 subunit. In addition, the E2 subunit UbcH5c engages the nucleosomal DNA. Combined, both contributions are responsible for exact positioning of the catalytic center of the ubiquitin carrying E2 to the target H2A K119 (**Figure 2B**).

Besides LANA, RCC1 and PRC1, other crystal structures of nucleosome complexes offered further insights into nucleosome recognition. In particular, the structure of the nucleosome complex of the SAGA DUB deubiquitination module showed a non-canonical acidic patch binding. Morgan *et al.* found that the SAGA nucleosome-binding DUB module possesses three equally crucial arginine residues distributed over an α -helix [56] (**Figure 2A**). This perhaps points towards yet other acidic patch interaction modes.

Recently, also cryo-EM-derived structures of nucleosome-protein complexes have been published. The first structure, solved in 2016, yielded the structure of the complex with 53BP1, a reader protein for post-translational histone modifications [40]. Subsequently, the structures of Snf2 and CENP-N were solved and published [71–73].

Since the first crystal structure two decades ago, the list of nucleosome complexes deposited in the RCSB PDB protein databank is continuously growing. Still, the 12 high-resolution structures solved to date only encompass a fraction of all nucleosome-protein interactions. This discrepancy highlights the need for alternative techniques in chromatin structural biology.

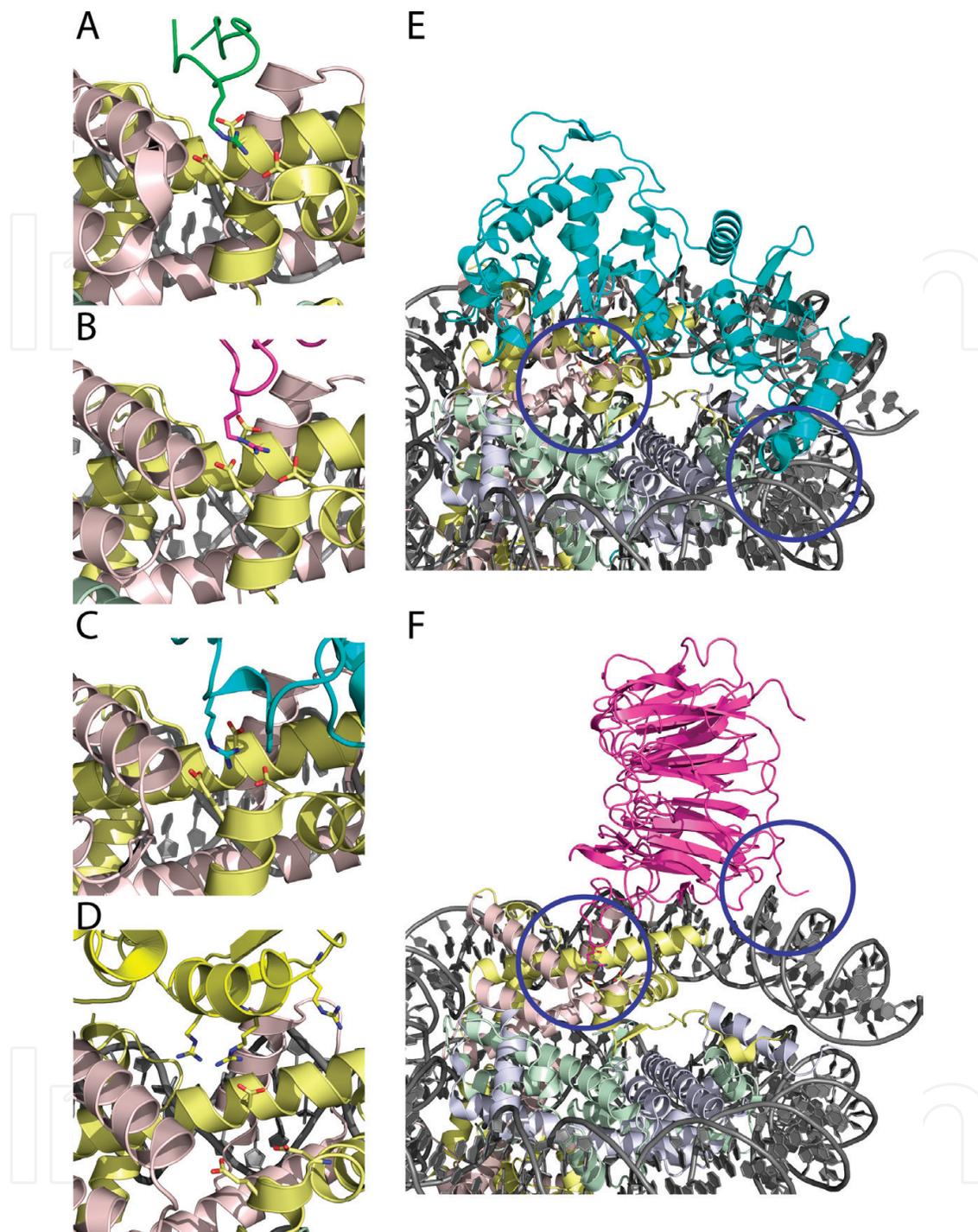


Figure 2. A detailed depiction of the acidic patch region and the triad of H2A that complexes the canonical arginine anchor. (A) LANA was the first peptide model of a binding domain to be crystalized (green, pdb: 1zla). (B) The first crystalized protein domain was RCC1 (pink, pdb: 3mvd). (C) They all share the acidic patch as a common binding epitope for the arginine anchor residue, as is also the case for the PRC1 complex with its acidic patch binding RING domain (cyan, 4r8p). (D) Interestingly, acidic patch binding is not necessarily limited to one single arginine anchor residue. As for the nucleosome-bound structure of the deubiquitinase complex SAGA-DUB (yellow, pdb: 4zux), three arginine residues are essential part of the acidic patch binding, of which none occupies the position in the center of the acidic triad. Combination of nucleosomal DNA and acidic patch binding is shown in the structure of RCC1 and PRC1. (E) PRC1 (cyan, pdb: 4r8p) that besides the acidic patch also engages DNA with its UbCh5c subunit. (F) The same holds true for RCC1 (pink, pdb: 3mvd) that contacts to DNA with the unstructured tail region.

4. Data-driven modeling

An attractive alternative to traditional structure determination methods is the modeling of structures of complexes based on some sort of experimental information on the interaction [79, 80]. In such data-driven modeling of a complex structure, the two interaction partners are docked together, guided by the experimental data, and respecting their biophysical properties. The exact binding interface and relative orientation of the binding partners are typically refined over several steps. Prerequisite for this approach is the availability of the 3D structures of the interacting partners. Several molecular docking programs allow the incorporation and use of experimental data and so increase the accuracy of resulting structures [81]. Hence, data from diverse biophysical techniques are translated into restraints guiding the docking process [82–84]. The type of information includes interaction interface, distances or shape of the complex and its subunits. Techniques that can provide these information are listed in **Table 3**.

Interestingly, all three classes of information can be provided by NMR spectroscopy. It is possible to gather data on intermolecular distances and shape by paramagnetic relaxation enhancement (PRE) and the nuclear Overhauser effect (NOE) as well as information on binding interfaces and binding affinity through chemical shift perturbation (CSP). The use of these NMR methods in docking studies is reviewed in detail elsewhere [79]. An overview of publications that used data-driven docking to investigate nucleosome-protein complexes is listed in **Table 4**.

4.1. Bringing data-driven modeling to nucleosome complexes (LSD1-CoREST)

A pioneer study for data-driven modeling of a nucleosome complex was successfully applied for the lysine-specific demethylase 1 and CoREST complex [86]. Both proteins cooperate in the demethylation of mono- and dimethylated H3K4. While it was possible to solve the crystal structure of LSD1-CoREST, their nucleosome-bound state remains elusive. Yang et al. gained insight into the molecular basis of LSD1-CoREST interaction by identifying point mutations that interfere with the LSD1-CoREST ability to demethylate a methylated peptide model of the histone H3 tail. Since it was previously shown that LSD1 recognizes a specific stretch of the H3 tail [94], it was possible to employ modeling to identify intermolecular interactions between the peptide and both the LSD1 active site and the LSD1-CoREST interface (**Figure 3B**). Lastly, NMR titration experiments of the CoREST SANT2 domain with DNA revealed a DNA-binding interface on SANT2. These pieces of interaction data were used to

Interaction interface	Distances	Shape
Mutagenesis	XL-MS	Cryo-EM
H/D exchange	Forster resonance energy transfer (FRET)	Small angle X-ray or neutron scattering (SAXS/SANS)
	Electron paramagnetic resonance (EPR)	Ion-mobility mass spectrometry (IM-MS)

Table 3. Biochemical and biophysical techniques for structural analysis of protein complexes.

guide a docking approach resulting in a complete structural model of the LSD1-CoREST-nucleosome complex (**Figure 3A**). With the lack of experimental data on the nucleosome interaction, this is a prime example of combining crystal structures, mutagenesis and NMR data to overcome limitations of the separate techniques.

