Review

Characterization of functional disordered regions within chromatin-associated proteins

Catherine A. Musselman[1,*] and Tatiana G. Kutateladze[2]

SUMMARY

Intrinsically disordered regions (IDRs) are abundant and play important roles in the function of chromatin-associated proteins (CAPs). These regions are often found at the N- and C-termini of CAPs and between structured domains, where they can act as more than just linkers, directly contributing to function. IDRs have been shown to contribute to substrate binding, act as auto-regulatory regions, and drive liquid-liquid droplet formation. Their disordered nature provides increased functional diversity and allows them to be easily regulated through post-translational modification. However, these regions can be especially challenging to characterize on a structural level. Here, we review the prevalence of IDRs in CAPs, highlighting several studies that address their importance in CAP function and show progress in structural characterization of these regions. A focus is placed on the unique opportunity to apply nuclear magnetic resonance (NMR) spectroscopy alongside cryo-electron microscopy to characterize IDRs in CAPs.

INTRINSIC DISORDER

The function of proteins has long been ascribed to their three-dimensional structure. Indeed, for many such proteins, fold is critical for proper activity. However, the eukaryotic genome encodes for a large number of intrinsically disordered proteins (IDPs), as well as intrinsically disordered regions (IDRs) within otherwise folded proteins (Dunker and Obradovic, 2001; Dunker et al., 2002) (Figure 1). IDPs/IDRs are characterized by low sequence complexity (Romero et al., 2001) and are typically enriched in hydrophilic amino acids which lead to a lack of secondary structure and stabilization of an unfolded state (Altman et al., 2000; Tompa, 2002). In recent years, many of these IDPs and IDRs have been identified as functionally important in various cellular processes (Dunker et al., 2002; Tantos et al., 2011; Dyson, 2011; Uversky, 2017). They have been shown to directly drive critical intermolecular interactions, including protein-protein and protein-nucleic acid interactions. They have also been found to be involved in the formation of phase-separated condensates. Recent remarkable advances in this field have led to a revisiting of the traditional structure-function paradigm (Wright and Dyson, 1999), with the disordered state now recognized to play fundamental roles in a broad range of cellular processes.

Compared to more rigid structured elements, unstructured elements are quite malleable which enables multiple interactions not achievable by a structured domain (Tompa, 2002). In addition, IDPs/IDRs undergo substantially greater post-translational modification in comparison with structured elements as they are enriched in commonly modified amino acids and are easily accessible to modifying enzymes (Bah and Forman-Kay, 2016). Together, these abilities of IDPs/IDRs lead to high functional diversity and ease of regulation that is unique as compared to structured elements.

Upon association with a binding partner, some IDPs/IDRs undergo an unstructured to structured transition. This transition can involve folding into secondary structure elements, or alternatively, association can stabilize a single conformation without inducing secondary structure (Figure 2, left and center). The mechanism of such transitions is not fully understood and likely varies between IDPs/IDRs. The two limiting cases are conformational selection and induced fit, with a broad spectrum between them (Mollica et al., 2016). In conformational selection, the IDP/IDR samples the bound state within its unbound conformational ensemble, and the binding partner recognizes and stabilizes this state. In fact, the unbound states of

[1]Biochemistry and Molecular Genetics, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA

[2]Department of Pharmacology, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA
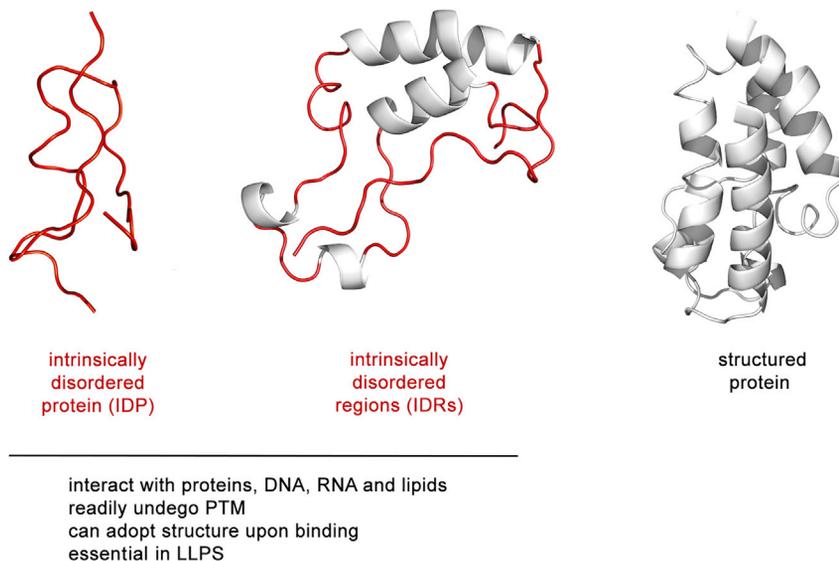
*Correspondence: catherine.musselman@cuanschutz.edu

https://doi.org/10.1016/j.isci.2021.102070

Figure 1. Architecture of disordered and structured proteins
Proteins can adopt a variety of conformations from fully disordered to fully structured to mixed structured/disordered. Intrinsically disordered proteins (IDPs, left) do not adopt any globular structure. Mixed proteins contain intrinsically disordered regions (IDRs, middle) at the N- or C-terminus or between structured regions. Fully structured proteins (right) adopt a unique globular fold, lacking IDRs. Structured regions are shown in gray and intrinsically disordered regions are shown in red.

IDPs/IDRs are predicted to not be fully random, rather having conformational bias that would encourage conformational selection (Fuxreiter et al., 2004). On the other end of the spectrum, induced fit suggests that contacts with the partner are required to enable the IDP/IDR to adopt the bound state conformation.

In contrast, some IDPs/IDRs retain conformational disorder even in the bound state. This can arise from the existence of multiple binding sites on the IDP/IDR for a single site on the partner (allovalency), multiple but distinct binding sites on both the IDP/IDR and partner (avidity), or multiple non-distinct binding sites on both the IDP/IDR and partner. The latter has been termed a fuzzy complex (Figure 2, right) and refers to the extreme disorder of an IDP/IDR in the bound state that arises when both the IDP/IDR and partner have multiple exchangeable interaction sites (Olsen et al., 2017; Tompa and Fuxreiter, 2008; Sharma et al., 2015; Fuxreiter, 2018). Recently, it has been demonstrated that despite their disorder, even fuzzy complexes can be formed via a very high affinity association (Borgia et al., 2018; Turner et al., 2018). It has been suggested that conformational disorder in the bound state might be favorable because it allows for easy regulation of the interaction (Tompa and Fuxreiter, 2008). Such regulation could be achieved through post-translational modification or quick exchange between partners, as partial dissociation can facilitate transfer.

STRUCTURAL CHARACTERIZATION OF IDPS/IDRS

Given the emerging functional significance of IDPs/IDRs, it is of great interest to understand their structural properties, molecular basis of binding, and effect of post-translational modification on both. However, the conformational heterogeneity of these regions causes a substantial challenge in their characterization. To date, the most common method used for protein structure determination is X-ray crystallography. However, the conformational heterogeneity of IDPs precludes the ability to crystallize them. Furthermore, the presence of IDRs can sometimes impede crystallization of macromolecules even if they do contain structured elements. As a result, IDRs are often deleted to facilitate crystallization. Alternatively, if systems can be crystallized with IDRs intact, these regions often do not resolve in the final structural model.

