Protein intrinsic disorder on a dynamic nucleosomal landscape

Sveinn Bjarnason[a,†], Sarah F. Ruidiaz[a,†], Jordan McIvor[b,†], Davide Mercadante[b,*], and Pétur O. Heidarsson[a,*]

[a]Department of Biochemistry, Science Institute, University of Iceland, Reykjavík, Iceland
[b]School of Chemical Science, University of Auckland, Auckland, New Zealand
*Corresponding authors: e-mail address: davide.mercadante@auckland.ac.nz; pheidarsson@hi.is

Contents

Abstract

The complex nucleoprotein landscape of the eukaryotic cell nucleus is rich in dynamic proteins that lack a stable three-dimensional structure. Many of these intrinsically disordered proteins operate directly on the first fundamental level of genome compaction: the nucleosome. Here we give an overview of how disordered interactions with and within nucleosomes shape the dynamics, architecture, and epigenetic regulation of the genetic material, controlling cellular transcription patterns. We highlight experimental and computational challenges in the study of protein disorder and illustrate how integrative approaches are increasingly unveiling the fine details of nuclear interaction

† These authors contributed equally.

Progress in Molecular Biology and Translational Science
ISSN 1877-1173
https://doi.org/10.1016/bs.pmbts.2021.06.006

networks. We finally dissect sequence properties encoded in disordered regions and assess common features of disordered nucleosome-binding proteins. As drivers of many critical biological processes, disordered proteins are integral to a comprehensive molecular view of the dynamic nuclear milieu.

Abbreviations

BD	Brownian dynamics
CG	coarse-grained
DBD	DNA binding domain
FRET	Förster resonance energy transfer
IDP	intrinsically disordered protein
IDR	intrinsically disordered region
MC	Monte Carlo
MD	molecular dynamics
NCP	nucleosome core particle
NMR	nuclear magnetic resonance
NRL	nucleosome repeat length
PTM	post-translational modification
TAD	transactivation domain
TF	transcription factor

1. Introduction

As organisms become increasingly complex, so too must they evolve a more sophisticated molecular alphabet. The recent discovery of proteins that can adopt multiple structural states is one way of addressing this complexity and it has dramatically changed our view of the protein structure-function paradigm.[1,2] Intrinsically disordered proteins (IDPs) either do not contain any well-defined secondary structure element or have long unstructured regions (IDRs), and they fluctuate between a multitude of isoenergetic structural states. These proteins, which comprise an estimated third of the human proteome,[3] are particularly prominent in the nucleus where as much as 70% have been shown or predicted to be IDPs.[4] The cell nucleus, which encompasses the genetic material, is a complex and moldable nucleoprotein landscape, shaped by frequent epigenetic changes that regulate the pattern of gene expression, and ultimately the organismal phenotype. It is thus unsurprising that a multivalent and dynamic nuclear proteome is needed to steer such a diverse environment. The conformational plasticity mediated by intrinsic disorder has been suggested to provide additional levels of functionality to complex cellular regulatory mechanisms.

In this chapter, we highlight protein disorder in the nucleus and emphasize the interplay between IDPs and the nucleosomal landscape that leads to a

functional output. We first define the general components of the nuclear environment, before discussing the challenges and recent advances in understanding structural disorder within the context of transcription. We then compile and dissect a subset of important molecular systems in the nucleus that involve disordered interactions, including the effects of chemical modifications, and overview the resulting biological consequences. We exclusively review the interactions of structural disorder within nucleosomes and chromatin, but for reviews on IDP interactions with nucleic acids, we refer to excellent work on those topics.[5–7] Deciphering the complexity of molecular disorder in the chromatin landscape is an exceedingly challenging task. Yet, recent work has begun to map the functions of many constituent proteins of the nucleus by using innovative biophysical strategies, moving us ever closer to a comprehensive molecular view of the cell nucleus.

2. Protein intrinsic disorder on a nucleosomal landscape

2.1 Components of the nuclear environment

The importance of IDPs and IDRs in cell biology is now well established, and their prevalence in signaling and regulatory pathways has been clearly demonstrated.[6] It is their unique conformational properties that make them ideally suited for their roles. High structural heterogeneity, a consequence of their low complexity and biased amino-acid sequences,[8] imparts IDPs with multivalency in many cases, allowing them to interact with more than one biomolecular partner.[9] Even though the presence of disordered proteins in the nucleus has been recognized for decades, it is only relatively recently that their functions have surfaced. IDPs, which have sometimes been called constituents of the dark proteome,[10] are now increasingly being illuminated as key players in the nucleus of eukaryotes.

To appreciate the many roles played by IDPs and IDRs in the nucleus, we first need to clearly define the nuclear architecture that they operate within (Fig. 1). The genetic material for a typical human cell is composed of ~4.6 million basepairs of DNA, which contain the instructions for generating the cell's proteome. The DNA is substantially compacted to fit this enormous amount into the relatively tiny nucleus, and at all stages of DNA compaction we encounter dynamic protein disorder in one form or another. The first level of compaction is to wrap the DNA around an octamer of the core histones (H2A, H2B, H3, and H4) containing two copies of each, forming the nucleosome.[12] A chromatosome is then constructed by binding of linker histone H1 (H1), which attaches to the dyad of a nucleosome (Fig. 1).[13] Both the core and linker histones contain a large amount of

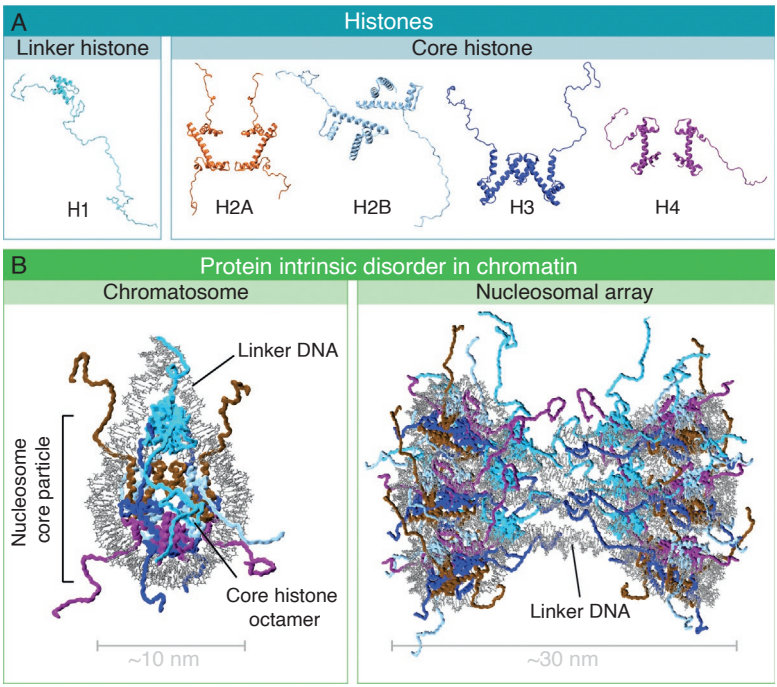

Fig. 1 Protein disorder on a nucleosomal landscape. Intrinsic disorder is a large component of the nucleosomal landscape, contributing to chromatin architecture, dynamics and overall function. (A and B) The nucleosome core particle (NCP) is composed of an octamer of core histones (H2A, H2B, H3, and H4), around which ~147 bp of DNA (gray) is wound in a left-handed super-helical manner.[11] Within the NCP, H2A (orange), H2B (light blue), H3 (dark blue) and H4 (magenta) homodimerize *via* interactions in the structured domains, while the intrinsically disordered N- and C-terminal regions extend into the local nucleosomal space. Linker histone H1 (cyan) binds on or close to the nucleosomal dyad, forming the chromatosome (B) and uses its long disordered and highly basic C-terminal domain to drive conformational changes in linker DNA, impacting the overall structure of poly-nucleosomal arrays and ultimately chromatin fibers.

disorder regulating nucleosome structure and dynamics and ultimately impacting global chromatin structure.[14] There are many histone variants, some cell- or tissue-specific, that can be dynamically exchanged to impart nucleosomes with distinct structural properties.[15] H1 rapidly exchanges between nucleosomes on the second to minute timescale *in vivo*,[16] using largely its positively charged and disordered C-terminal tail to drive orientational changes in linker DNA connecting adjacent nucleosomes.[17] Local interactions between nucleosomes, involving the disordered histone regions, and binding of various regulatory proteins modulate nucleosomal structure and dynamics (recently reviewed in[18]), and subsequently

chromatin condensation into higher-order structures. Protein disorder thus plays an integral role in the formation, regulation, recognition, and modification of genome architecture.