Protein	Role	Data source	Reference
PSIP1-PWWP	Trimethyl lysine reader H3K36	NMR	[67, 68, 85]
CoREST/LSD1	Demethylase	Crystallography/NMR	[86]
Rad6-Bre1	Ubiquitin ligase	XL-MS	[70]
LANA	Viral protein	ssNMR	[87]
NSD1	Methyltransferase H3K36	Mutagenesis	[88]
RNF169	Ubiquitin reader	NMR, SAXS	[69, 89]
H1	Linker histone	NMR	[43, 90]
ISW2	Chromatin remodeler	XL-MS	[91]
Rad18	DNA repair factor	NMR	[89]
RCC1	Ran-recruitment	Crystallography	[62]
PHF1 Tudor	Trimethyl lysine reader H3K36	Crystallography/NMR	[92]
HMGN2	Chromatin decompaction	NMR	[93]

Table 4. Structural models of nucleosome-protein complexes based on biophysical data.

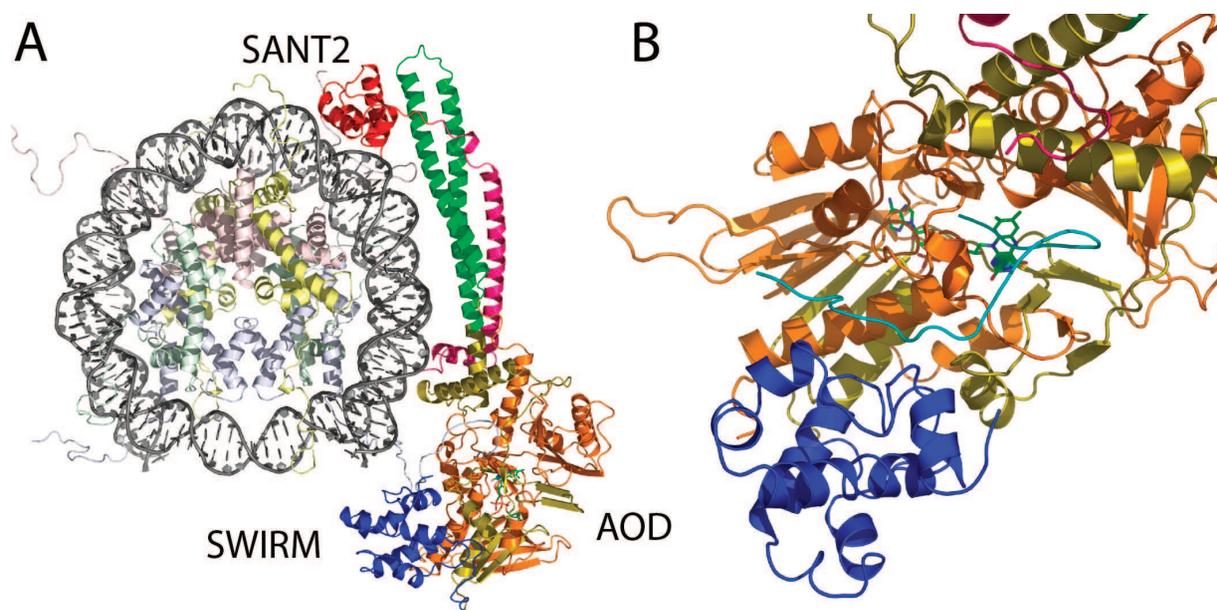


Figure 3. (A) Structural model of LSD1-CoREST bound to the nucleosome. The DNA binding of the SANT2 domain was elucidated by NMR spectroscopy. A previously identified binding motif in the H3 tail sequence was docked onto the interface of amine oxidase (AOD) and SWIRM domain revealing a second binding epitope. (B) The resulting model of the model peptide binding to AOD-SWIRM is shown as a close-up, highlighting how the tail is positioned on the interface of both domains. Figure generated using the author-provided PDB file [86].

4.2. NMR-based structural biology of nucleosome-protein complexes

Over recent years, several studies have demonstrated that state-of-the-art solution NMR can offer high-resolution and site-specific characterization of the structures and dynamics of nucleosome-protein complexes. NMR has the particular advantage of its sensitivity to dynamics and the ease with which interactions can be studied, allowing detailed insights into molecular recognition processes. NMR allows studies when systems are dynamic, or (partially) disordered, while this typically hampers high-resolution structure determination by crystallography and cryo-EM.

The molecular size of nucleosomes, and even more so of complexes with effector proteins, poses a challenge to traditional NMR methods. However, this challenge can be overcome through the use of methodologies designed for high-molecular weight systems. This method, methyl group-based transverse-relaxation-optimized spectroscopy (methyl-TROSY), relies on the highly sensitive observation of NMR signals of protein methyl groups [95]. Here, a specific isotope-labeling scheme is used, which typically results in observation of isoleucine, leucine, valine (ILV) methyl groups. The methyl-TROSY NMR spectra can subsequently be used to delineate binding sites of effector proteins on the nucleosome surface and vice versa [68, 69, 93, 96]. Extracting more detailed structural information is possible through the use of so-called spin-labels that can generate long-range distance restraints between the interaction partners [97, 98]. Whichever way used, NMR-based interaction data are of unique value in the modeling of nucleosome-protein complexes.

4.3. Expanding data sources for nucleosome complex models to NMR (HMGN2)

Kato *et al.* were the first to use the methyl-TROSY approach for the study of nucleosome-protein interactions [93]. Importantly, they reported the NMR signal assignments of the ILV-methyl groups for all histones in the nucleosomes. These assignments are essential in determining protein-binding sites on the nucleosome surface. The approach was demonstrated using high mobility group nucleosomal protein 2 (HMGN2), which regulates a variety of chromatin functions. HMGN2 was found to bind both the acidic patch and nucleosomal DNA. Based on these NMR data, supported by mutagenesis, it was possible to determine a structural model of the complex (**Figure 4A**). HMGN2 binds to the nucleosome as a staple, using two main interaction sites. On one side, HMGN2 is anchored to the acidic patch using a canonical arginine anchor in the N-terminal region of the binding domain, while the lysine-rich motif in its C-terminal region binds to nucleosomal DNA (**Figure 4B**). This binding mode provided a structural basis for the antagonistic function of HMGN2 towards linker histone H1 for nucleosome binding.