Recently, cryo-electron microscopy (cryo-EM), in particular single-particle cryo-EM, has emerged as a very promising approach to investigate the structure of complexes containing compositional and/or conformational heterogeneity (Cheng, 2018). Here, the target does not need to be crystallized, rather it is frozen in a thin-layer hydrated state on a grid. The images from many molecules are then computationally combined
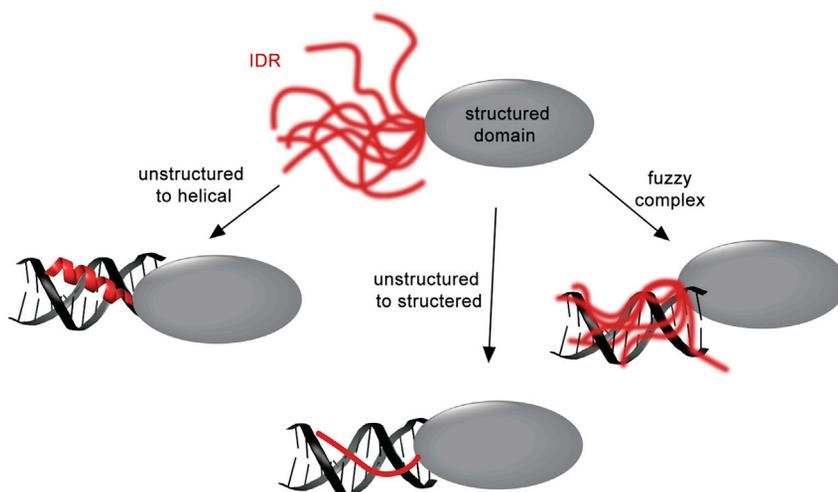
Figure 2. IDR conformational transitions upon ligand binding

Shown is a globular domain (gray) flanked by an IDR (red). IDRs sample a broad conformational ensemble in the unbound state, with fast exchange between states (represented by a blur). Upon binding, IDRs have been observed to adopt secondary structure such as an alpha helix (left), adopt a stable conformation without folding into secondary structure (middle), or maintain substantial conformational dynamics even in the bound state, known as a fuzzy complex (right).

to reconstruct the 3D structure. Importantly, moderate heterogeneity in both composition and conformation can be accommodated as this can be computationally managed. Thus, the presence of IDRs does not preclude the ability to generate a structural model. In addition, no labeling and very little sample is required, opening the avenue to studying native complexes and regulation by the presence of native post-translatiponal modifications (PTMs). The problem remains, however, that the regions of substantial conformational heterogeneity will still resolve poorly in the final structural model. Thus, while this approach can allow for positioning of the IDR in the complex, it often cannot yield detailed structural analysis.

In contrast to X-ray crystallography and cryo-EM, solution nuclear magnetic resonance (NMR) spectroscopy is highly suited for investigating IDPs/IDRs at amino acid-level resolution (Gibbs et al., 2017; Brutscher et al., 2015). Solution NMR provides an equilibrium averaged readout that does not require conformational homogeneity and readily resolves conformationally dynamic regions. Thus, NMR is amenable to investigating proteins that sample a broad conformational space. Various experiments allow for investigation of structural properties, including the identification of transient secondary structure and assessment of the magnitude and timescale of conformational dynamics (Gibbs et al., 2017; Jensen et al., 2014). In addition, the effects of PTMs and mechanisms of binding can straightforwardly be studied (Theillet et al., 2012).

This is not to say that there are no challenges in using NMR to study IDPs/IDRs. In contrast to single-particle cryo-EM, NMR requires isotope labeling and high concentrations of sample, generally requiring overexpression in *E. coli*. Compared to structured proteins, there is substantial lack of dispersion of resonances in NMR spectra recorded using standard biomolecular approaches. Although the often sharp linewidth partially overcomes this issue, it can sometimes preclude assignment of the resonances to particular amino acids. This in turn limits the capability to obtain amino acid-level information that makes NMR so powerful. To help overcome this problem, a variety of specialized NMR experimental approaches have been developed, including ^{13}C direct detect approaches (Sahu et al., 2014). Finally, while NMR spectroscopy is a fantastic option for investigating IDP/IDR structure and dynamics at high resolution, this method is size limited, and thus, IDRs are often taken out of a larger context.

Thus, it is becoming increasingly clear that a more powerful approach would be to combine NMR spectroscopy and cryo-EM to investigate IDRs (Geraets et al., 2020). Higher resolution information about the conformation and substrate binding can be garnered from NMR, while cryo-EM can provide a framework for modeling this into the larger context.

IDRS IN CHROMATIN-ASSOCIATED PROTEINS

IDPs/IDRs are especially enriched in the nuclear proteome and are particularly abundant in chromatin and chromatin-associated proteins (CAPs). Perhaps the best characterized functional IDRs within chromatin are the core histone tails of histone proteins H2A, H2B, H3, and H4. These N- and C-termini of the core histones are heavily post-translationally modified (Rothbart and Strahl, 2014; Zentner and Henikoff, 2013). These modifications are known to be important in stabilizing CAPs at particular regions of the genome and in regulating their activity. In addition to the core histone tails, the C-terminal tail of the linker histone H1 has been identified to be important in formation of higher order chromatin structure.

IDRs are also abundant in histone chaperones, transcription factors, chromatin remodeling proteins, chromatin modifying proteins, and architectural proteins (Tantos et al., 2011; Sandhu, 2009; Lazar et al., 2016; Hadidy and Uversky, 2019). The IDRs are now being recognized to play a much more direct role in chromatin regulation than simply dynamically linking folded functional domains. In this review, we discuss several examples of IDRs within CAPs, highlighting their structural and functional characterization, focusing on how NMR spectroscopy and cryo-EM can be used together to build an understanding the molecular mechanisms underlying IDR activities.

FUNCTIONAL ROLES OF IDPS/IDRS IN CHROMATIN REGULATION

IDRs have been found to play essential biological roles in a wide variety of CAPs. As noted above, these include the histones themselves, chaperones, chromatin remodeling and modifying complexes, and architectural proteins. Below, we outline five distinct IDRs that demonstrate the functional diversity of these regions in chromatin regulation.