2.2 Post-translational modifications fine-tune disordered interactions

To add yet another layer of complexity, most IDPs are chemically and reversibly modified after translation from the ribosome.[19] Histones and their variants have multiple post-translational modification (PTM) sites, mostly in their IDRs, where the pattern and number of modifications can fine-tune their interactions with nucleosomes and other biomolecules.[20] In general, PTMs render the proteome far more vast than the genome, with hundreds of thousands or even up to a million chemically distinct proteins at any given time in the cell.[21] Chemical modifications can change stability, concentration, localization, conformations, and interaction patterns of proteins, providing an important form of regulation and signaling. The most common modifications include (but are not limited by) covalent yet reversible chemical additions such as phosphorylations, acetylations, methylations, hydroxylations, and amidations, as well as attachments of sugar moieties or entire proteins involving sumoylation or ubiquitinylation.[19] In addition to protein modifications, DNA can be modified, most commonly involving cytosine methylation, and when located in CpG islands on promoters, this covalent modification is normally associated with gene repression.[22] Together, these modifications form an almost unfathomably complex and constantly evolving molecular surroundings that dictate the state of a cell.

Protein PTM sites are frequently located in IDRs, partly due to their accessibility to modifying enzymes such as kinases, acetylases, and methylases.[21] PTMs can induce or relieve secondary structure propensity or have a global effect on the structural ensemble sampled by the disordered region, potentially shifting the ensemble to a certain functional state, resembling conformational selection. They can also affect disorder-to-order transitions, which are a common interaction-mode for IDPs,[23] or affect the degree of disorder in fuzzy[24] or fully[25] disordered complexes. PTMs that affect charges will influence intrachain electrostatic interactions, which have an important role in determining the compactness of a disordered region.[26] In general, PTMs modulate the structural and dynamical properties of IDPs, fine-tuning their functional repertoire. We now explore the arsenal of experimental and computational approaches that can and have been used to engage with IDPs, ranging from simple gel-based binding experiments to sophisticated atomistic models.

2.3 Challenges in studying disordered protein interactions with nucleosomes

Quantitative measurements of structurally heterogeneous polypeptides binding to the dynamic nucleoprotein landscape is a daunting task. Nonetheless, technological advances that enable access to various levels of molecular detail are continuously emerging.[27] As an initial characterization of protein-DNA interactions, classical binding experiments have often involved using an electrophoretic mobility shift assay (EMSA). EMSA is a simple and rapid way to monitor the binding of proteins (structured or disordered) to DNA by observing the changed migration pattern of DNA as a result of protein binding.[28] The EMSA can provide information on binding affinity and specificity but may underestimate these parameters as during the electrophoresis the system is out of equilibrium. In addition, the EMSA does not give direct information on actual binding sites, *i.e.*, it does not detect the exact base pair sequence which is recognized. Exact sequence with base-pair resolution can be determined using footprinting assays (*e.g.*, hydroxyl radical footprinting[29]) or nuclease digestion (*e.g.*, Micrococcal nuclease or MNase). Isothermal titration calorimetry (ITC) and more recently microscale thermophoresis, enable quantitative determination of protein-DNA binding affinity and specificity.[30] Chromatin immunoprecipitation, which relies on chemical crosslinking of the target protein to DNA, combined with sequencing (ChIP-Seq[31]) is a powerful method to find protein binding sites *in vivo*. Similarly, ATAC-seq[32] (Assay for Transposase-Accessible Chromatin using sequencing) reveals genome-wide chromatin accessibility as a consequence of chromatin remodeling or other processes. In this elegant method a transposase is used to incorporate next-generation sequencing adapters into chromatin, which after sequencing provides a map of genome-wide chromatin accessibility. To understand local contributions from the polypeptide sequence, the before mentioned approaches can be combined with genetic and biochemical modifications of target proteins, such as introducing domain deletions/additions, charge reversal, domain swapping or local mutations. Still, without a view into microscopic molecular-level details, the underlying physical principles of protein function can be challenging to deconvolute. Cryogenic electron microscopy (cryo-EM) and X-ray crystallography enable determining atomic-resolution three-dimensional structures of macromolecules.[33] X-ray crystallography determines structures from diffraction patterns and it is the most widely used technique in structural biology.[34] Modern cryo-EM is rapidly catching up through recent advances in deep-frozen sample preparations, direct electron detection

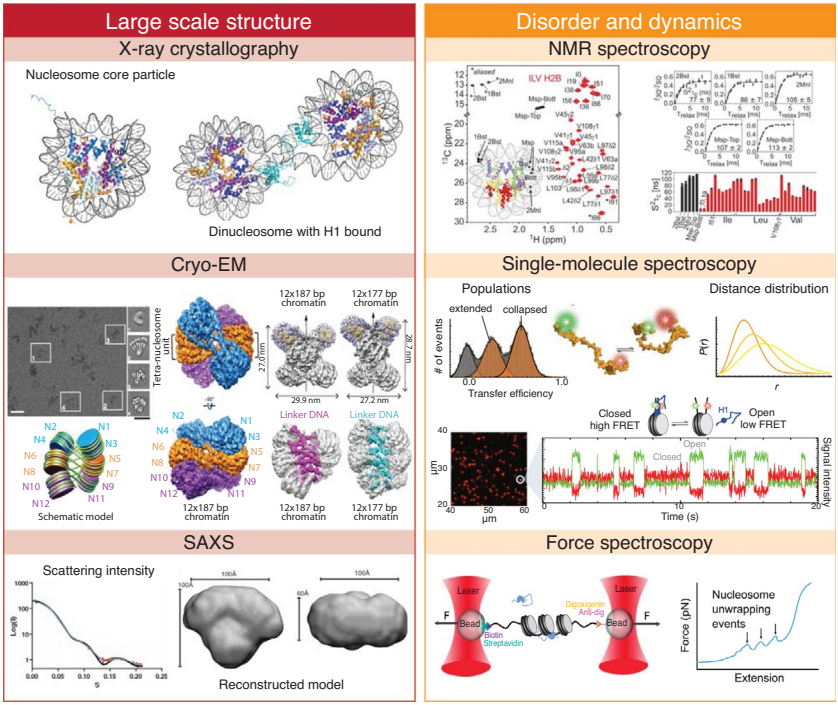

Fig. 2 Methods to study chromatin and intrinsically disordered proteins. Atomic-resolution structures can be determined from X-ray crystallography and cryo-EM while SAXS gives lower resolution information on the overall dimensions of molecules. NMR spectroscopy yields both atomic resolution three-dimensional models of biomolecules and their inter- and intramolecular dynamics, as well as providing residue-specific information on protein-protein or protein-DNA interactions. Single-molecule techniques can be used to study heterogeneous conformational ensembles, at equilibrium, intra- and intermolecular distance distributions, and reaction kinetics. The X-ray crystal structures show the nucleosome core particle[37] and a dinucleosome[38] with bound H1 (PDB codes 1AOI and 6LAB). *Cryo-EM, SAXS, and NMR data shown is reproduced from Song F, Chen P, Sun D, et al. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units.* Science. *2014;344:376–380, Abramov G, Velyvis A, Rennella E, Wong LE, Kay LE. A methyl-TROSY approach for NMR studies of high-molecular-weight DNA with application to the nucleosome core particle.* Proc Natl Acad Sci U S A. *2020;117 (23):12836–12846, Yang C, Van Der Woerd MJ, Muthurajan UM, Hansen JC, Luger K. Biophysical analysis and small-angle X-ray scattering-derived structures of MeCP2-nucleosome complexes.* Nucleic Acids Res. *2011;39:4122–4135 with permission.*

cameras and sophisticated image analysis,[35] which take advantage of graphics processing units (GPU) acceleration. Recent studies using these methods have supplied us with an impressive view of large molecular assemblies, such as a translating ribosome[36] and entire chromatin fibers (Fig. 2).[38–40]

However, biomolecular processes involving extensive disordered interactions lie outside the scope of current structural biology efforts and thus require different approaches to understand their molecular underpinnings.