4.4. Latest applications of NMR to investigate structures of nucleosome complexes (RNF169 & Rad18)

Two recent studies relied on methyl-TROSY NMR-derived binding data to elucidate the recognition of ubiquitinated nucleosomes. Both focused on the interaction between ubiquitylated H2A K13/15 and the DNA repair factor RNF169. The work of Kitevski-LeBlanc *et al.*

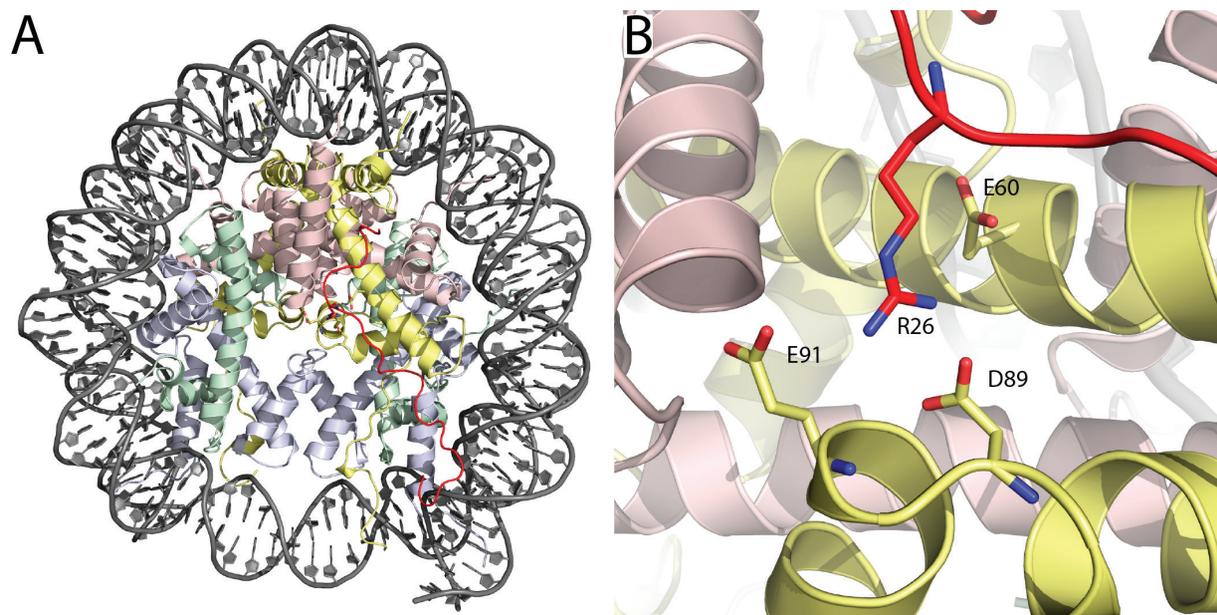


Figure 4. (A) Structural model of HMGN2 (red) bound to the nucleosome. The binding occurs along the nucleosome surface and is driven by interactions with the acidic patch and nucleosomal DNA, resulting in HMGN2 competing with H1 for nucleosome binding. (B) Close view on the acidic patch binding N-terminal HMGN2 region depicting the canonical arginine anchor R26 surrounded by the Glu 91, Asp 89, Glu 60 acidic triad motif of H2A. Figure generated using the author-provided PDB file [93].

established the molecular basis of this interaction. The α -helical MIU2 (motif interacting with ubiquitin) domain binds to a hydrophobic patch on the K13/15-conjugated ubiquitin while a disordered region anchors RNF169 on the nucleosome by binding to the acidic patch. They subsequently reconstructed a model structure that presents both epitopes in their nucleosome-bound state (**Figure 5A**). The work of Hu *et al.* combined traditional NOESY-based structure determination at the level of histone-dimers with interaction studies at the nucleosome level and complemented these with SAXS data into a final model [89]. The authors also extended their findings to an NMR-based structural model for the complex with DNA repair factor Rad18. Both RNF169 and Rad18 are known to interfere with the binding of 53BP1 to nucleosomes ubiquitinated at H2A K13/15. These NMR-based structural models have allowed to hypothesize on the molecular mechanism for this interference.

4.5. Importance of the nucleosomal context in epigenetic read-out (PSIP1-PWWP & PHF1-Tudor)

The complexity of nucleosome recognition by reader proteins is well illustrated by the NMR-based studies on the recognition of H3K36me-nucleosomes by the PWWP domain of PSIP1(Ledgf). NMR studies of this reader interaction found that the PWWP domain has binding affinity orders of magnitude lower for a H3K36me peptide compared to H3K36me₃ in a nucleosomal context. Interestingly, a similar observation was made for the Tudor domain of the H3K36me reader PHF1 [85]. Here, an isolated peptide model of the H3 tail showed decreased affinity as well. Due to the proximity of H3K36 to nucleosomal DNA, a role of DNA binding was hypothesized for both proteins. NMR studies showed for PSIP1 and PHF1 alike

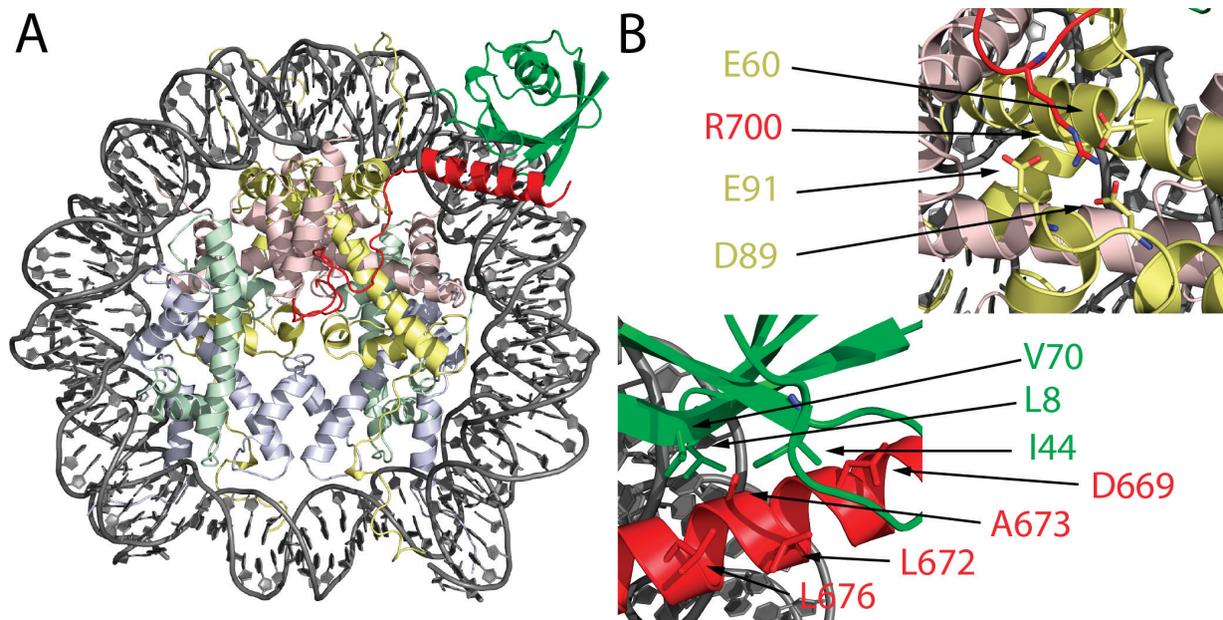


Figure 5. (A) Structural model of nucleosome-bound RNF169 (red) and ubiquitin (green). (B, top) The proposed main acidic patch anchoring residue R700 (conserved position throughout the docking solutions) is shown in the conserved arginine anchor position between the acidic triads (Glu 60, Asp 89, Glu 91). (B, bottom) Side chain interactions between RNF169 MIU2 (red) and ubiquitin (green). Figure generated using the author-provided PDB file [69].

a binding site for nucleosomal DNA, resulting in a simultaneous binding mechanism of both trimethyl lysine and nucleosomal DNA.

For PHF1-Tudor, a crystal structure bound to a trimethylated H3 tail peptide was already available to use. The additional importance of the nucleosomal context and synergetic binding mechanism can be understood from the corresponding nucleosome-bound structure (**Figure 6A**). In case of PSIP1-PWWP, the domain structure was solved by NMR and, together with NMR titration data, used to determine a structural model of nucleosome-bound protein (**Figure 6B**) [67, 68, 85]. The structural models of both highlighted the importance of the nucleosomal context in H3K36me3 recognition, emphasizing that complex formation critically depends on two synergetic binding processes. Firstly, the aromatic residues that form the aromatic cage bind to trimethylated lysine H3K36me3. This recognition of the PTM is crucial for the binding, but the readers reach their full binding affinity only when their positive surface residues interact with the nucleosomal DNA. This makes both studies outstanding examples of synergetic interplay of epitopes in nucleosome-binding proteins (**Figure 6C, D**).

The insights derived from these structural models were used to design experiments to validate the structural model and may offer possible tools for further research approaches. In case of PSIP1-PWWP, the structural model sparked current efforts in the design of nucleosome-mimicking peptides to modulate the PSIP1-chromatin interaction.