PRC2 stimulation response motif

Polycomb repressive complex 2 (PRC2) is a histone lysine methyltransferase complex. It consists of four main subunits: EED, SUZ12, RbAp46/48, and the catalytic subunit EZH2. PRC2 specifically methylates lysine 27 of histone H3, producing the epigenetic mark H3K27me3 (Margueron and Reinberg, 2011). The complex has been found to be allosterically stimulated by the H3K27me3 product, which is thought to be important in spreading this mark to form heterochromatin domains. It can also be stimulated by methylated JARID2, a PRC2 co-factor. In the past five years, X-ray crystallography and cryo-EM structures of this complex from several groups have suggested that the molecular basis of this allosteric activation is mediated in part by an IDR in EZH2, termed the stimulation response motif (SRM) (Jiao and Liu, 2015; Justin et al., 2016; Poepsel et al., 2018; Kasinath et al., 2018). In basal state structures of the PRC2 complex, the SRM does not resolve (Figure 3, left). In contrast, in compact structures of active PRC2 bound to methylated substrates, the SRM is resolved and adopts an alpha-helical structure (Figure 3, right). The helical SRM forms contacts with the SET domain which are suggested to stimulate PRC2 activity. Notably, the active state complex was also seen to adopt an extended state in which the SRM is not visible. It was suggested that this conformational plasticity may allow for further fine-tuning of PRC2 activity. Notably, a cryo-EM structure of PRC2 in complex with a di-nucleosome in which one nucleosome contains H3K27me3 and one is unmodified revealed that the methylated nucleosome leads to stabilization of the SRM in an alpha helix conformation (Poepsel et al., 2018). Furthermore, positioning of the complex would allow for methylation of the adjacent unmodified nucleosome, supporting the proposed spreading mechanism.

The lack of resolution of the SRM in the basal and extended active states was interpreted to imply that it is unstructured. However, the exact structural state of the SRM is not accessible by X-ray crystallography or cryo-EM. NMR analysis complimented these structures. While this analysis could not be carried out on the full complex due to size limitations, NMR allowed for study of the isolated SRM structural characteristics. Chemical shift and linewidth analysis revealed a largely disordered conformational state of the SRM but with moderate alpha-helical propensity (Weaver et al., 2019). This indicates that the activate conformation is sampled in the conformational ensemble of the isolated SRM (Figure 3, left). Together, this suggests that modulating the conformational ensemble of the SRM within PRC2, in particular the population of the active state, is one mechanism for fine-tuning the activity of the PRC2 complex. Notably, this would be an example of a conformational selection mechanism.

In addition to stimulating PRC2 activity, a recent study has indicated that SRM can act as a transactivation domain (TAD) outside of the canonical PRC2 complex. While this SRM-containing TAD is sequestered by
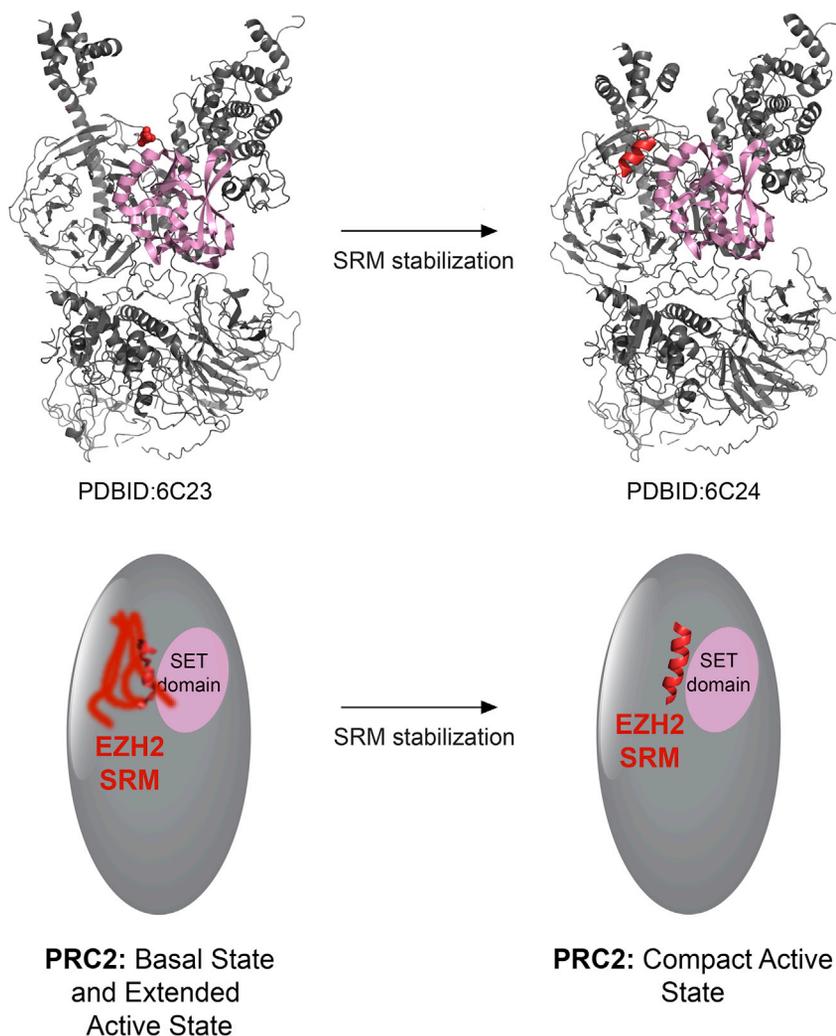
Figure 3. PRC2 SRM function

Top are shown the cryo-EM structures of the extended active (left, PDBID:6C23) and the compact active (right, PDBID:6C24) states of the PRC2 complex. In these structures, the EZH2 SRM IDR (sequence TFIEELIKNY, shown in red) only resolves in the compact active structure (right), and its position is marked by a red sphere in the basal state (left). Below are shown models of the PRC2 complex (depicted as a gray oval) in the basal and extended actives states (left). The SRM (red) is modeled as a conformationally dynamic ensemble with fast exchange between states (blur), only transiently sampling helical structure, as determined by NMR spectroscopy. Upon adopting the compact active state (right), the SRM helical conformation is stabilized forming important contacts with the catalytic SET domain (purple).

intramolecular interactions and oligomerization of EZH2 outside canonical PRC2, cancer-associated phosphorylation can disrupt these interactions and release the TAD, leading to aberrant gene activation (Jiao et al., 2020).

HP1

The heterochromatin protein 1 (HP1) family contains important architectural proteins involved in gene regulation. In humans, there are three paralogs: HP1α, HP1β, and HP1γ. Of these, HP1α is strongly associated with heterochromatin and gene repression, whereas HP1β and HP1γ are associated with both gene repression and activation. All HP1 family members contain a structured chromodomain (CD) that can associate with methylated lysines. In particular, recognition of H3K9me3 was found to be important in formation of heterochromatin. Another structured domain present in HP1, the chromoshadow domain (CSD), is capable of homodimerization and heterodimerization (e.g. across paralogs). This dimerization was proposed to promote bridging of nucleosomes to facilitate heterochromatin formation (Canzio
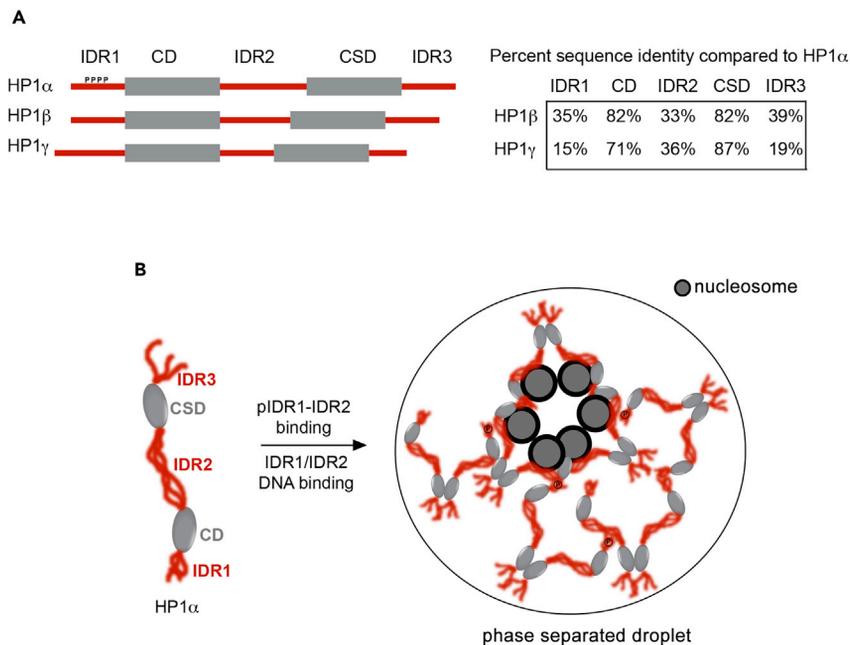
Figure 4. HP1 IDR function