2.4 Integrative modeling of disordered protein interactions

Modern research on structurally heterogeneous systems such as IDPs often combines multiple techniques to decipher their underlying physical mechanisms (Figs. 2 and 3). To study dynamic and disordered systems, techniques that can resolve conformational subpopulations in bulk have proven particularly useful. Nuclear magnetic resonance (NMR) spectroscopy can be used to obtain three-dimensional structural models of well-folded proteins, and it has also been extensively used to study protein dynamics and disorder,[25,46,47] even in live cells.[48] After assignments of chemical shifts, protein NMR gives residue-specific information on structure, stability, binding sites, and dynamics on a wide timescale.[49] Despite still being limited to relatively small to medium-sized systems for structure determination, NMR spectroscopy has revealed dynamic movements of the disordered core histones and their interactions, even within entire nucleosomes (Fig. 2).[50–52] Small-angle X-ray scattering (SAXS), the solution-state counterpart to X-ray crystallography, gives information on the shapes of molecules, including IDPs and their dynamic populations, often aided by computer simulations.[53,54] Especially relevant to DNA binding proteins, fluorescence recovery after photobleaching (FRAP) probes the mobility of fluorescently labeled proteins, inside the cell nucleus, and has been used to study the dynamic exchange of histone H1 between nucleosomes.[55] These and other methods have over the years been extraordinarily influential in shaping our perception of IDPs.

Techniques that probe the behavior of individual molecules, and thus access molecular distributions, are an attractive approach to understanding disordered interactions and have been used to complement traditional ensemble methods. Single-molecule spectroscopy, usually in combination with Förster resonance energy transfer (smFRET), has emerged in recent years as an exceedingly powerful technique to study structured and unstructured proteins, *in vitro* and in living cells.[56–58] SmFRET enables sensitive site-specific probing of the distance and dynamics between two or more

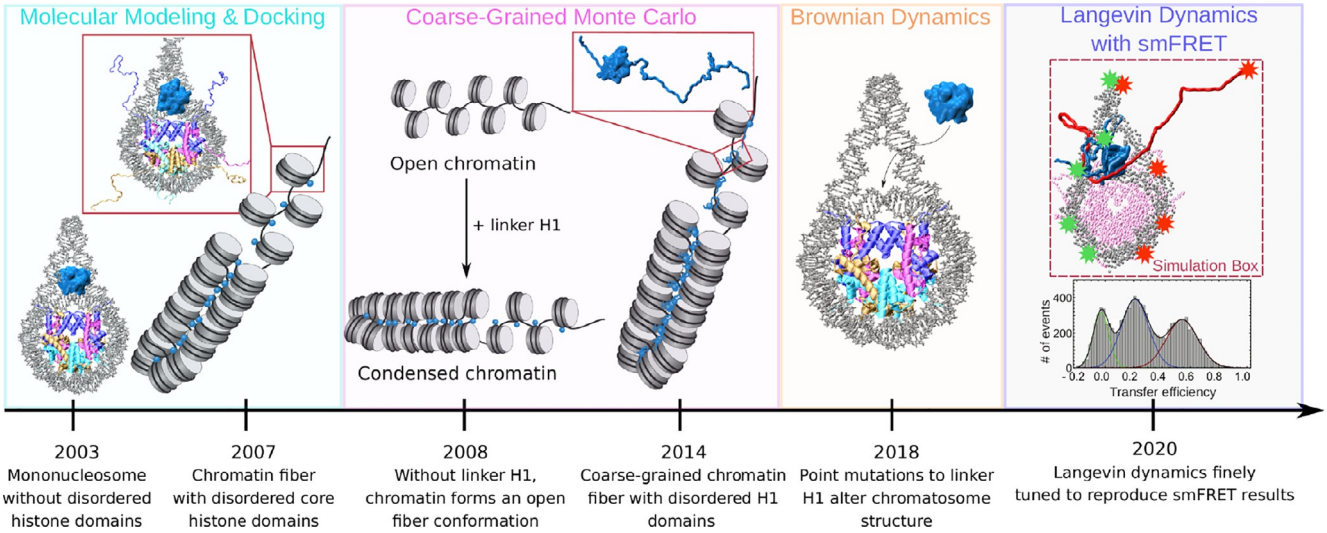

Fig. 3 Chronological overview of computational approaches adopted to study intrinsic disorder in chromatin topology and dynamics. Schematic illustration of the evolution of computational attempts used to investigate the interaction between proteins and DNA within the context of nucleosomes, chromatosomes and chromatin fibers. Early molecular modeling and docking investigated the binding of linker histone H1 to the nucleosome dyad and functioned as preliminary approaches to later attempts featuring molecular simulations on increasingly larger systems.[41,42] Coarse-grained Monte Carlo simulations were used to generate several possible topological arrangements of chromatin fibers with and without the linker histone H1.[43,44] Brownian dynamics,[45] which increases the amount of attainable sampling by scarifying internal motions, facilitated the understanding of how linker histone H1 diffuses toward and binds to the nucleosome dyad. Most recently, modeling and simulations featuring customized potentials, finely tuned to reproduce experimental findings, have provided a semi-quantitative overview of the disorder-mediated interactions between linker histone H1 and fully disordered chaperones involved in its displacement from the nucleosomal dyad.[17]

fluorescent dyes, *e.g.*, within a disordered region of a protein.[56] The rate of energy transfer between a donor and acceptor fluorophore is steeply dependent on the distance between them, where the useful range is typically on a convenient molecular scale of 1–10 nm. Importantly, smFRET can be applied to structurally heterogeneous systems one molecule at a time, avoiding the complication of ensemble-averaging, which can mask transient yet important molecular events. Confocal fluorescence microscopy offers a wide array of experiments that probe the thermodynamics and kinetics of biomolecular interactions through timescales covering 15 orders of magnitude,[59] as well as enabling high-resolution imaging in cells through fluorescence lifetime imaging (FLIM) and stimulated emission depletion (STED) microscopy.[60] Multi-parameter analysis of fluorescence intensity and photon timings allows quantitative investigation into molecular processes such as binding thermodynamics and kinetics, translational and intrachain diffusion, complex stoichiometries, misfolding and aggregation.[58] With total internal reflection fluorescence (TIRF), several surface-immobilized molecules can be excited and detected simultaneously, offering higher-throughput data analysis of FRET trajectories.[58] The versatility of the method has over the years provided new insights into fundamental biological processes such as DNA maintenance and repair, signaling, translation, transcription, and molecular transport.[57,61] On the flip side of the single-molecule coin are force spectroscopy techniques, such as optical tweezers or the atomic force microscope (AFM), that allow direct tethering and manipulation of individual proteins or DNA. Force spectroscopy can probe the microscopic molecular forces involved in biomolecular interactions and has enabled a fresh view into the energetics and mechanisms of protein-nucleosome interactions.[62,63] Single-molecule methods hold great promise for understanding chromatin interactions and when combined with technologies probing ensemble biophysics[64] and genome-wide approaches, these methods can provide a comprehensive view of dynamic and disordered protein-DNA interactions.[27]

In recent years, a plethora of computational techniques have been used alongside experiments to study chromatin and chromatin binding proteins (Fig. 3). By undertaking multiscale approaches, the finer molecular details of chromatin dynamics and interactions are now better understood. Early on, computational techniques were confined to molecular modeling and docking studies. All-atom molecular dynamics (MD) simulations are a gold standard technique for modeling biomolecular behavior, as they provide

atomic-resolved information on the movements and interactions of molecules in their given environments.[65] By solving Newton's equations of motion for each atom, modeled as a van der Waals sphere, interatomic forces and their corresponding energies are calculated using molecular mechanics force fields describing both bonded and non-bonded interactions, either in implicit or explicit solvent conditions. In implicit solvent models, the solvent is treated as a structureless continuum, thereby reducing the number of interacting particles and degrees of freedom. In contrast to explicit descriptions, where the presence of each solvent molecule is explicitly accounted, implicit models do not include solute-solvent interactions. Although all-atom simulations provide an unparalleled level of detail, simulating nucleosomal arrays in this manner is unreasonable because of the high computational cost leading to insufficient sampling.