4.6. LANA goes solid state

The studies mentioned above illustrate the potential of data-driven modeling of nucleosome-protein complexes based on state-of the-art solution NMR. Recent advances in solid-state

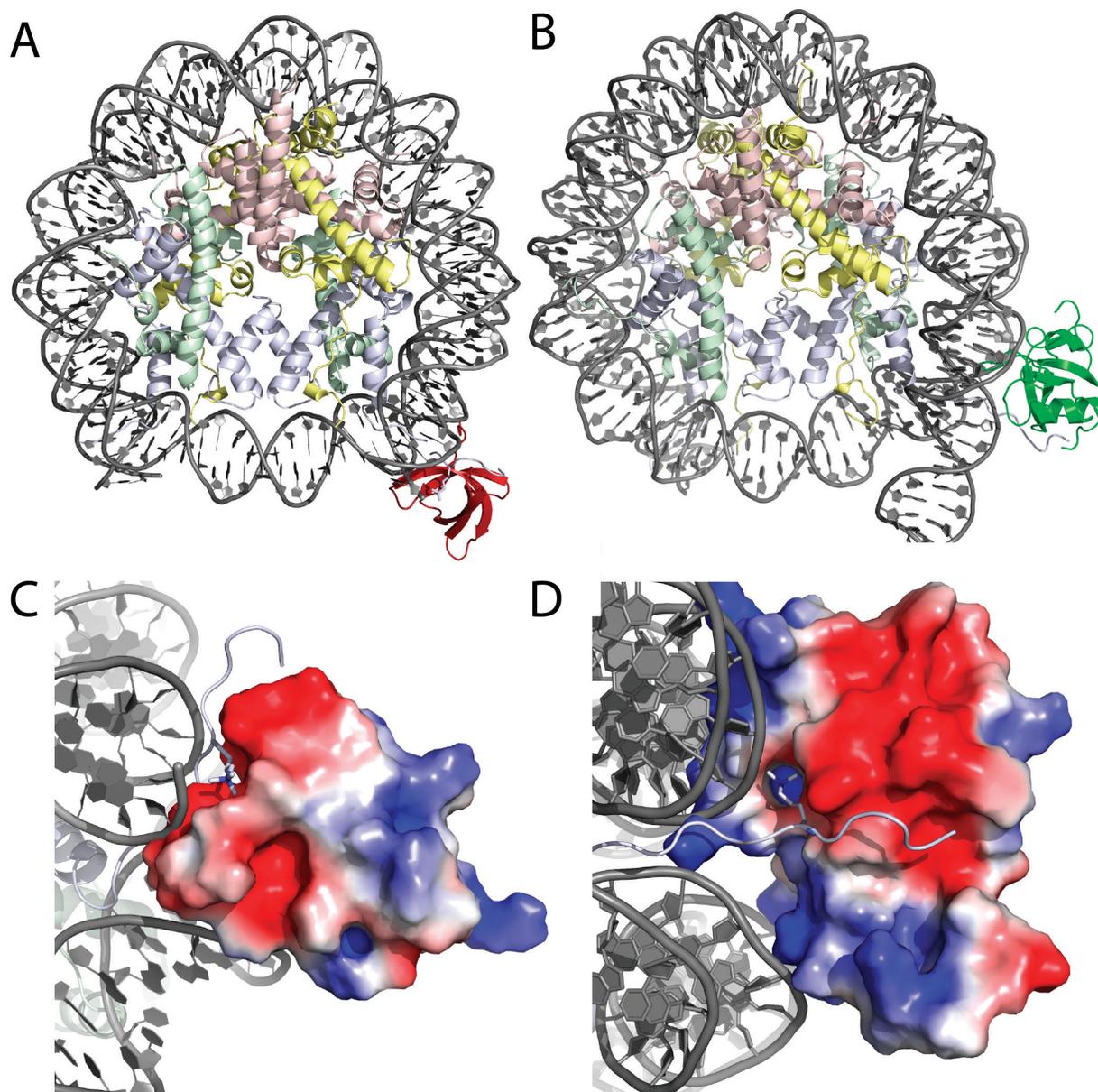


Figure 6. Structural model of nucleosome-bound PHF1 (red; A) and PSIP1-PWWP (green; B). The electrostatic potential of nucleosomal DNA and the surface of PHF1 (C) and PWWP (D), respectively, act in combination with H3K36me3 recognition by the aromatic cage motif (trimethyl lysine side chain shown as sticks). Figure generated using the author-provided PDB file [85].

NMR (ssNMR) have enabled the detailed investigation of large, soluble biomolecular complexes. Very recently, our lab capitalized on these advances and tailored them for application to nucleosome-protein complexes [87]. Unlike the methyl-TROSY methods, this approach allows observation of all residues, in principle allowing for a more complete mapping of binding interfaces. In this approach, NMR spectra are recorded on sediments, generated by ultracentrifugation, of nucleosomes or their complexes. After assignments of NMR signals of histone H2A in the unbound nucleosome, spectra were recorded on the nucleosome complex with the LANA peptide, analogous to the LANA crystal structure (**Figure 7A**) [61, 87]. Based on the chemical shift changes, the binding site of LANA could be mapped to the acidic patch

and a structural model generated. The large agreement between the crystal structure and ssNMR-derived structural model (**Figure 7B**) illustrates the power of this approach. In our view, ssNMR, just as the solution NMR approach, is an attractive alternative for structure determination for nucleosome-protein complexes. While its application awaits to be extended to larger nucleosome-binding domains, we anticipate that it will be a valuable addition to the tool kit in chromatin structural biology.

4.7. Modeling nucleosome-bound Rad6-Bre1 based on cross-linking MS

Next to NMR, cross-linking mass spectrometry has found increasing application as a data source on nucleosome-protein interactions. With cross-linking, intermolecular contacts between the proteins of interest are captured and converted to covalent connections. These connections are introduced by small molecule linkers, specific for the fusion of well-defined side chains or less specific as radical-forming photo cross-linkers. Furthermore, cross-linkers possess a spacer between their terminal functional groups to define the range of cross-linking ability [99, 100]. Both characteristics can be tuned for the study of a specific system, resulting in a manifold of reported linker molecules. After cross-linking, the protein complex undergoes trypsin digestion resulting in peptide fragments of the complex. Here, covalently cross-linked fragments stay connected. An analysis of these fragments by liquid chromatography mass spectrometry (LC-MS) enables identification of the sequence positions. The cross-links can thus be converted to distance restraints between two residues, with the distance depending on the length of the cross-linker. These restraints can be used to guide structural modelling of the complex [80]. In one of the earliest examples for nucleosome complexes, XL-MS was used to map the binding sites of the various nucleosome-binding domains of the chromatin

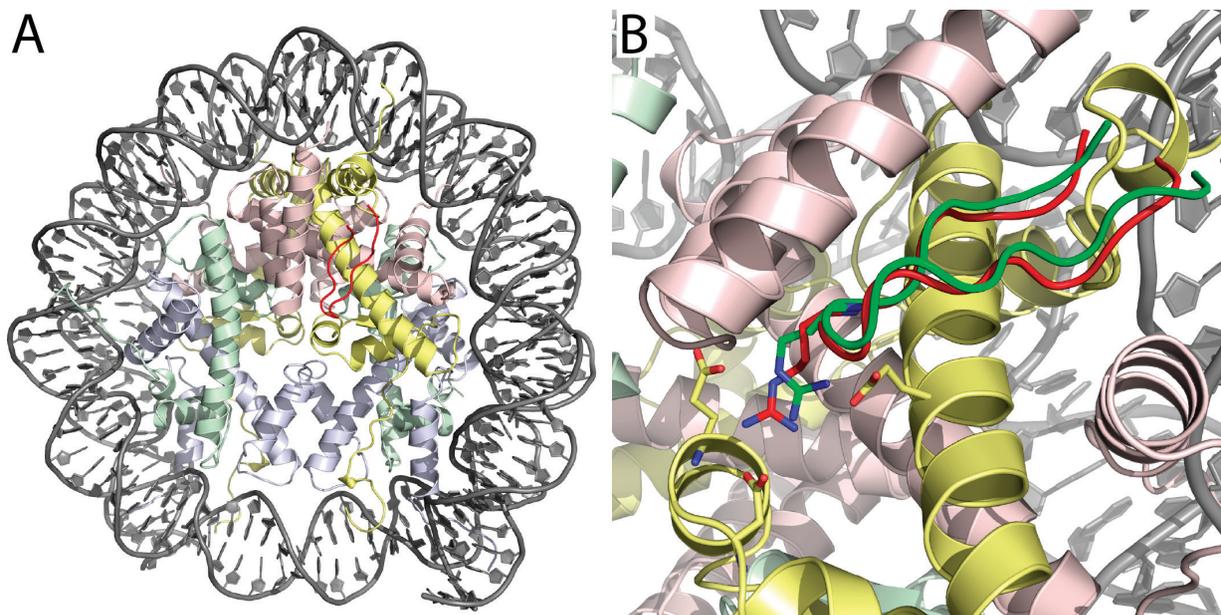


Figure 7. (A) Structural model for nucleosome-bound LANA peptide. ssNMR data derived from NMR titration experiments were used to direct the docking simulation. (B) Alignment of the ssNMR-derived model for LANA (red) and the crystal structure (green, pdb: 1zla) shows remarkable accuracy of the docking-derived solution. For both, the canonical arginine anchor is depicted as sticks in the typical central position between the acidic triad of H2A (yellow).