(A) HP1 has three paralogs (HP1α, HP1β, HP1γ) each containing a chromodomain (CD), a chromoshadow domain (CSD), and three IDRs (IDR1, IDR2, IDR3). The sequence identity of HP1β and HP1γ is shown compared to HP1α for each region (right).

(B) HP1α three IDRs (shown in red) modeled as conformational ensembles with fast exchange between states (blur). Phosphorylated IDR1 has been found to associate with IDR2 intermolecularly, and IDR1 and IDR2 have both been found to form contacts with DNA. Together, these interactions are proposed to lead to phase-separated droplets sequestering nucleosomes [shown as a complex of histones (dark gray circles) wrapped by DNA (black rings)] and compacting DNA critical in the formation of heterochromatin.

et al., 2011). Notably, the CD and CSD are highly similar between paralogs (Figure 4A), and thus, the differences in paralog function is not thought to arise from these structured domains.

In addition to the CD and CSD, HP1 contains three low complexity regions (Figure 4A). Modeling and NMR spectroscopy have revealed that these regions are intrinsically disordered (Munari et al., 2012; Velez et al., 2015); IDR1 (also known as the N-terminal extension or NTE) is N-terminal to the CD, IDR2 (also referred to as the hinge region) links the CD and CSD, and IDR3 is C-terminal to the CSD. The major differences between paralogs are found in these IDRs (Figure 4A), and as such, it has been hypothesized that it is these regions that uniquely regulate each HP1 paralog.

IDR1 and IDR2 have been found to interact with DNA in all paralogs, contributing to multivalent association with nucleosomes (Munari et al., 2012; Meehan et al., 2003; Zhao et al., 2000). Due to the smaller size of the HP1 proteins, NMR spectroscopy studies can be carried out on the full-length proteins. Chemical shift and linewidth analysis reveal that the IDRs retain their disordered conformation even in the DNA-bound state, consistent with a fuzzy complex. It was proposed that IDR2 may function to associate with linker DNA to aid in bridging two nucleosomes. Though this was not resolvable by NMR studies, cryo-EM structures of HP1α, HP1β, and HP1γ with a di-nucleosome allowed this to be addressed. Surprisingly, these structures revealed that IDR2 associates primarily with nucleosomal DNA leaving the linker DNA open. Consistent with NMR studies, IDR2 is not well resolved in these structures supporting that it remains highly conformationally dynamic in the bound state (Munari et al., 2012; Machida et al., 2018). Notably, the mechanism of association with di-nucleosomes was found to be nearly identical between all paralogs.

Only recently has a major difference in IDR function between paralogs been identified. Specifically, phosphorylation of IDR1 in HP1α was shown to promote oligomerization through intermolecular interactions

with IDR2, driving phase separation and formation of heterochromatin (Figure 4B) (Strom et al., 2017; Larson et al., 2017). Importantly, these phosphorylation sites are missing in HP1β and HP1γ, making this a unique property of HP1α. The interaction of HP1α IDR2 with DNA also promoted phase separation, and while HP1β IDR2 can associate with DNA, this interaction did not promote phase separation. Thus, the composition and modifiability of IDRs in HP1 paralogs can uniquely regulate the activity of these proteins.

MBD2 IDR

Methyl-CpG binding domain protein 2 (MBD2) contributes to gene repression by associating with methylated CpG islands in DNA and recruiting the nucleosome remodeling and deacetylase (NuRD) co-repressor complex. The NuRD complex harbors the CHD4 chromatin remodeling protein and histone deacetylase 1/2 (HDAC1/2), which together promote a repressive chromatin structure. MBD2 contains a large IDR that plays an important role in protein and DNA interactions (Desai et al., 2015; Ghosh et al., 2010). Methyl-CpG association is mediated by a structured MBD domain (Scarsdale et al., 2011). An adjacent IDR enhances DNA affinity. NMR chemical shift and linewidth indicate that the IDR does not become ordered upon DNA binding and instead adopts a fuzzy complex (Figure 5A, left) (Desai et al., 2015). On longer pieces of DNA, this IDR substantially reduces exchange between methyl-CpG sites, as well as dissociation from the DNA, anchoring the MBD (Figure 5A, right) (Pan et al., 2017). Within the same IDR but C-terminal to the DNA binding region is a stretch that facilitates interaction with HDAC1/2, leading to gene repression. While cryo-EM studies have been carried out on CHD4 in complex with the nucleosome (Farnung et al., 2020), a structure of the MBD2/NURD complex has yet to be completed. This would provide substantial insight into the mechanism of MBD2 IDR function in the proper context.

H1 C-terminal tail

Linker histone H1 associates with the linker DNA between nucleosomes and promotes the formation of higher order chromatin structures associated with gene repression (Fang et al., 2016; Caterino and Hayes, 2011). H1 consists of an unstructured N-terminal region, followed by a structured winged-helix domain, followed by a 100 residue C-terminal tail. NMR chemical shift and linewidth analysis have revealed that the C-terminal tail is intrinsically disordered (Turner et al., 2018) and that it remains disordered in the DNA-bound state, forming a high-affinity fuzzy complex (Turner et al., 2018). The association with DNA was further shown to promote phase separation, which is modulated by phosphorylation of the C-terminal tail. This modification is known to increase mobility in chromatin and promote a more open chromatin structure (Contreras et al., 2003). NMR chemical shift and linewidth analysis of modified H1 reveals that the C-terminal tail remains disordered upon phosphorylation but has decreased affinity for DNA and associated decreased phase separation properties.

A combination of cryo-EM and X-ray crystallography allowed for structural analysis of H1 in complex with a nucleosome. The structure is consistent with the NMR data, as low resolution of the C-terminal tail indicates a fuzzy complex formed with the linker DNA (Figure 5B) (Bednar et al., 2017). However, despite the low density seen for the C-terminal tail, the structure revealed that the tail preferentially associates with one linker arm over the other. This binding mode introduces asymmetry to the nucleosome, which is proposed to be necessary in the formation of higher order chromatin structures.