Due to the high computational costs associated with simulating these large and complex systems, an understanding of nucleosome and chromatin organization has started from simply creating models of single nucleosomes[66] or chromatosomes[41] that would fit experimental constraints. In particular, such studies focused on elucidating the binding mechanism of linker histone H1 to nucleosomes and were able to provide an idea, resembling that observed in electron microscopy studies,[41] of how H1 and other histones shape the conformational dynamics of single and di-nucleosomes[66] as well as nucleosomal arrays composed of up to 100 nucleosomes,[42] reporting on the polymorphic nature of chromatin.

A different approach involves the use of Monte Carlo (MC) simulations of coarse-grained (CG) representations. By generating conformational states according to Boltzmann probabilities, MC can be used to sample a Boltzmann distribution of configurations. Coarse-graining offers a computationally less expensive approach. In CG models, groups of atoms are embedded into beads, thereby reducing the degrees of freedom and allowing efficient generation of conformations, without the explicit time dependence of MD simulations. Consequently, larger models, such as those encompassing entire chromatin fibers, can be simulated yet at the expense of fine molecular details. MC simulations of CG chromatin fiber models have revealed many aspects of chromatin compaction. In particular, studies carried out using CG-MC found that H1 is required in the formation of higher order chromatin structures[43] and that, without H1, chromosomal arrays adopt an open fiber conformation.[44] Additionally, chromatin structures with highly variable nucleosome repeat lengths (NRL) produce more

compact and uniform fibers,[67] while fibers with a longer NRL, corresponding to a more open chromatin structure, are likely to have a higher number of binding sites for chromatin binding proteins.[68]

Similar to MC, Brownian Dynamics (BD) simulations, which treat the simulated macromolecules as rigid bodies in implicit solvent, have been used to simplify the complexity of chromatosomes or chromatin. In BD simulations, the diffusion of solutes in a continuum solvent is simulated, and electrostatic interactions, which are particularly dominant in chromatin, are calculated by solving the Poisson-Boltzmann equation. BD approaches have successfully demonstrated the effects of sequence variation and PTMs on the H1-nucleosome ensemble. Due to lowering the complexity of the simulated system, docking simulations following the methodological paradigm of BD simulations, allowed a large number of PTMs to be considered, due to the increased computational efficiency associated with this technique.

Another approach to circumvent the computational constraints of all-atom MD is to coarse-grain (CG) the system. As discussed above, coarse-graining simplifies the system by reducing the degrees of freedom, making it an attractive technique for studies on chromatin fibers. In a recent study carried out by Watanabe et al., CG molecular dynamics simulations were used to create a model of an HP1α dimer bound to the histone H3 tail in a di-nucleosome complex.[69] In addition, reverse mapping of the CG structure was carried out upon the completion of the simulation to regain some of the lost atomic details. Within the context of CG models simulated by means of BD or MC, empirical potentials have added a certain level of integration with experimental data and specifically provided an experimentally-derived picture of the conformational ensembles of nucleosomes. Recently, smFRET and CG molecular simulations were tightly coupled in a complementary approach where the distance of multiple sites across protein-DNA within a nucleosome were mapped by smFRET and matched closely by the fine tuning of a single force field parameter describing van der Waals interactions between modeled beads. Integrative modeling has immense potential to deliver finer details of complex molecular systems and has already begun to uncover the physical principles governing IDP interactions.[64,70,71] We now move on to describe some recent work on disordered interactions with chromatin, highlighting new insights that biophysical methods have yielded on the role IDPs and IDRs in the nuclear environment.

3. Disordered interactions with nucleosomes

3.1 Nucleosome architectural proteins

3.1.1 Linker histone H1

Many proteins involved in generating and maintaining the overall nucleosomal architecture contain long disordered regions. H1 is involved in chromatin condensation through stabilization of compact chromatin structures, and thus functions generally as a transcriptional repressor.[72,73] The polypeptide sequence is highly positively charged with two long disordered regions (N-terminal domain, NTD; C-terminal domain, CTD) flanking a small folded globular domain.[74] The globular domain of H1 is known to bind to the dyad axis of the nucleosome (Fig. 1), thus interacting with the nucleosomal core and both entry- and exit DNA linkers.[13,75] By binding to the nucleosome dyad, the H1 tails are free to form non-specific electrostatic interactions with linker DNA to minimize charge repulsion, thereby facilitating chromatin condensation. Therefore, although the binding mode is facilitated through the structured domain of H1, function is largely conferred through the disordered tails and an on-dyad binding mode may provide the freedom required for the H1 CTD to interact with one or both linker DNA arms. However, the conformational distributions of the disordered regions of H1 on the nucleosome have been more difficult to elucidate because of their pronounced dynamics and seeming lack of persistent structure.[76] Importantly, single point mutations on linker H1 significantly affect chromatosome structure, indicating that small changes may alter the overarching chromatin structure and, consequently, transcriptional regulation.[45] It has been suggested that H1 draws the two linker arms together, thereby reducing their mobility, and introducing a strong degree of asymmetry to the nucleosome.[13] Recent integrative studies of full-length H1 in complex with nucleosomes gave insight into the behavior of the long disordered tails.[17] The authors studied binding of human linker histone H1 to reconstituted nucleosomes using confocal single molecule spectroscopy and CG molecular simulations. Fluorescent labeling of the approximately 100 residue-long disordered CTD of H1 revealed that it becomes considerably more compact in complex with the nucleosome. This can be explained by screening of H1´s positive charges by the negatively charged nucleosomal DNA, which otherwise renders H1 highly expanded due to charge repulsion. Labeling on the terminal end of nucleosomal linker DNA arms and

addition of unlabeled H1, resulted in the expected closure of the linker DNA arms in the H1-bound nucleosome (in agreement with a crystal structure of the chromatosome). Comprehensive mapping of FRET efficiencies within the H1-nucleosome complex combined with nanosecond fluorescence correlation spectroscopy (nsFCS) showed that H1 lacks persistent structure and is extremely dynamic on the nucleosome, displaying sub-μs chain reconfiguration times. Using a simple CG model describing the system in terms of non-specific short-range and electrostatic interactions, combined with existing structural information on the nucleosome and H1´s globular domain, the entire complex was simulated and the distances between the corresponding FRET pair locations back-calculated. After tuning the only free parameter in the model—the inter-bead interaction strength that was set globally for all beads—the simulation was able to capture the dynamic conformations of H1 on the nucleosome with high accuracy. Simulations were indeed in excellent agreement with the FRET-derived distances (a total of 57 FRET pairs), illustrating well how closely simulations can reconstruct experimentally determined parameters even in very complex systems.

The presence and conservation of multiple H1 variants within cells suggests that different variants may be linked to specific cellular functions. Within each variant, the structured globular domain shows the highest degree of conservation, while disordered tails are, expectedly, more variable. However, when H1 tail regions from different species are compared, a high degree of conservation is observed between orthologs. For example, human H1.4 and its mouse ortholog, H1e, share 93.5% sequence similarity,[77] indicating that H1 tail regions may confer a high degree of functional selectivity in cells.

In addition to the highly basic charge in the linker H1 CTD, recent studies suggest a direct link between CTD length and chromatin affinity. FRAP experiments on human linker histones found that H1 variants with shorter CTD tails, such as H1.1 and H1.2, have rapid recovery times compared to variants with longer CTD tails, such as H1.4 and H1.5. Moreover, longer histone H1 tails were found to have two or more cyclin-dependent kinase (CDK)-dependent S/T-P-X-K phosphorylation motifs. Therefore, recovery times may be dependent on the density of lysine residues, the CTD length and the distribution of DNA-binding S/T-P-X-K motifs.[78] Because the H1 CTD directly interacts with the linker DNA, each variant will have a different effect on NRL. A higher degree of chromatin folding will likely be achieved if neutralization occurs across the chromatin fiber. Although chromatin condensation is not a direct consequence of linker

histone binding, it does stabilize higher order chromatin structures to an extent that depends on the corresponding H1 variant. In general, the affinity of H1 for chromatin increases with its compacting properties. In agreement with this observation, H1.0, H1.4 and H1.5 have a longer CTD and were found to stabilize higher order chromatin.