remodeling complex ISW2 onto the nucleosome surface [91]. These data were subsequently used to build a structural model of the ISW2-nucleosome complex. A recent case of cross-linking-based modeling in nucleosome research is the E2/E3 ubiquitin ligase complex Rad6-Bre1 (**Figure 8A**). Bre1 is known to act as a homodimer in a complex with Rad6 to specifically ubiquitinate H2B K123 [101, 102]. However, the molecular mechanism of specific ubiquitination remained unknown without any nucleosome-bound complex structure available. Gallego et al. addressed exactly this problem by using XL-MS data to identify the binding interface between the Bre1 RING domain and the nucleosome. Next to nucleosomal DNA binding, they observed binding of the homodimer to the acidic patch (**Figure 8B**), which was verified by LANA-induced inhibition of Bre1 RING nucleosome binding. As a first step in the modeling, the authors modeled the Rad6-Bre1 complex structures based on homology with known E2/E3 RING ligases. Importantly, the resulting model was supported by the observed cross-links. The Rad6-Bre1 model could then be docked onto the nucleosome guided by the observed cross-links. This provided the structural basis for the specificity of Bre1 towards H2B K123 ubiquitination [70].

4.8. Adding new perspective on binding modes

Data-driven structural models complement high-resolution structures in many ways. An interesting example is the RCC1-nucleosome interaction, which serves as binding platform for subsequent binding of Ran, a protein relevant during mitosis (see Section 3.2). Biochemical data have shown that Ran activity is increased in the nucleosome-bound complex. The crystal structure suggests no nucleosome-Ran interactions upon modeling Ran to the RCC1

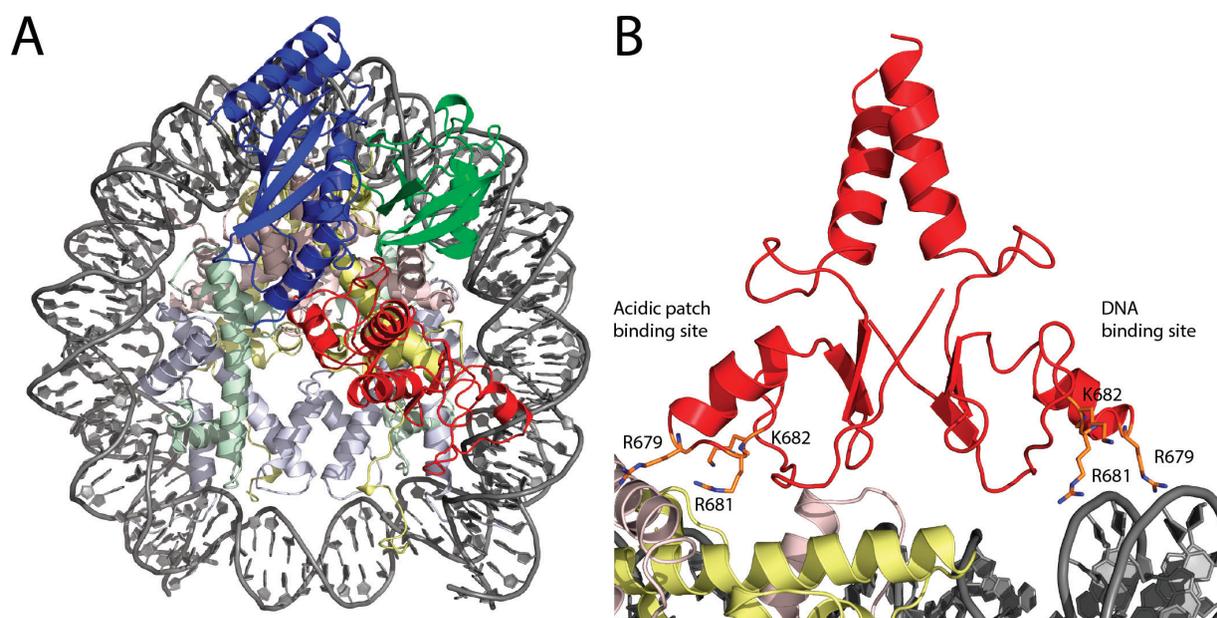


Figure 8. (A) Structural model of homodimeric Bre1 (red) bound to the nucleosome together with the E2 ligase Rad6 (blue) with attached ubiquitin (green). The study was conducted by identifying the interactions between positive Bre1 RING residues and the acidic patch. The docking was further facilitated due to the known target lysine residue. (B) Close view on Bre1 bound to both the acidic patch and nucleosomal DNA. The homodimeric nature allows the engagement of both epitopes in simultaneous binding. Figure generated using the author-provided PDB file [70].

Ran-binding interface. Before the crystal structure of nucleosome-bound RCC1 was solved, a data-driven model was reported, which does feature Ran-nucleosome interactions. [62]. The authors suggest that, upon Ran binding, the nucleosomal DNA contacts with RCC1 N-terminal tail observed in the crystal are broken in favor of Ran-nucleosome interactions as observed in model. Even though additional studies have to elucidate the exact mechanism of RCC1-Ran nucleosome binding, the use of crystal structure and data-driven model in combination outlines a possible mechanism to further investigate.

4.9. Debating H1

Another cardinal topic is the nucleosome-bound state of linker histone H1. To date, the structure of the chromatosome, consisting of the four canonical histones and 166bp of DNA in a complex with linker histones, is strongly debated. In this case as well, there are contradictions between structural models and a nucleosome-bound crystal structure of the chromatosome. The crystal structure reported by Zhou *et al.* displays the globular domain of linker histone H5 (chicken H5) with truncated tails in an on-dyad binding mode encountering both entering and leaving ends of linker DNA [75]. As for linker histone H1 (*X. laevis* H1.0b, human H1.5), a similar on-dyad binding mode was reported by cryo-EM and crystallography independently from absence or presence of H1 tails [45]. In fact, while not vital for linker histone positioning, the H1 C-terminal domain engages in binding of one of both linker DNAs preferably, introducing asymmetry into the nucleosome-bound complex.

In contrast to the proposed on-dyad complex, computational studies on linker histone binding suggest an alternative, off-dyad binding geometry of the complex in which the linker histone shows interactions with but one strand of linker DNA [103]. This binding mode was shown experimentally in the case of the globular domain of linker histone H1 (*D. melanogaster*). Here, NMR-based distance information, obtained through paramagnetic relaxation enhancement (PRE), was used to derive the nucleosome-binding mode of H1, showing an asymmetric, off-dyad binding [43]. Interestingly, it was shown by PRE as well that the mutation of a set of five crucial amino acids in H5 to its equivalents in H1 is sufficient to change the binding mode of H5 from on-dyad (crystal) to off-dyad [90]. This points out the importance of linker histone subtype sequence and the interacting residues in determining the binding mode towards the nucleosome [44].