FACT C-terminal domain

FACT is a histone chaperone that can assist both in nucleosome assembly and disassembly through rearrangements of the H2A/H2B dimer. FACT contains two subunits, SSRP1 and Spt16. At the C-terminus of Spt16 is an acidic IDR that is critical in H2A/H2B binding and referred to as the acidic intrinsically disordered (AID) segment or the C-terminal domain. Two recent cryo-EM structures of FACT or Spt16 in complex with nucleosome substrates reveal that the unstructured AID adopts a structured conformation in complex with nucleosome substrate (Figure 5C) (Liu et al., 2020; Mayanagi et al., 2019). Intriguingly, it acts as a DNA mimetic, wrapping around the H2A/H2B dimer. It was proposed that this process depends on the phosphorylation of residues in the AID (pAID) to mimic the DNA (Mayanagi et al., 2019). As several studies have found that the H3 N-terminal tails interact with DNA (Kan et al., 2007; Mutskov et al., 1998; Pilotto et al., 2015; Gatchalian et al., 2017; Shaytan et al., 2015; Li and Kono, 2016; Lehmann et al., 2020; Stützer et al., 2016; Morrison et al., 2018), it was also proposed that pAID association might affect the H3 tails.

Histone tails generally do not resolve in the cryo-EM structures and did not resolve in the FACT/nucleosome structures. However, NMR spectroscopy has proven very powerful in the study of histone tails in
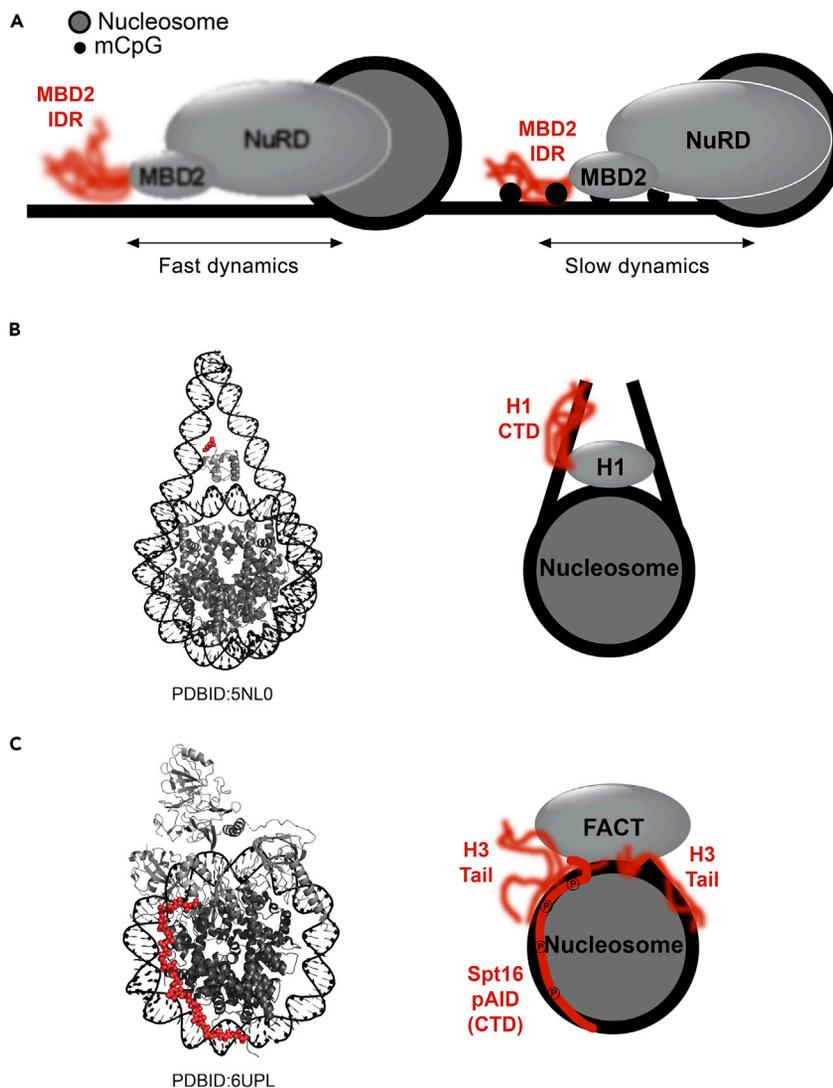
Figure 5. MBD2, H1, and FACT IDR function

(A) MBD2 functions with the NuRD complex to remodel nucleosomes and facilitate formation of heterochromatin at methylated DNA regions. DNA is shown in black, and nucleosomes are shown as a complex of histones (dark gray circles) wrapped by the DNA. MBD2 contains a methyl-DNA binding domain (light gray) and an adjacent DNA binding IDR (red). NMR has shown that the IDR forms a fuzzy complex with DNA, modeled as a broad conformational ensemble with fast exchange between states (highly blurred). In the absence of DNA methylation, MBD2/NuRD can move quickly along DNA. Upon encountering DNA methylation (black circles), the IDR reduces exchange between methyl-DNA sites and dissociation from the DNA, leading to much slower movement of the MBD2/NuRD complex along chromatin (slightly blurred).

(B) Linker histone H1 contains a globular domain (light gray) and IDR called the C-terminal domain (CTD) (red). Shown on the left is the crystal structure of the H1/nucleosome complex (PDBID:5NL0), and the position of the unresolved C-terminus is marked by a red sphere. A model integrating cryo-EM and NMR data is depicted on the right. NMR spectroscopy has shown that the CTD forms a fuzzy complex with DNA; thus, the CTD is shown as a broad conformational ensemble with fast exchange between states (blur). The cryo-EM structure of the H1-nucleosome complex [the nucleosome is shown as a complex of histones (dark gray circle) wrapped by DNA (black ring)] reveals that the CTD favors one strand of linker DNA.

(C) Cryo-EM demonstrates that an IDR referred to as the AID in the Spt16 subunit of the FACT chaperone adopts a stable structure and mimics DNA when in complex with a nucleosome. Shown on the left is a representative structure of FACT in complex with the nucleosome (PDBID:6UPL), with the AID colored red. Shown on the right is the model combining the cryo-EM structure and NMR analysis of the H3 tails upon pAID binding. NMR studies have demonstrated that both H3 tails (red) form fuzzy interactions with DNA, and they are modeled as a broad conformational ensemble with fast exchange between states (blur). However, binding of pAID (sharp red line) was found to alter the conformational ensemble of the adjacent H3 tail, making it more accessible.

the context of the nucleosome (Emmerik and Ingen, 2019). Through NMR chemical shift and linewidth analysis, it was found that the tails are intrinsically disordered and adopt a fuzzy complex with DNA in the nucleosome context (Stützer et al., 2016; Morrison et al., 2018). A recent NMR study also provided the ability to determine the effect of pAID association upon the H3 tails. Analysis of the H3 tails chemical shift revealed that pAID replacement of DNA alters the conformational ensemble of the adjacent H3 tail, while having little to no effect on the other tail (Tsunaka et al., 2020). While the affected H3 tail still adopts a fuzzy complex with the remaining DNA strand, it becomes more accessible in the presence of pAID (Figure 5C).