3.1.2 Post-translational modifications of linker histones

Linker histones are subject to a variety of PTMs, in both their IDRs, adding a large degree of compositional complexity to this protein family. The presence of multiple PTM sites in the H1 IDRs enables a number of regulatory mechanisms for H1 and finely regulates the affinity of each H1 variant for chromatin. The most prominent PTM for H1 is certainly phosphorylation, which occurs in a highly complex and dynamic fashion.[79] Phosphorylation mainly occurs in the CTD where S/T-P-X-K motifs (X is any amino-acid) are recognized by CDKs.[73] Although counterintuitive, H1 phosphorylation can trigger both chromatin expansion and contraction, based on the progression of the cell cycle. Such effects are likely to be a result of conformational rearrangements within H1, arising from site-specific modifications.[80] Phosphorylation levels are the lowest during the G1 phase, rise during the S phase and peak during mitosis, followed by a sharp decrease in the telophase.[11] *In vivo* studies showed that serine residues in H1.4 are generally modified during G1 and S phases, while threonine is phosphorylated in mitosis,[81] outlining the cell cycle dependence of phosphorylation. CDK1 and Cyclin B are primarily responsible for H1 phosphorylation during the mitotic phase. Additionally, recent studies suggested that several kinases phosphorylate the H1 NTD.[82–84]

The conversion of lysine into its methylated analogs (methyllysine, di-methyllysine or tri-methyllysine) is another important modification that can compete with or complement phosphorylation on a functional basis. For instance, the methylation of lysine 26 in the H1.4 NTD recruits heterochromatin protein-1 (HP1), resulting in heterochromatin formation,[85] and is controlled by a phospho-switch: when H1.4 is phosphorylated at serine 17, the interaction between HP1 and methyllysine 26 is inhibited, demonstrating the importance of crosstalk between PTMs.[85] In the cell, H1 phosphorylation is CDK2 dependent and is required for progression through the S-phase. Because CDK2 colocalizes with replication sites and H1 is crucial in the formation of higher order chromatin, CDK2 recruitment to replication foci by Cdc45 may result in H1 phosphorylation and drive fork progression,[86] linking H1 phosphorylation and active transcription.

Methylation of the H1 NTD and CTD may elicit variant-specific cellular responses *in vivo*. For instance, the two predominant human H1 variants, H1.2 and H1.4,[87] are methylated differently by the same methyltransferases.[88]

Interestingly, the methylation of H1.4 lysine 26, a highly conserved PTM in vertebrates, creates HP1 binding conditions, likely because H1.4 lysine 26 is part of an "ARKS" motif[85,88]: a conserved motif assumed to have a regulatory function in heterochromatin. Therefore, a link may be present between chromatin compaction and lysine 26 methylation.

Like phosphorylation, acetylation of the NTD leads to both heterochromatin formation and activation of transcription. Acetylation of H1.4 lysine 26 is related to the formation of facultative heterochromatin, which forms parts of the genome not shared across cell types and usually contains poorly expressed genes, which are task specific and mostly associated with cellular differentiation. Deacetylation by SIRT1, on the other hand, results in the formation of repressive heterochromatin.[89] The presence of an acetyl group on lysine 26 prevents methylation and subsequent recruitment of HP1, providing an additional level of regulation. Interestingly, an *in vivo* study using T47D cells expressing a lysine 26 to alanine H1.4 mutant, reported defects in gene regulation and cell proliferation, compared to wild-type H1.4.[90] Additionally, acetylation of H1.4 lysine 34 is also associated with transcription activation *in vivo*. In this position acetylation is, however, suggested to reduce H1-chromatin affinity and recruit TAF1; a subunit transcription factor TFIID.[91]

3.1.3 Core histones

The core histones, which are the main structural support of nucleosomes, contain relatively short yet crucial IDRs when compared to the linker histone. Core histones form an octamer around which DNA is wrapped in the initial stages of chromatin condensation.[73] Each core histone shares a common histone-fold domain of three helices connected by two loop regions. To complete the octamer, each core histone homodimerizes, followed by the formation of specific H2A-H2B and H3-H4 heterodimers to create the "handshake" shaped core[92,93] (Fig. 1). In addition to the structured domains, each core histone has an intrinsically disordered, solvent-exposed N-terminal tail; with only H2A having an additional C-terminal tail.[94] Like H1, these tails are enriched with highly basic residues that form electrostatic interactions with nucleosomal DNA, linker DNA and acidic patches on neighboring nucleosomes. Such interactions are believed to stabilize the histone-DNA and nucleosome-nucleosome associations.

Moreover, the dynamic nature of the tail regions makes them a target for PTMs and subsequent recruitment of histone chaperones, architectural binding proteins and chromatin remodelers.[95] The inter- and intramolecular contacts between the nucleosome and disordered core histone tails are crucial in the formation of the nucleosome core particle (NCP) and for the stabilization of higher order chromatin structures.

The intrinsically disordered tails of core histones have also been implicated in the formation of higher order chromatin structures and chromatin condensation through inter-nucleosomal interactions.[96] In the first nucleosome crystal structure published in 1997, an inter-nucleosome interaction between the N-terminal H4 tail and an acidic patch on the H2A/H2B dimer interface of a neighboring nucleosome was identified.[37] As for H1, the *in vivo* functions of core histones are complex and their ability to modulate transcription is largely dependent on their PTMs.[77] In the nucleosome, DNA accessibility is controlled by transient unwrapping from the NCP; a process that is modulated by each core histone to different degrees.[97,98] While the histone H3 tail suppresses nucleosome unwinding, the histone H4 tail enhances it.[99] Consequently, PTMs in the histone tails are especially important in their role of modulating protein-nucleosome interactions and regulating unwrapping.

Acetylation is an abundant modification in core histone tails, affecting chromatin compaction *via* the neutralization of positive charges.[20] Such effects have been demonstrated *in vitro*, where compaction of nucleosomal arrays required residues 14–19 in the NTD of the human histone H4[100] and the acetylation of lysine 16 prevents array compaction.[101] Moreover, acetylation is likely to reduce the electrostatic cross-talk between the DNA and the histone tails, decreasing the force required to unwrap nucleosomes[102] and increasing DNA accessibility to transcription factors and other modifying enzymes.[103] In line with this notion, histone acetylation has been shown to be strongly associated with transcription.[104] Additionally, many histone acetyltransferases interact with tri-methylated lysine 4 on H3; a modification associated with transcriptional activation.[20] Relevantly, *in vivo*, inhibition of transcription was shown to result in rapid histone deacetylation in mouse embryonic cells, indicating that much of histone acetylation occurs as a result of transcription.[105]

Phosphorylation of histones predominantly occurs in the intrinsically disordered NTD.[106] For instance, Aurora B kinase is known to phosphorylate H3 serine 10[107] and serine 28[108] during the mitotic phase *in vivo*, however, there is no evidence of both modifications being present on a single

histone tail.[109] Nevertheless, phosphorylation of the histone H3 tails is likely required for cell cycle progression. Importantly, the Aurora B kinase is overexpressed in a number of human cancers,[110] suggesting that phosphorylation plays a significant role in nucleosome availability to transcription.