5. Conclusions

Chromatin structural biology is an equally important as demanding field. This is not only clear from the tremendous efforts necessary for the first nucleosome structure but also from the limited number of structures for nucleosome-protein complexes. While crystallography and cryo-EM resulted in various high-resolution structures, not every interaction is accessible this way due to either of many experimental limitations, such as the need for crystallization, the fleeting nature of some complexes or the pervasive role of highly dynamic protein regions. Here, an increasing number of studies shift towards a combined approach utilizing

various sources of interaction data to direct sophisticated data-driven docking. This way all knowledge on a nucleosome-interacting system can be integrated into a structural model that is otherwise inaccessible. These models strongly depend on the quality and quantity of data and contain an inherent ambiguity. However, as in the case of linker histone H1, structural models can point to alternative binding modes and thus result in new, testable hypotheses. Additionally, crucial residues for nucleosome binding can be identified, allowing design of, for example, loss of function or loss of binding mutants to silence specific pathways. It also offers the possibility to drive the design of competing small molecule or peptide structures as potential candidates for epigenetic drugs interfering with specific effector binding. Remarkably, these developments might be otherwise lost due to the lack of a structure. However, as for now, a database for such structural models, akin to the RCSB protein databank, remains to be established. This might however be essential to advance the study of chromatin effector proteins. Publicly available structures including their data-based restraints could be used for further refinements upon availability of new, additional datasets from an array of techniques. It also would offer the possibility of negative results, otherwise rarely reported, to contribute to drive or score the quality of already reported models. Data-driven modeling of nucleosome-protein complexes has the potential to yield unique fundamental insights into nucleosome-binding dynamics and enable advances in modulation of chromatin effector proteins, which would be otherwise inaccessible.

Acknowledgements

We thank all authors of the studies included in this work who kindly provided us with files of their structural models for review. This work is supported by the Netherlands Organization for Scientific Research (NWO) through a VIDI grant (723.013.010) to Hugo van Ingen.

Conflict of interest

The authors of this work declare no conflict of interest.

Author details

Velten Horn^{1,2} and Hugo van Ingen^{2*}

*Address all correspondence to: h.vaningen@uu.nl

1 Department of Macromolecular Biochemistry, Leiden Institute of Chemistry, Leiden University, Leiden, The Netherlands

2 NMR Group, Bijvoet Center for Biomolecular Research, Utrecht University, Utrecht, The Netherlands

References

- [1] Kornberg RD. Chromatin structure: A repeating unit of histones and DNA. *Science*. 1974;**184**(4139):868-871
- [2] Thomas JO, Kornberg RD. An octamer of histones in chromatin and free in solution. *Proceedings of the National Academy of Sciences*. 1975;**72**(7):2626
- [3] Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997;**389**:251
- [4] Luger K, Dechassa ML, Tremethick DJ. New insights into nucleosome and chromatin structure: An ordered state or a disordered affair? *Nature Reviews. Molecular Cell Biology*. 2012;**13**:436
- [5] Sarah H, Karolina L, Nicole S, Meghan L, Sarah R, Sibaji S. Use of epigenetic drugs in disease: An overview. *Genetics & Epigenetics*. 2014;**6**:GEG.S12270
- [6] Gijbbers R, Vets S, De Rijck J, Ocwieja KE, Ronen K, Malani N, et al. Role of the PWWP domain of Lens epithelium-derived growth factor (LEDGF)/p75 cofactor in lentiviral integration targeting. *Journal of Biological Chemistry*. 2011;**286**(48):41812-41825
- [7] Yokoyama A, Cleary ML. Menin critically links MLL proteins with LEDGF on cancer-associated target genes. *Cancer Cell*. 2008;**14**(1):36-46
- [8] Egger G, Liang G, Aparicio A, Jones PA. Epigenetics in human disease and prospects for epigenetic therapy. *Nature*. 2004;**429**:457
- [9] Yu X, Li Z, Shen J. BRD7: A novel tumor suppressor gene in different cancers. *American Journal of Translational Research*. 2016;**8**(2):742-748
- [10] Pérez-Salvia M, Esteller M. Bromodomain inhibitors and cancer therapy: From structures to applications. *Epigenetics*. 2017;**12**(5):323-339
- [11] Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;**128**(4):693-705
- [12] Bannister AJ, Kouzarides T. Regulation of chromatin by histone modifications. *Cell Research*. 2011;**21**(3):381-395
- [13] Rothbart SB, Strahl BD. Interpreting the language of histone and DNA modifications. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2014;**1839**(8):627-643
- [14] Li Y, Sabari Benjamin R, Panchenko T, Wen H, Zhao D, Guan H, et al. Molecular coupling of histone crotonylation and active transcription by AF9 YEATS domain. *Molecular Cell*. 2016;**62**(2):181-193
- [15] Andrews FH, Shinsky SA, Shanle EK, Bridgers JB, Gest A, Tsun IK, et al. The Taf14 YEATS domain is a reader of histone crotonylation. *Nature Chemical Biology*. 2016;**12**:396
- [16] Zhao D, Guan H, Zhao S, Mi W, Wen H, Li Y, et al. YEATS2 is a selective histone crotonylation reader. *Cell Research*. 2016;**26**:629

- [17] Filippakopoulos P, Knapp S. Structural genomics and drug discovery for chromatin-related protein complexes involved in histone tail recognition. In: Emili A, Greenblatt J, Wodak S, editors. *Systems Analysis of Chromatin-Related Protein Complexes in Cancer*. New York, NY: Springer New York; 2014. pp. 211-225
- [18] Hughes RM, Wiggins KR, Khorasanizadeh S, Waters ML. Recognition of trimethyllysine by a chromodomain is not driven by the hydrophobic effect. *Proceedings of the National Academy of Sciences*. 2007;**104**(27):11184
- [19] Chen C, Nott TJ, Jin J, Pawson T. Deciphering arginine methylation: Tudor tells the tale. *Nature Reviews. Molecular Cell Biology*. 2011;**12**:629
- [20] Maurer-Stroh S, Dickens NJ, Hughes-Davies L, Kouzarides T, Eisenhaber F, Ponting CP. The Tudor domain 'Royal Family': Tudor, plant Agenet, chromo, PWWP and MBT domains. *Trends in Biochemical Sciences*. 2003;**28**(2):69-74
- [21] Teske KA, Hadden MK. Methyllysine binding domains: Structural insight and small molecule probe development. *European Journal of Medicinal Chemistry*. 2017;**136**:14-35
- [22] Yun M, Wu J, Workman JL, Li B. Readers of histone modifications. *Cell Research*. 2011;**21**(4):564-578
- [23] Barbieri I, Cannizzaro E, Dawson MA. Bromodomains as therapeutic targets in cancer. *Briefings in Functional Genomics*. 2013;**12**(3):219-230
- [24] Gupta A, Hunt CR, Chakraborty S, Pandita RK, Yordy J, Ramnarain DB, et al. Role of 53BP1 in the regulation of DNA double-Strand break repair pathway choice. *Radiation Research*. 2014;**181**(1):1-8
- [25] Yang Y, Lu Y, Espejo A, Wu J, Xu W, Liang S, et al. TDRD3 is an effector molecule for arginine-methylated histone Marks. *Molecular Cell*. 2010;**40**(6):1016-1023
- [26] Trojer P, Li G, Sims RJ, Vaquero A, Kalakonda N, Boccuni P, et al. L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell*. 2007;**129**(5):915-928
- [27] Boccuni P, MacGrogan D, Scandura JM, Nimer SD. The human L(3)MBT polycomb group protein is a transcriptional repressor and interacts physically and functionally with TEL (ETV6). *Journal of Biological Chemistry*. 2003;**278**(17):15412-15420
- [28] Daugaard M, Baude A, Fugger K, Povlsen LK, Beck H, Sørensen CS, et al. LEDGF (p75) promotes DNA-end resection and homologous recombination. *Nature Structural & Molecular Biology*. 2012;**19**:803
- [29] Ge H, Si Y, Roeder RG. Isolation of cDNAs encoding novel transcription coactivators p52 and p75 reveals an alternate regulatory mechanism of transcriptional activation. *The EMBO Journal*. 1998;**17**(22):6723-6729
- [30] Flanagan JF, Mi L-Z, Chruszcz M, Cymborowski M, Clines KL, Kim Y, et al. Double chromodomains cooperate to recognize the methylated histone H3 tail. *Nature*. 2005;**438**:1181