FUTURE DIRECTIONS

IDRs are being increasingly recognized to play critical roles in fundamental cellular processes, operating in a manner that differs from structured domains. They are abundant in CAPs and implicated in diverse biological functions, including regulation of chromatin association, enzymatic activity, and liquid-liquid phase separation. Post-translation modification can rapidly alter IDR function, allowing for dynamic process. Given their important regulatory roles, these regions hold tremendous therapeutic potential. However, a poor understanding of the mechanistic basis for their activities has made traditional structure-function approaches insufficient. Thus, new approaches to characterizing the conformation, dynamics, and regulation of IDRs are critical.

Advances in NMR methodology as well as cryo-EM have opened the door to better understand IDRs. As studies outlined here highlight, NMR and cryo-EM analysis of IDRs is highly synergistic. NMR offers the ability to readily investigate conformational dynamics, the mechanism of binding, and effect of PTMs. However, it is limited with respect to molecular size and therefore often cannot be used to investigate IDRs in their proper context. On the other hand, cryo-EM is well suited to characterization of large complexes and is tolerant to conformational heterogeneity. While IDRs are typically not well resolved, they are often visible at low resolution. Together, these techniques allow for building a model of the IDR in its physiologically relevant environment and provide substantial insight into the mechanistic basis of function. Moving forward, this combined approach could not only reveal fundamental mechanisms of chromatin regulation but also aid in the development of potential therapeutics that target IDRs.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

REFERENCES

Altman, R.B., Dunker, A.K., Hunker, L., Lauderdale, K., Klein, T.E., Williams, R.M., Obradovic, Z., Mathura, V., Braun, W., Garner, E.C., Young, J., Takayama, S., Brown, C.J., and Dunker, A.K. (2000). The protein non-folding problem: amino acid determinants of intrinsic order and disorder. Pac Symposium Biocomput Pac Symposium Biocomput, 89–100, https://doi.org/10.1142/9789814447362_0010.

Bah, A., and Forman-Kay, J.D. (2016). Modulation of Intrinsically Disordered Protein Function by Post-translational Modifications. J Biological Chem 291, 6696–6705, https://doi.org/10.1074/jbc.r115.695056.

Bednar, J., Garcia-Saez, I., Boopathi, R., Cutter, A.R., Papai, G., Reymer, A., Syed, S.H., Lone, I.N., Tonchev, O., Crucifix, C., Menoni, H., Papin, C., Skoufias, D.A., Kurumizaka, H., Lavery, R., Hamiche, A., Hayes, J.J., Schultz, P., Angelov, D., Petosa, C., and Dimitrov, S. (2017). Structure and Dynamics of a 197 bp Nucleosome in Complex with Linker Histone H1. Molecular Cell 66, 384–397.e8, https://doi.org/10.1016/j.molcel.2017.04.012.

Borgia, A., Borgia, M.B., Bugge, K., Kissling, V.M., Heidarsson, P.O., Fernandes, C.B., Sottini, A., Soranno, A., Buholzer, K.J., Nettels, D., Kragelund, B.B., Best, R.B., and Schuler, B. (2018). Extreme disorder in an ultrahigh-affinity protein complex. Nature 555, 61–66, https://doi.org/10.1038/nature25762.

Brutscher, B., Felli, I.C., Gil-Caballero, S., Hošek, T., Kümmerle, R., Piai, A., Pierattelli, R., and Sólyom, Z. (2015). Intrinsically Disordered Proteins Studied by NMR Spectroscopy. Adv Exp Med Biol 870, 49–122, https://doi.org/10.1007/978-3-319-20164-1_3.

Canzio, D., Chang, E.Y., Shankar, S., Kuchenbecker, K.M., Simon, M.D., Madhani, H.D., Narlikar, G.J., and Al-Sady, B. (2011). Chromodomain-mediated oligomerization of HP1 suggests a nucleosome-bridging mechanism for heterochromatin assembly. Molecular Cell 41, 67–81, https://doi.org/10.1016/j.molcel.2010.12.016.

Caterino, T.L., and Hayes, J.J. (2011). Structure of the H1 C-terminal domain and function in chromatin condensation. Biochem Cell Biology Biochimie Et Biologie Cell 89, 35–44, https://doi.org/10.1139/o10-024.

Cheng, Y. (2018). Single-particle cryo-EM-How did it get here and where will it go. Science 361, 876–880, https://doi.org/10.1126/science.aat4346.

Contreras, A., Hale, T.K., Stenoien, D.L., Rosen, J.M., Mancini, M.A., and Herrera, R.E. (2003). The Dynamic Mobility of Histone H1 Is Regulated by Cyclin/CDK Phosphorylation. Mol Cell Biol 23, 8626–8636, https://doi.org/10.1128/mcb.23.23.8626-8636.2003.

Desai, M.A., Webb, H.D., Sinanan, L.M., Scarsdale, J.N., Walavalkar, N.M., Ginder, G.D., and Williams, D.C. (2015). An intrinsically disordered region of methyl-CpG binding domain protein 2 (MBD2) recruits the histone deacetylase core of the NuRD complex. Nucleic Acids Res 43, 3100–3113, https://doi.org/10.1093/nar/gkv168.

Dunker, A.K., Brown, C.J., Lawson, J.D., Iakoucheva, L.M., and Obradović, Z. (2002). Intrinsic Disorder and Protein Function †. Biochemistry-us 41, 6573–6582, https://doi.org/10.1021/bi012159+.

Dunker, A.K., and Obradovic, Z. (2001). The protein trinity—linking function and disorder. Nat Biotechnol 19, 805–806, https://doi.org/10.1038/nbt0901-805.

Dyson, H.J. (2011). Roles of intrinsic disorder in protein-nucleic acid interactions. Mol Biosyst 8, 97–104, https://doi.org/10.1039/c1mb05258f.

van Emmerik, C., and van Ingen, H. (2019). Unspinning chromatin: Revealing the dynamic nucleosome landscape by NMR. Prog Nucl Mag Res Sp 110, 1–19, https://doi.org/10.1016/j.pnmrs.2019.01.002.

Fang, H., Wei, S., Lee, T.-H., and Hayes, J.J. (2016). Chromatin structure-dependent conformations of the H1 CTD. Nucleic acids research 44, 9131–9141, https://doi.org/10.1093/nar/gkw586.

Farnung, L., Ochmann, M., and Cramer, P. (2020). Nucleosome-CHD4 chromatin remodeller structure maps human disease mutations. Elife 9, e56178, https://doi.org/10.7554/elife.56178.

Fuxreiter, M. (2018). Fuzziness in Protein Interactions—A Historical Perspective. J Mol Biol 430, 2278–2287, https://doi.org/10.1016/j.jmb.2018.02.015.

Fuxreiter, M., Simon, I., Friedrich, P., and Tompa, P. (2004). Preformed Structural Elements Feature in Partner Recognition by Intrinsically Unstructured Proteins. J Mol Biol 338, 1015–1026, https://doi.org/10.1016/j.jmb.2004.03.017.