Methylation is another predominant modification in core histone tails and for the cross-talk between different PTMs. For example, H4 arginine 3 methylation is recognized by p300; the acetyltransferase responsible for acetylation in histone H4. Methylation in this position is indeed an important PTM in gene transcription, as it is required for the subsequent acetylation that reduces electrostatic repulsion in the chromatin fibers.[111] In contrast, methylation of H3 arginine 8 is associated with gene repression, outlining the importance of site specificity in PTMs.[112]

3.1.4 HP1 proteins

Gene expression within the context of heterochromatin is facilitated by a series of important, conserved proteins called heterochromatin proteins (HP). These are fundamental units of chromatin packing that can be subdivided into families, with HP1 being the dominant family composed of three isoforms in humans—HP1α, HP1β and HP1γ—, all of which have two highly conserved structured domains; the amino-terminal chromo domain (CD) and the carboxyl chromo shadow domain (CSD). The structured domains are separated by a disordered hinge region (HR), of varying length across paralogs. Additionally, shorter intrinsically disordered extensions are present at the N- and C-termini of HP1.[113] HP1 is a major component of heterochromatin and is involved in the regulation of DNA-mediated processes including heterochromatin formation, stabilization of telomeres and gene silencing in pericentric heterochromatin.[69,114]

In general terms, HPs are multivalent, structural chromatin effectors[115] that cause transcriptional repression by recognizing and binding di- or tri-methylated lysine 9 in histone H3 (H3K9me2/3) *via* the CD,[116] while remaining highly dynamic. Methylation of H3 provides an epigenetic mark, suitable for the hydrophobic binding pocket created by the CD. Despite the high degree of specificity between HP1 and H3K9me2/3, the binding affinity spans widely depending on the paralog.[117] Varying affinity is believed to provide a dynamic range in which HP1 paralogs are able to elicit different cellular functions. The dynamic nature of HP1α has recently been probed using *in vitro* techniques. By employing a chemically defined assay, well suited to cellular measurements, it was found that HP1α residence time increases with H3K9me3 density, due to rapid re-binding of dissociated

factors on neighboring sites. Moreover, dimeric HP1α exhibited accelerated association rates; a key feature of effector multivalency, allowing fast and efficient binding in a competitive environment.[115]

PTMs of HPs also play a fundamental role in regulating affinity. For instance, HP1α phosphorylation further strengthens its multivalency, while simultaneously reducing DNA binding, ultimately increasing HP1α residence time.[118] Phosphorylation of serine residues in the disordered NTD in mouse HP1α was found to increase CD-H3K9me2/3 affinity such that the overall affinity close to that of mouse HP1β and HP1γ.[117] Moreover, in HP1β and HP1γ the serine residues are replaced by glutamate. Such findings suggest that CD-H3K9me2/3 affinity may be partially modulated by charge differences in distal regions. Therefore, regulating charge *via* phosphorylation of the serine residues within the NTD in HP1α may contribute to the protein binding to H3. In turn, this interaction may impact the activity of kinases or phosphatases, increasing or decreasing binding.

Between the HP1 paralogs, both the underlying amino acid sequence and length of the HR are variable and such differences may control localization and function.[113] For instance, there are 41 and 36 residues in the HP1α and HP1β hinge regions, respectively, and both variants localize to heterochromatic regions[119,120] mediating transcriptional gene silencing.[121] Comparatively, HP1γ, where the HR has only 31 residues, localizes to euchromatin[119,120] and plays a role in transcriptional elongation and RNA processing.[121] The molecular basis for functional divergence is suggested to arise from the non-conserved residues in the HR, since the positively charged domains (KRK and KKK) are conserved across all three variants.[121] These domains are crucial for the specificity of HP1-H3K9me2/3 binding *in vitro*[122] and for intranuclear localization *in vivo*.[123] PTMs in the hinge regions have been shown to affect HP1 functionality.[113,124] Phosphorylation of serine 83 in the hinge region of HP1γ increases its interactions with Ku70, a DNA repair protein, thereby increasing its localization to euchromatin.[113] Taken together, this may suggest that modifications in the disordered HR of HP1 paralogs are able to elicit specific cellular functions.

Recent studies show that HP1 proteins play an important role in heterochromatin by interacting with histones H3 and H4 and methyltransferase enzymes.[106,125] The binding of the HP1 CD to poly-methylated H3 lysine 9 (H3K9me2/3) and H1.4K26me[85] triggers a silencing mechanism, resulting in the formation of heterochromatin.[116] Moreover, this interaction may be influenced by PTMs,[126] especially those in the intrinsically

disordered regions. In particular, phosphorylation of HP1α NTD poly-serine stretch 11–14, increases chromatin binding affinity by reducing tail flexibility in human and mouse cells.[117,127] Phosphorylation changes the conformation of the NTD, such that neighboring acidic residues (15E-DEE-E19) are able to interact with basic residues surrounding H3K9 (8R-Kme-STGGKAPR-K18).[128] Addition of the negatively charged residues formed upon phosphorylation results in repulsion, causing the HP1α NTD to behave as an extended intrinsically disordered region, in turn allowing the CD to dynamically bind H3K9me2/3.[128] Interestingly, in HP1β and HP1γ the residues corresponding to the poly-serine stretch of HP1α (12E-VL-E15 and 21K-VE-E24, respectively) are partially negatively charged.[117]

3.2 Intrinsically disordered proteins that interact or compete with linker histone H1

The state of chromatin compaction is tightly linked to the presence of H1. Therefore, the cell has evolved various regulatory mechanisms to actively remove H1 from nucleosomes. One such mechanism is proteins that compete with H1 for binding to the nucleosome or otherwise lead to its eviction. As mentioned above, H1 is highly disordered outside of the globular domain, a feature that is commonly shared among the diverse H1 competitors outlined here.

3.2.1 Protamines

Protamines are short (25–100 residues), highly basic, and disordered nuclear proteins[129,130] suggested to have evolved from histone H1.[131] Protamines replace core histones during the last stages of male germ terminal differentiation of spermiogenesis, where they are found to be the major packing units of DNA (Fig. 4B).[132] Most mammals have only one gene coding for protamine 1 (PMR1 or P1) which is expressed in spermatids as a mature protein[133] and is responsible for chromatin condensation in sperm. However, some mammals, including humans and mice, have a second protamine, PMR2 or P2. Protamines from the protamine 2 family are longer compared to P1 and are generated by proteolytic cleavage of a precursor. DNA packed by protamines in mature sperm cells is transcriptionally inactive and forms higher order structures vital for normal sperm function.[132,134]

Protamine packaging of DNA has been studied with chemical and physical studies of both natural sperm chromatin and synthetic DNA.[135] Earlier studies demonstrated that protamines can precipitate DNA from both

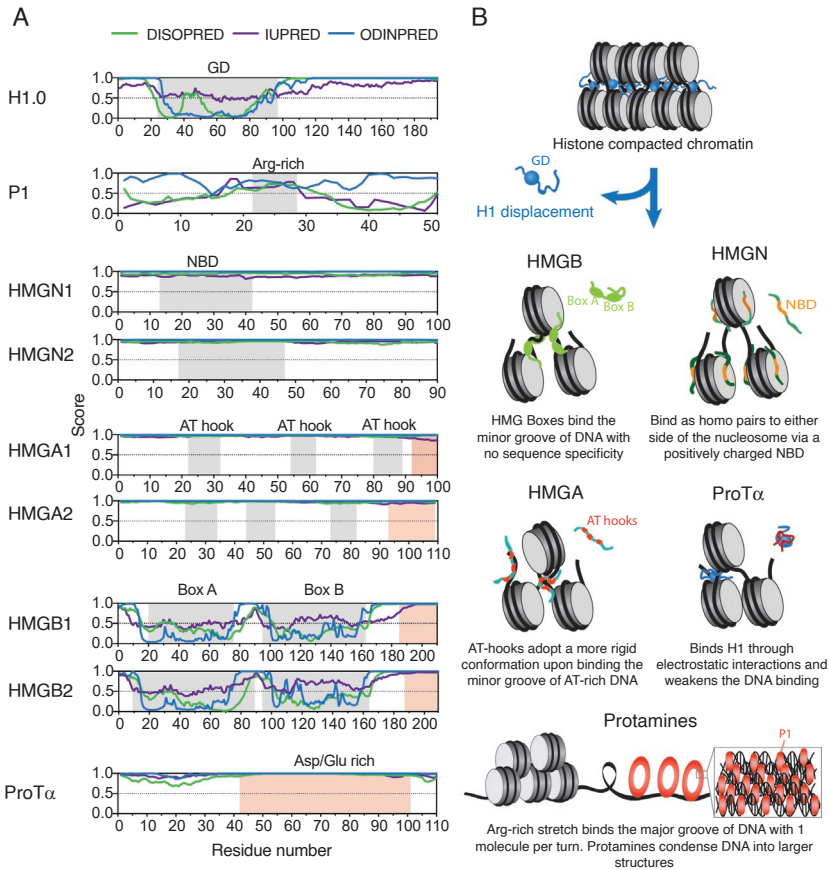

Fig. 4 Disordered H1 competitors and their nucleosome binding modes. (A) Disorder predictions of H1 and competitors are shown using three predictors. Domains are highlighted in gray: Linker histone H1.0 contains a folded globular domain, HMGN1 and 2 each have an NBD, HMGA1 and 2 each contain three AT-hooks, and HMGB1 and 2 are characterized by two folded Box domains. Acidic stretches are indicated in red. (B) Schematic illustration of the binding modes of H1 competitors in a nucleosomal context. The tightly packed structure of chromatin with bound H1 is remodeled by the disordered competitors, through eviction of H1. HMGB, HMGA, HMGN, ProTα and protamines all have distinct binding modes and their domains adopt different degrees of disorder essential to association with the nucleosomes.