- [31] Sims RJ, Chen C-F, Santos-Rosa H, Kouzarides T, Patel SS, Reinberg D. Human but not yeast CHD1 binds directly and selectively to histone H3 methylated at lysine 4 via its tandem chromodomains. *Journal of Biological Chemistry*. 2005;**280**(51):41789-41792
- [32] Eissenberg JC, Elgin SCR. The HP1 protein family: Getting a grip on chromatin. *Current Opinion in Genetics & Development*. 2000;**10**(2):204-210
- [33] Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. Regulation of alternative splicing by histone modifications. *Science*. 2010;**327**(5968):996
- [34] Alpatov R, Lesch Bluma J, Nakamoto-Kinoshita M, Blanco A, Chen S, Stützer A, et al. A chromatin-dependent role of the fragile X mental retardation protein FMRP in the DNA damage response. *Cell*. 2014;**157**(4):869-881
- [35] LeRoy G, Rickards B, Flint SJ. The double bromodomain proteins Brd2 and Brd3 couple histone acetylation to transcription. *Molecular Cell*. 2008;**30**(1):51-60
- [36] Winter S, Simboeck E, Fischle W, Zupkovitz G, Dohnal I, Mechtler K, et al. 14-3-3 proteins recognize a histone code at histone H3 and are required for transcriptional activation. *The EMBO Journal*. 2008;**27**(1):88-99
- [37] Xu C, Bian C, Yang W, Galka M, Ouyang H, Chen C, et al. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). *Proceedings of the National Academy of Sciences*. 2010;**107**(45):19266
- [38] McGinty RK, Tan S. Recognition of the nucleosome by chromatin factors and enzymes. *Current Opinion in Structural Biology*. 2016;**37**:54-61
- [39] Armache K-J, Garlick JD, Canzio D, Narlikar GJ, Kingston RE. Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science (New York, NY)*. 2011;**334**(6058):977-982
- [40] Wilson MD, Benlekbir S, Fradet-Turcotte A, Sherker A, Julien J-P, McEwan A, et al. The structural basis of modified nucleosome recognition by 53BP1. *Nature*. 2016;**536**:100
- [41] Carroll CW, Milks KJ, Straight AF. Dual recognition of CENP-A nucleosomes is required for centromere assembly. *The Journal of Cell Biology*. 2010;**189**(7):1143
- [42] Carroll CW, Silva MC, Godek KM, Jansen LE, Straight AF. Centromere assembly requires the direct recognition of CENP-A nucleosomes by CENP-N. *Nature Cell Biology*. 2009;**11**(7):896-902
- [43] Zhou BR, Feng H, Kato H, Dai L, Yang Y, Zhou Y, et al. Structural insights into the histone H1-nucleosome complex. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;**110**(48):19390-19395
- [44] Fyodorov DV, Zhou B-R, Skoultchi AI, Bai Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nature Reviews. Molecular Cell Biology*. 2017;**19**:192

- [45] Bednar J, Garcia-Saez I, Boopathi R, Cutter AR, Papai G, Reymer A, et al. Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. *Molecular Cell*. 2017;**66**(3):384-397. e8
- [46] Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M, Zaret KS. Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Molecular Cell*. 2002;**9**(2):279-289
- [47] Soufi A, Donahue G, Zaret Kenneth S. Facilitators and impediments of the pluripotency reprogramming factors' initial engagement with the genome. *Cell*. 2012;**151**(5):994-1004
- [48] Soufi A, Garcia Meilin F, Jaroszewicz A, Osman N, Pellegrini M, Zaret Kenneth S. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell*. 2015;**161**(3):555-568
- [49] Harp JM, Uberbacher EC, Roberson AE, Palmer EL, Gewiess A, Bunick GJ. X-ray diffraction analysis of crystals containing twofold symmetric nucleosome core particles. *Acta Crystallographica Section D*. 1996;**52**(2):283-288
- [50] Makde RD, England JR, Yennawar HP, Tan S. Structure of RCC1 chromatin factor bound to the nucleosome core particle. *Nature*. 2010;**467**(7315):562-566
- [51] Wang F, Li G, Altaf M, Lu C, Currie MA, Johnson A, et al. Heterochromatin protein Sir3 induces contacts between the amino terminus of histone H4 and nucleosomal DNA. *Proceedings of the National Academy of Sciences of the United States of America*. 2013;**110**(21):8495-8500
- [52] Arnaudo N, Fernandez IS, McLaughlin SH, Peak-Chew SY, Rhodes D, Martino F. The N-terminal acetylation of Sir3 stabilizes its binding to the nucleosome core particle. *Nature Structural & Molecular Biology*. 2013;**20**(9):1119-1121
- [53] Yang D, Fang Q, Wang M, Ren R, Wang H, He M, et al. N-alpha-acetylated Sir3 stabilizes the conformation of a nucleosome-binding loop in the BAH domain. *Nature Structural & Molecular Biology*. 2013;**20**(9):1116-1118
- [54] Kato H, Jiang J, Zhou B-R, Rozendaal M, Feng H, Ghirlando R, et al. A conserved mechanism for centromeric nucleosome recognition by centromere protein CENP-C. *Science (New York, N.Y.)*. 2013;**340**(6136):1110-1113
- [55] McGinty RK, Henrici RC, Tan S. Crystal structure of the PRC1 ubiquitylation module bound to the nucleosome. *Nature*. 2014;**514**(7524):591-596
- [56] Morgan MT, Haj-Yahya M, Ringel AE, Bandi P, Brik A, Wolberger C. Structural basis for histone H2B deubiquitination by the SAGA DUB module. *Science (New York, N.Y.)*. 2016;**351**(6274):725-728
- [57] Girish TS, McGinty RK, Tan S. Multivalent interactions by the Set8 histone methyltransferase with its nucleosome substrate. *Journal of Molecular Biology*. 2016;**428**(8):1531-1543

- [58] Pinotsis N, Waksman G. Crystal structure of the legionella pneumophila Lpg2936 in complex with the cofactor S-adenosyl-L-methionine reveals novel insights into the mechanism of RsmE family methyltransferases. *Protein Science: A Publication of the Protein Society*. 2017;**26**(12):2381-2391
- [59] Eustermann S, Schall K, Kostrewa D, Lakomek K, Strauss M, Moldt M, et al. Structural basis for ATP-dependent chromatin remodelling by the INO80 complex. *Nature*. 2018;**556**(7701):386-390
- [60] Ayala R, Willhoft O, Aramayo RJ, Wilkinson M, McCormack EA, Ocloo L, et al. Structure and regulation of the human INO80–nucleosome complex. *Nature*. 2018;**556**(7701):391-395
- [61] Barbera AJ, Chodaparambil JV, Kelley-Clarke B, Joukov V, Walter JC, Luger K, et al. The nucleosomal surface as a docking station for Kaposi's sarcoma herpesvirus LANA. *Science*. 2006;**311**(5762):856
- [62] England JR, Huang J, Jennings MJ, Makde RD, Tan S. RCC1 uses a conformationally diverse loop region to interact with the nucleosome: A model for the RCC1-nucleosome complex. *Journal of Molecular Biology*. 2010;**398**(4):518-529
- [63] Kan P-Y, Caterino TL, Hayes JJ. The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Molecular and Cellular Biology*. 2009;**29**(2):538-546
- [64] Mattioli F, Uckelmann M, Sahtoe DD, van Dijk WJ, Sixma TK. The nucleosome acidic patch plays a critical role in RNF168-dependent ubiquitination of histone H2A. *Nature Communications*. 2014;**5**:3291
- [65] Clarke PR, Zhang C. Spatial and temporal coordination of mitosis by ran GTPase. *Nature Reviews. Molecular Cell Biology*. 2008;**9**:464
- [66] Nishitani H, Ohtsubo M, Yamashita K, Iida H, Pines J, Yasudo H, et al. Loss of RCC1, a nuclear DNA-binding protein, uncouples the completion of DNA replication from the activation of cdc2 protein kinase and mitosis. *The EMBO Journal*. 1991;**10**(6):1555-1564
- [67] Eidahl JO, Crowe BL, North JA, McKee CJ, Shkriabai N, Feng L, et al. Structural basis for high-affinity binding of LEDGF PWWP to mononucleosomes. *Nucleic Acids Research*. 2013;**41**(6):3924-3936
- [68] van Nuland R, van Schaik FM, Simonis M, van Heesch S, Cuppen E, Boelens R, et al. Nucleosomal DNA binding drives the recognition of H3K36-methylated nucleosomes by the PSIP1-PWWP domain. *Epigenetics & Chromatin*. 2013;**6**(1):12
- [69] Kitevski-LeBlanc J, Fradet-Turcotte A, Kukic P, Wilson MD, Portella G, Yuwen T, et al. The RNF168 paralog RNF169 defines a new class of ubiquitylated histone reader involved in the response to DNA damage. *eLife*. 2017;**6**
- [70] Gallego LD, Ghodgaonkar Steger M, Polyansky AA, Schubert T, Zagrovic B, Zheng N, et al. Structural mechanism for the recognition and ubiquitination of a single nucleosome