Gatchalian, J., Wang, X., Ikebe, J., Cox, K.L., Tencer, A.H., Zhang, Y., Burge, N.L., Di, L., Gibson, M.D., Musselman, C.A., Poirier, M.G., Kono, H., Hayes, J.J., and Kutateladze, T.G. (2017). Accessibility of the histone H3 tail in the nucleosome for binding of paired readers. Nat Commun 8, 1489, https://doi.org/10.1038/s41467-017-01598-x.

Geraets, J.A., Pothula, K.R., and Schröder, G.F. (2020). Integrating cryo-EM and NMR data. Curr Opin Struc Biol 61, 173–181, https://doi.org/10.1016/j.sbi.2020.01.008.

Ghosh, R.P., Nikitina, T., Horowitz-Scherer, R.A., Gierasch, L.M., Uversky, V.N., Hite, K., Hansen, J.C., and Woodcock, C.L. (2010). Unique Physical Properties and Interactions of the Domains of Methylated DNA Binding Protein 2. Biochemistry-us 49, 4395–4410, https://doi.org/10.1021/bi9019753.

Gibbs, E.B., Cook, E.C., and Showalter, S.A. (2017). Application of NMR to studies of intrinsically disordered proteins. Arch Biochem Biophys 628, 57–70, https://doi.org/10.1016/j.abb.2017.05.008.

Hadidy, N.E., and Uversky, V.N. (2019). Intrinsic Disorder of the BAF Complex: Roles in Chromatin Remodeling and Disease Development. Int J Mol Sci 20, 5260, https://doi.org/10.3390/ijms20215260.

Jensen, M.R., Zweckstetter, M., Huang, J., and Blackledge, M. (2014). Exploring Free-Energy Landscapes of Intrinsically Disordered Proteins at Atomic Resolution Using NMR Spectroscopy. Chem Rev 114, 6632–6660, https://doi.org/10.1021/cr400688u.

Jiao, L., and Liu, X. (2015). Structural basis of histone H3K27 trimethylation by an active polycomb repressive complex 2. Science 350, aac4383. .aac4383. https://doi.org/10.1002/jcc.20084.

Jiao, L., Shubbar, M., Yang, X., Zhang, Q., Chen, S., Wu, Q., Chen, Z., Rizo, J., and Liu, X. (2020). A partially disordered region connects gene repression and activation functions of EZH2. P Natl Acad Sci Usa 117, 16992–17002, https://doi.org/10.1073/pnas.1914866117.

Justin, N., Zhang, Y., Tarricone, C., Martin, S.R., Chen, S., Underwood, E., Marco, V.D., Haire, L.F., Walker, P.A., Reinberg, D., Wilson, J.R., and Gamblin, S.J. (2016). Structural basis of oncogenic histone H3K27M inhibition of human polycomb repressive complex 2. Nature communications 7, 11316, https://doi.org/10.1038/ncomms11316.

Kan, P.-Y., Lu, X., Hansen, J.C., and Hayes, J.J. (2007). The H3 tail domain participates in multiple interactions during folding and self-association of nucleosome arrays. Mol Cell Biol 27, 2084–2091, https://doi.org/10.1128/mcb.02181-06.

Kasinath, V., Faini, M., Poepsel, S., Reif, D., Feng, X.A., Stjepanovic, G., Aebersold, R., and Nogales, E. (2018). Structures of human PRC2 with its cofactors AEBP2 and JARID2. Science 359, 940–944, https://doi.org/10.1126/science.aar5700.

Larson, A.G., Elnatan, D., Keenen, M.M., Trnka, M.J., Johnston, J.B., Burlingame, A.L., Agard, D.A., Redding, S., and Narlikar, G.J. (2017). Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. Nature 547, 236–240, https://doi.org/10.1038/nature22822.

Lazar, T., Schad, E., Szabo, B., Horvath, T., Meszaros, A., Tompa, P., and Tantos, A. (2016). Intrinsic protein disorder in histone lysine methylation. Biol Direct 11, 30, https://doi.org/10.1186/s13062-016-0129-2.

Lehmann, K., Felekyan, S., Kühnemuth, R., Dimura, M., Tóth, K., Seidel, C.A.M., and Langowski, J. (2020). Dynamics of the nucleosomal histone H3 N-terminal tail revealed by high precision single-molecule FRET. Nucleic Acids Res 48, 1551–1571, https://doi.org/10.1093/nar/gkz1186.

Li, Z., and Kono, H. (2016). Distinct Roles of Histone H3 and H2A Tails in Nucleosome Stability. Scientific reports 6, 31437, https://doi.org/10.1038/srep31437.

Liu, Y., Zhou, K., Zhang, N., Wei, H., Tan, Y.Z., Zhang, Z., Carragher, B., Potter, C.S., D'Arcy, S., and Luger, K. (2020). FACT caught in the act of manipulating the nucleosome. Nature 577, 426–431, https://doi.org/10.1038/s41586-019-1820-0.

Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J.-I., Wolf, M., and Kurumizaka, H. (2018). Structural Basis of Heterochromatin Formation by Human HP1.

Molecular Cell. https://doi.org/10.1016/j.molcel.2017.12.011.

Margueron, R., and Reinberg, D. (2011). The Polycomb complex PRC2 and its mark in life. Nature 469, 343–349, https://doi.org/10.1038/nature09784.

Mayanagi, K., Saikusa, K., Miyazaki, N., Akashi, S., Iwasaki, K., Nishimura, Y., Morikawa, K., and Tsunaka, Y. (2019). Structural visualization of key steps in nucleosome reorganization by human FACT. Sci Rep-uk 9, 10183, https://doi.org/10.1038/s41598-019-46617-7.

Meehan, R.R., Kao, C., and Pennings, S. (2003). HP1 binding to native chromatin in vitro is determined by the hinge region and not by the chromodomain. Embo J 22, 3164–3174, https://doi.org/10.1093/emboj/cdg306.

Mollica, L., Bessa, L.M., Hanoulle, X., Jensen, M.R., Blackledge, M., and Schneider, R. (2016). Binding Mechanisms of Intrinsically Disordered Proteins: Theory, Simulation, and Experiment. Frontiers Mol Biosci 3, 52, https://doi.org/10.3389/fmolb.2016.00052.

Morrison, E.A., Bowerman, S., Sylvers, K.L., Wereszczynski, J., and Musselman, C.A. (2018). The conformation of the histone H3 tail inhibits association of the BPTF PHD finger with the nucleosome. Elife 7, e31481, https://doi.org/10.7554/elife.31481.

Munari, F., Soeroes, S., Zenn, H.M., Schomburg, A., Kost, N., Schröder, S., Klingberg, R., Rezaei-Ghaleh, N., Stützer, A., Gelato, K.A., Walla, P.J., Becker, S., Schwarzer, D., Zimmermann, B., Fischle, W., and Zweckstetter, M. (2012). Methylation of Lysine 9 in Histone H3 Directs Alternative Modes of Highly Dynamic Interaction of Heterochromatin Protein hHP1β with the Nucleosome. J Biol Chem 287, 33756–33765, https://doi.org/10.1074/jbc.m112.390849.

Mutskov, V., Gerber, D., Angelov, D., Ausio, J., Workman, J., and Dimitrov, S. (1998). Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding. Molecular and cellular biology 18, 6293–6304.