assembled chromatin and native chromatin extracted from calf thymus tissue in a concentration dependent manner. By analyzing the supernatant composition using gel electrophoresis after addition of protamine to a chromatin solution, H1 was discovered to be the first histone to appear in solution. However, the release of H1 was slower in native chromatin compared to

reconstituted nucleosomes although somewhat affected by sample preparation, suggesting that the mechanistic picture of H1 competition and histone-protamine transition is more intricate.[136] Later work unveiled that protamines replace histones though a complex and progressive transition mechanism.[134] After meiosis in spermiogenesis, the canonical histones are replaced with testis-specific histones, and subsequently replaced by transition proteins that causes alteration in DNA structure. Many protamine molecules then bind and reorganize DNA into tightly packed structures. Thus, the conversion from histone-packed to protamine-packed chromatin is a complex interplay between different DNA-binders, regulated by PTMs including hyperacetylation of histones, in steps that are crucial for the correct progression of chromatin maturation and spermiogenesis.

The primary structure of protamines is characterized by a conserved arginine-rich core functioning as a DNA anchor[135] and cysteine-rich N- and C-termini. The overall dimensions of different protamines have been predicted by simulations to be controlled by their net charge per residue, which shifts their conformational ensembles from collapsed globule to coil-like.[137] The central arginine-rich region provides a high net positive charge that facilitates strong binding to DNA. Protamines wrap around DNA in the major groove[138] and bind with one protamine molecule per turn of DNA helix (Fig. 4).[139] Even though protamines are known to be disordered when free in solution, it is challenging to probe conformational changes of the individual molecules within the context of the large chromatin structures because of the large number of protamines associating with chromatin. Potential disorder-to-order transitions upon DNA binding are yet to be clearly demonstrated experimentally for protamines. However, many cysteines partake in multiple intra- and intermolecular disulfide bridges that provide rigidity, which is essential to stabilizing the structure of sperm cell chromatin.[140,141] Single-molecule studies demonstrated that protamines can bend DNA into loops through multiple steps[142] leading to the formation of higher order structures similar to those induced by H1 DNA packing,[143] which suggests a common pathway for positively charged IDPs in chromatin condensation.

Protamines contain several conserved phosphorylation sites and have been shown to undergo various PTMs. These include phosphorylation, acetylation and methylation, which have been detected in protamines of mice sperm using peptide-based tandem mass spectrometry.[144] The data indicated that methylation, acetylation and phosphorylation do not occur at the same time on a single protamine, suggesting a complex network of PTMs affecting the epigenetic landscape of sperm cells.[144] However, the

exact role of PTMs on the conformational ensembles of protamines and within the context of histone eviction in sperm remains unknown.

3.2.2 HMG proteins

High-mobility group (HMG) proteins belong to a family of disordered architectural transcription factors known to interact with nucleosomes and, together with H1, were the first nuclear proteins known to affect the structure of chromatin.[145] They were first discovered in isolated chromatin together with histones and named due to their unusually high electrophoretic mobility.[146] HMG proteins modulate local and global chromatin architecture by inducing formation of distorted DNA structures and promoting chromatin decompaction. The decompaction in turn enhances various DNA-dependent activities such as transcription, replication, and repair. HMG proteins are divided into three families—HMGA, HMGB and HMGN—, depending on their structural and functional properties, and we discuss here the role of their intrinsic disorder in nucleosome binding and H1 competition.[145,147]

The HMGN subgroup has five members; HMGN1 and HMGN2 were the first to be discovered, whereas HMGN3-5 were identified later. HMGNs are fully disordered (Fig. 4A) and bind the nucleosomal structure with high specificity in pairs to form complexes containing two molecules of either HMGN1 or HMGN2.[148–150] These proteins are characterized by a positively charged and conserved nucleosome binding domain (NBD), a nuclear localization signal (NLS), and an acidic C-terminal chromatin regulatory domain (CHUD) involved in modulating acetylation of histones.[147,151] HMGN proteins recognize generic nucleosome structures without specificity for DNA sequence or histones *via* the ~30 amino acid-long NBD.[152] The domain contains the canonical motif *RRSARLSA* which serves as an anchoring point on the nucleosome to a negatively charged patch formed by the H2A-H2B dimer surface.[153] The C-terminal domain of the HMGN protein interacts with the DNA in the two major grooves flanking the nucleosome dyad axis and is in close proximity to the N-terminal tail of histone H3.[149,154,155] HMGN also influences H3 phosphorylation and acetylation, by inducing local structural change that alters the accessibility of enzymes and thus the equilibrium of nucleosomal PTMs.[156,157] HMGNs are themselves modulated through phosphorylation,[157] which decreases the affinity to chromatin and allows kinases to access histone H3.[156] HMGNs also have many lysines placed in the NBD that are acetylated, which leads to less efficient binding to nucleosomes.[158]

NMR studies using paramagnetic relaxation enhancement (PRE) experiments, which probe long-range interactions, and methyl-labeled histones in assembled nucleosomes, showed that, upon addition of HMGN, an interaction occurs between an arginine-rich region of the NBD and the folded core of the H2A/H2B dimer.[155] In addition, it was demonstrated that the several lysine residues in the C-terminal end of the NBD have an affinity for DNA non-specifically. The experimental data was then used as restraints to model the binding orientation of the NBD on the nucleosome. In the calculated structural model, the N-terminal end of the two NBDs stacked on each side of the nucleosome are predicted to bind the H2A/H2B acidic patch, while the lys-containing ends associate with DNA near the exit/entry point.[155] Due to the orientation of the NBD on the nucleosome, the disordered C-terminal tail was expected to be located where H1 associates with nucleosomes. This suggests that the chromatin decompaction function of HMGN is a result of disrupting H1 binding to DNA and the core histone tails.

Surprisingly, native gel-shift assays and cross-linking studies showed that HMGN1 binds nucleosomes already bound to H1, without interfering with the specific contacts made between the H1 globular domain and the nucleosome.[159] This observation agreed with previous MNase digestion results on mammalian chromatin, which reported nucleosomes bound to two HMGN proteins and H1.[160] This implies that binding of HMGN proteins to nucleosomes engenders a dynamic rearrangement of H1 interactions leading to modulation of chromatin structure. However, the molecular mechanism is still unclear and whether HMGN folds upon binding to the nucleosomes or if the two bound HMGN remain fully unstructured in the complex, is yet to be determined.

HMGB proteins (HMGB1 and HMGB2) have two structurally conserved DNA binding domains (DBDs, Box A and B) that each fold into three helices when in complex with DNA.[145] A disorder-to-order transition observed by NMR spectroscopy on HMG domains from Sox-proteins,[161] demonstrated that HMG box domains retain a flexible structure in solution that folds upon DNA binding. In addition, HMGB proteins contain a disordered C-terminal region of ~30 amino acids (Fig. 4A) enriched with acidic residues. The Box DBDs of HMGB bind with low affinity to single-stranded, linear duplex, and supercoiled DNA,[162,163] but have a preference for bent and distorted DNA[164–166] which is increased upon acetylation.[167] Based on structures of closely related HMG-Box-DNA structures, the box

domains distort DNA through intercalation of bulky hydrophobic amino acid residues into the DNA minor groove, resulting in bending the molecule toward its major groove (reviewed in[165]).