- residue by Rad6–Bre1. *Proceedings of the National Academy of Sciences of the United States of America*. 2016;**113**(38):10553-10558
- [71] Liu X, Li M, Xia X, Li X, Chen Z. Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. *Nature*. 2017;**544**(7651):440-445
- [72] Chittori S, Hong J, Saunders H, Feng H, Ghirlando R, Kelly AE, et al. Structural mechanisms of centromeric nucleosome recognition by the kinetochore protein CENP-N. *Science*. 2018;**359**(6373):339
- [73] Pentakota S, Zhou K, Smith C, Maffini S, Petrovic A, Morgan GP, et al. Decoding the centromeric nucleosome through CENP-N. *eLife*. 2017;**6**
- [74] Farnung L, Vos SM, Wigge C, Cramer P. Nucleosome-Chd1 structure and implications for chromatin remodelling. *Nature*. 2017;**550**(7677):539-542
- [75] Zhou B-R, Jiang J, Feng H, Ghirlando R, Xiao TS, Bai Y. Structural mechanisms of nucleosome recognition by linker histones. *Molecular Cell*. 2015;**59**(4):628-638
- [76] Amamoto Y, Aoi Y, Nagashima N, Suto H, Yoshidome D, Arimura Y, et al. Synthetic post-translational modifications: Chemical catalyst-driven regioselective histone acylation of native chromatin. *Journal of the American Chemical Society*. 2017;**139**(22):7568-7576
- [77] Fang Q, Chen P, Wang M, Fang J, Yang N, Li G, et al. Human cytomegalovirus IE1 protein alters the higher-order chromatin structure by targeting the acidic patch of the nucleosome. *eLife*. 2016;**5**:e11911
- [78] Lesbats P, Serrao E, Maskell DP, Pye VE, O'Reilly N, Lindemann D, et al. Structural basis for spumavirus GAG tethering to chromatin. *Proceedings of the National Academy of Sciences*. 2017;**114**(21):5509
- [79] van Ingen H, Bonvin AMJJ. Information-driven modeling of large macromolecular assemblies using NMR data. *Journal of Magnetic Resonance*. 2014;**241**:103-114
- [80] Rappsilber J. The beginning of a beautiful friendship: Cross-linking/mass spectrometry and modelling of proteins and multi-protein complexes. *Journal of Structural Biology*. 2011;**173**(3):530-540
- [81] Xue LC, Dobbs D, Bonvin AMJJ, Honavar V. Computational prediction of protein interfaces: A review of data driven methods. *FEBS Letters*. 2015;**589**(23):3516-3526
- [82] Clore GM, Schwieters CD. Docking of protein–protein complexes on the basis of highly ambiguous intermolecular distance restraints derived from ¹HN/¹⁵N chemical shift mapping and backbone ¹⁵N–¹H residual dipolar couplings using conjoined rigid body/torsion angle dynamics. *Journal of the American Chemical Society*. 2003;**125**(10):2902-2912
- [83] Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A protein–protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*. 2003;**125**(7):1731-1737

- [84] van Dijk Aalt DJ, Boelens R, Bonvin Alexandre MJJ. Data-driven docking for the study of biomolecular complexes. *The FEBS Journal*. 2004;**272**(2):293-312
- [85] Musselman CA, Gibson MD, Hartwick EW, North JA, Gatchalian J, Poirier MG, et al. Binding of PHF1 Tudor to H3K36me3 enhances nucleosome accessibility. *Nature Communications*. 2013;**4**:2969
- [86] Yang M, Gocke CB, Luo X, Borek D, Tomchick DR, Machius M, et al. Structural basis for CoREST-dependent demethylation of nucleosomes by the human LSD1 histone demethylase. *Molecular Cell*. 2006;**23**(3):377-387
- [87] Xiang S, Le Paige UB, Horn V, Houben K, Baldus M, van Ingen H. Site-specific studies of nucleosome interactions by solid-state NMR. *Angewandte Chemie International Edition*. 2018;**57**(17):4571-4575
- [88] Qiao Q, Li Y, Chen Z, Wang M, Reinberg D, Xu RM. The structure of NSD1 reveals an autoregulatory mechanism underlying histone H3K36 methylation. *The Journal of Biological Chemistry*. 2011;**286**(10):8361-8368
- [89] Hu Q, Botuyan MV, Cui G, Zhao D, Mer G. Mechanisms of ubiquitin-nucleosome recognition and regulation of 53BP1 chromatin recruitment by RNF168/169 and RAD18. *Molecular Cell*. 2017;**66**(4):473-487. e9
- [90] Zhou BR, Feng H, Ghirlando R, Li S, Schwieters CD, Bai Y. A small number of residues can determine if linker histones are bound on or off dyad in the chromatosome. *Journal of Molecular Biology*. 2016;**428**(20):3948-3959
- [91] Dang W, Bartholomew B. Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. *Molecular and Cellular Biology*. 2007;**27**(23):8306-8317
- [92] Musselman CA, Avvakumov N, Watanabe R, Abraham CG, Lalonde M-E, Hong Z, et al. Molecular basis for H3K36me3 recognition by the Tudor domain of PHF1. *Nature Structural & Molecular Biology*. 2012;**19**:1266
- [93] Kato H, van Ingen H, Zhou BR, Feng H, Bustin M, Kay LE, et al. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proceedings of the National Academy of Sciences of the United States of America*. 2011;**108**(30):12283-12288
- [94] Forneris F, Binda C, Vanoni MA, Battaglioli E, Mattevi A. Human histone demethylase LSD1 reads the histone code. *Journal of Biological Chemistry*. 2005;**280**(50):41360-41365
- [95] Ollerenshaw JE, Tugarinov V, Kay LE. Methyl TROSY: Explanation and experimental verification. *Magnetic Resonance in Chemistry*. 2003;**41**(10):843-852
- [96] Miller TCR, Simon B, Rybin V, Grötsch H, Curtet S, Khochbin S, et al. A bromodomain-DNA interaction facilitates acetylation-dependent bivalent nucleosome recognition by the BET protein BRDT. *Nature Communications*. 2016;**7**:13855

- [97] Hass MAS, Ubbink M. Structure determination of protein–protein complexes with long-range anisotropic paramagnetic NMR restraints. *Current Opinion in Structural Biology*. 2014;**24**:45-53
- [98] Schilder J, Liu WM, Kumar P, Overhand M, Huber M, Ubbink M. Protein docking using an ensemble of spin labels optimized by intra-molecular paramagnetic relaxation enhancement. *Physical Chemistry Chemical Physics*. 2016;**18**(8):5729-5742
- [99] Back JW, de Jong L, Muijsers AO, de Koster CG. Chemical cross-linking and mass spectrometry for protein structural modeling. *Journal of Molecular Biology*. 2003;**331**(2):303-313
- [100] Sinz A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *Journal of Mass Spectrometry*. 2003;**38**(12):1225-1237
- [101] Turco E, Gallego LD, Schneider M, Köhler A. Monoubiquitination of histone H2B is intrinsic to the bre1 RING domain-Rad6 interaction and augmented by a second Rad6-binding site on Bre1. *Journal of Biological Chemistry*. 2015;**290**(9):5298-5310
- [102] Kumar P, Wolberger C. Structure of the yeast Bre1 RING domain. *Proteins*. 2015;**83**(6):1185-1190
- [103] Pachov GV, Gabdoulline RR, Wade RC. On the structure and dynamics of the complex of the nucleosome and the linker histone. *Nucleic Acids Research*. 2011;**39**(12):5255-5263