Olsen, J.G., Teilum, K., and Kragelund, B.B. (2017). Behaviour of intrinsically disordered proteins in protein-protein complexes with an emphasis on fuzziness. Cell Mol Life Sci Cmls 74, 3175–3183, https://doi.org/10.1007/s00018-017-2560-7.

Pan, H., Bilinovich, S.M., Kaur, P., Riehn, R., Wang, H., and Williams, D.C. (2017). CpG and methylation-dependent DNA binding and dynamics of the methylcytosine binding domain 2 protein at the single-molecule level. Nucleic acids research. https://doi.org/10.1093/nar/gkx548.

Pilotto, S., Speranzini, V., Tortorici, M., Durand, D., Fish, A., Valente, S., Forneris, F., Mai, A., Sixma, T.K., Vachette, P., and Mattevi, A. (2015). Interplay among nucleosomal DNA, histone tails, and corepressor CoREST underlies LSD1-mediated H3 demethylation. Proc National Acad Sci 112, 2752–2757, https://doi.org/10.1073/pnas.1419468112.

Poepsel, S., Kasinath, V., and Nogales, E. (2018). Cryo-EM structures of PRC2 simultaneously engaged with two functionally distinct

nucleosomes. Nature structural & molecular biology 1, https://doi.org/10.1038/s41594-018-0023-y.

Romero, P., Obradovic, Z., Li, X., Garner, E.C., Brown, C.J., and Dunker, A.K. (2001). Sequence complexity of disordered protein. Proteins Struct Funct Bioinform 42, 38–48, https://doi.org/10.1002/1097-0134(20010101)42:1<38::aid-prot50>3.0.co;2-3.

Rothbart, S.B., and Strahl, B.D. (2014). Interpreting the language of histone and DNA modifications. Biochimica et biophysica acta 1839, 627–643, https://doi.org/10.1016/j.bbagrm.2014.03.001.

Sahu, D., Bastidas, M., and Showalter, S.A. (2014). Generating NMR chemical shift assignments of intrinsically disordered proteins using carbon-detected NMR methods. Anal Biochem 449, 17–25, https://doi.org/10.1016/j.ab.2013.12.005.

Sandhu, K.S. (2009). Intrinsic disorder explains diverse nuclear roles of chromatin remodeling proteins. J Mol Recognit 22, 1–8, https://doi.org/10.1002/jmr.915.

Scarsdale, J.N., Webb, H.D., Ginder, G.D., and Williams, D.C. (2011). Solution structure and dynamic analysis of chicken MBD2 methyl binding domain bound to a target-methylated DNA sequence. Nucleic Acids Res 39, 6741–6752, https://doi.org/10.1093/nar/gkr262.

Sharma, R., Raduly, Z., Miskei, M., and Fuxreiter, M. (2015). Fuzzy complexes: Specific binding without complete folding. Febs Lett 589, 2533–2542, https://doi.org/10.1016/j.febslet.2015.07.022.

Shaytan, A.K., Armeev, G.A., Goncearenco, A., Zhurkin, V.B., Landsman, D., and Panchenko, A.R. (2015). Coupling between Histone Conformations and DNA Geometry in Nucleosomes on a Microsecond Timescale:

Atomistic Insights into Nucleosome Functions. Journal of molecular biology. https://doi.org/10.1016/j.jmb.2015.12.004.

Strom, A.R., Emelyanov, A.V., Mir, M., Fyodorov, D.V., Darzacq, X., and Karpen, G.H. (2017). Phase separation drives heterochromatin domain formation. Nature 547, 241–245, https://doi.org/10.1038/nature22989.

Stützer, A., Liokatis, S., Kiesel, A., Schwarzer, D., Sprangers, R., Söding, J., Selenko, P., and Fischle, W. (2016). Modulations of DNA Contacts by Linker Histones and Post-translational Modifications Determine the Mobility and Modifiability of Nucleosomal H3 Tails. Molecular Cell 61, 247–259, https://doi.org/10.1016/j.molcel.2015.12.015.

Tantos, A., Han, K.-H., and Tompa, P. (2011). Intrinsic disorder in cell signaling and gene transcription. Mol Cell Endocrinol 348, 457–465, https://doi.org/10.1016/j.mce.2011.07.015.

Theillet, F.-X., Smet-Nocca, C., Liokatis, S., Thongwichian, R., Kosten, J., Yoon, M.-K., Kriwacki, R.W., Landrieu, I., Lippens, G., and Selenko, P. (2012). Cell signaling, post-translational protein modifications and NMR spectroscopy. J Biomol Nmr 54, 217–236, https://doi.org/10.1007/s10858-012-9674-x.

Tompa, P. (2002). Intrinsically unstructured proteins. Trends Biochem Sci 27, 527–533, https://doi.org/10.1016/s0968-0004(02)02169-2.

Tompa, P., and Fuxreiter, M. (2008). Fuzzy complexes: polymorphism and structural disorder in protein–protein interactions. Trends Biochem Sci 33, 2–8, https://doi.org/10.1016/j.tibs.2007.10.003.

Tsunaka, Y., Ohtomo, H., Morikawa, K., and Nishimura, Y. (2020). Partial replacement of nucleosomal DNA with human FACT induces dynamic exposure and acetylation of histone H3

N-terminal tails. Iscience 23, 101641, https://doi.org/10.1016/j.isci.2020.101641.

Turner, A.L., Watson, M., Wilkins, O.G., Cato, L., Travers, A., Thomas, J.O., and Stott, K. (2018). Highly disordered histone H1-DNA model complexes and their condensates. Proceedings of the National Academy of Sciences, 201805943, https://doi.org/10.1073/pnas.1805943115.

Uversky, V.N. (2017). Intrinsic Disorder, Protein-Protein Interactions, and Disease. Adv Protein Chem Str 110, 85–121, https://doi.org/10.1016/bs.apcsb.2017.06.005.

Velez, G., Lin, M., Christensen, T., Faubion, W.A., Lomberk, G., and Urrutia, R. (2015). Evidence supporting a critical contribution of intrinsically disordered regions to the biochemical behavior of full-length human HP1γ. J Mol Model 22, 12, https://doi.org/10.1007/s00894-015-2874-z.

Weaver, T.M., Liu, J., Connelly, K.E., Coble, C., Varzavand, K., Dykhuizen, E.C., and Musselman, C.A. (2019). The EZH2 SANT1 domain is a histone reader providing sensitivity to the modification state of the H4 tail. Sci Rep-uk 9, 987, https://doi.org/10.1038/s41598-018-37699-w.

Wright, P.E., and Dyson, H.J. (1999). Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. J Mol Biol 293, 321–331, https://doi.org/10.1006/jmbi.1999.3110.

Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nature structural & molecular biology 20, 259–266, https://doi.org/10.1038/nsmb.2470.

Zhao, T., Heyduk, T., Allis, C.D., and Eissenberg, J.C. (2000). Heterochromatin Protein 1 Binds to Nucleosomes and DNA in Vitro. J Biol Chem 275, 28332–28338, https://doi.org/10.1074/jbc.m003493200.