The acidic C-terminal region forms a flexible extended structure which was characterized by NMR spectroscopy and demonstrated narrow dispersion in ^{1}H-^{15}N-HSQC typical for IDPs.[168] The flexible tail is involved in dynamic intramolecular interactions, with the highest affinity for DBD Box B. SAXS and NMR studies, including PRE measurements, suggested that the C-terminal tail promotes a more compact conformation where the two basic boxes get closer to each other.[169] Apart from intramolecular interactions of the acidic tail with the HMG-boxes, the acidic C-terminal tail of HMGB1 is also engaged in intermolecular interactions with other proteins. Notably the HMGB tail associates with histone H3, an interaction suggested to modulate the biological functions of HMGB proteins.[170] MNase digestion data suggested that HMGB protects linker DNA on one side of the NCP at the entry/exit of nucleosomes opposite to the linker histone H1 binding site.[171] As observed for other HMG proteins, early studies using chromatin fractionation experiments reported strongly enriched HMGB1 and 2 in H1-depleted fractions of salt-soluble chromatin.[172] FRAP experiments showed that HMGB also enhances H1 mobility in cells, indicating its ability to displace H1 from chromatin.[173] Studies of HMGB and H1 interactions by chemical cross linking and gel filtration experiments showed that they form a 1:1 complex. The complex persists at physiological ionic strength, where it was reported by NMR spectroscopy that H1 binds through its basic C-terminal domain to the acidic tail of HMGB1, disrupting its interaction with HMG boxes. A consequence of this interaction is enhanced DNA binding and bending by HMGB1, followed by a lowered affinity of H1 for DNA.[174] This might facilitate H1 eviction in a chromatin context and supports the data showing increased H1 mobility in cells in presence of HMBG.[173]

As outlined above, several lines of evidence support that the disordered C-terminus of HMGB1 has a crucial role in its function, including orchestrating many different interaction partners (review in[175]) and modulating chromatin structure. HMGB function is also regulated by PTMs like acetylation, phosphorylation, methylation, as well as its oxidative state as formation of a disulfide bridge in Box A leads to reduced H1 displacement from hemicatenated DNA loops.[176] Most PTMs found or predicted in HMGBs are placed in the folded box domains, but several acetylation sites have been

identified in disordered stretches within NLS regions. Although acetylation within the Box domain is known to affect DNA binding affinity, more studies point toward PTMs having a large impact on controlling the nuclear localization and export of HMGBs. However, how PTMs affect the conformational ensemble of the acidic disordered tail of HMGB and its interaction with H1 has yet to be fully defined.

The third class of highly disordered HMG chromatin binders are the *HMGA proteins (previously named HMG-I(Y))*. There are two genes coding for HMGA proteins, HMGA1 (and its splicing variants HMGA1a, b and c) and HMGA2, characterized by their very short DNA-binding AT-hook motifs. HMGA proteins are fully disordered in solution in absence of DNA as shown by biophysical techniques such as circular dichroism[177] and NMR spectroscopy.[178] An NMR study reported that the AT-hook DBD of HMGA transits from disordered to a well-defined crescent-shaped configuration upon binding to the minor groove of short AT-rich DNA stretches.[179] The specificity for AT-rich DNA regions was also demonstrated by a PCR-based systematic evolution of ligands by exponential enrichment (SELEX) approach, identifying nucleotide consensus sequences with two AT-rich stretches of 5–6 base pairs separated by four GC-rich base pairs.[180] EMSA studies show that HMGA binds isolated nucleosome particles with much higher affinity than to naked DNA.[181,182] It was previously suggested that HMGA1 also associates with core histones based on DNase footprinting and chemical cross-linking studies,[183] possibly explaining the preference for nucleosomal DNA.

HMGA1 proteins were found to co-localize with histone H1 at AT-rich DNA stretches called scaffold attachment regions (SARs) in mammalian cells.[184] T7 polymerase assays in combination with DNA binding assays showed that HMGA can compete with H1 on SAR, and even redistribute H1 onto non-SAR DNA.[183] Purification of HMGA and H1 from HeLa cell chromatin also demonstrated that HMGA is strongly enriched in H1-depleted fractions of active chromatin.[183] In a study of chromatin condensation in neural precursor cells of mice, it was found that HMGA proteins are essential for chromatin opening in the early developmental stage. Overexpression of either HMGA1a or HMGA2 in cells increase the sensitivity to MNase digestion of chromatin from extracted nuclei, whereas depletion of HMGA mRNA led to reduction in MNase digested DNA. This clear effect of HMGA proteins on MNase digestion of extracted DNA from nuclei suggests that HMGA induce chromatin opening and accessibility.[185] The chromatin was more resistant to digestion in absence

of HMGA, which supports the notion that chromatin becomes more accessible in presence of HMGA, suggesting an inhibition of H1 driven compaction. Another work observed increased H1 mobility caused by HMGA by measuring FRAP in cells expressing GFP-H1 and microinjected with purified HMGA into the cytoplasm.[173] The apparent H1 displacement in the different studies was due to competition for chromatin binding sites, since HMGA mutants incapable of binding DNA did not increase H1 mobility in a similar manner nor compete with H1 for SAR binding.[173,183] This strongly indicates that HMGA competes with H1 on chromatin resulting in destabilization of higher order chromatin structure.

Like the other HMG proteins, HMGA undergoes various PTMs which have been extensively studied. In fact, HMGA1 proteins are among the most phosphorylated proteins in the nucleus by the action of various kinases like cdc2, protein kinase C (PKC), and casein kinase II (CK2).[157] Phosphorylation of HMGA leads to considerably lower affinity toward DNA, in part because two of the main phosphorylation sites are near the positively charged AT hooks and thus disrupt their binding to the negatively charged DNA. The acidic C-terminal tail of HMGA also undergoes phosphorylations *in vivo* which can lead to a conformational change.[186] NOE measurements and 1D proton spectra of C-terminally phosphorylated HMGA indicated a more rigid structure compared to the native and free HMGA which is fully disordered. Pull-down assays of truncated HMGA containing AT-hooks with the acidic C-terminal peptides showed a clear interaction to nucleosomes driven by electrostatics as the affinity increased with the number of phosphorylations in the C-terminal peptides. This supports the hypothesis that the phosphorylated acidic tail folds back onto positively charged clusters on the HMGA and through charge neutralization impairs binding mediated by these Arg/Lys-rich regions.

To summarize, all HMG proteins compete with H1 for chromatin binding sites (in a dose dependent fashion) although each HMG subfamily has distinct effects on the interaction of H1 with chromatin (Fig. 4).[152] HMG proteins all contain disordered regions—HMGA and HMGN being completely disordered in absence of DNA—but they have different structural features and folded domains. This results in distinct modes of action in their modulation of chromatin structures, although a common feature appears to be recognition of DNA conformation. Even though the regions where the HMG proteins bind nucleosomes is known, the sequence of events leading to H1 eviction from nucleosomes is still not fully clear. It has also been suggested that the different classes of HMG proteins can

weaken H1 binding cooperatively without competing with each other, hinting that they distinctly affect H1 binding to the nucleosome.[173] Overall, the HMG proteins are part of a dynamic and elaborate interaction network that leads to H1 displacement, where disorder plays a fundamental role.

3.2.3 FoxA1

Forkhead box A (FoxA) transcription factor (previously called HNF-3) is part of a group of transcription factors evolutionary conserved in eukaryotes and is crucial in regulation of biological processes such as cell development, signal transduction, cell differentiation, and regeneration. FoxA1 is as a so-called pioneer transcription factor due to its ability to engage target sites on nucleosomal DNA.[187,188] FoxA1 is disordered outside of a highly conserved winged helix DBD which is structurally similar to that of histone H1.[189] The DBD contains a helix–turn–helix (HTH) motif that makes base-specific DNA contacts as well as two flanking loops (wings) that contact the phosphodiester backbone of DNA. FoxA1 is known to stably bind nucleosomes *in vitro* and *in vivo* near the nucleosome dyad[190,191] and decompact repressed chromatin compacted by H1 to make it accessible for other DNA binding factors.[192,193]

In vitro sequential binding experiments with purified proteins showed that FoxA1 displaces H1 prebound on assembled nucleosomes.[190] Further DNase footprinting of H1-compacted nucleosome arrays with and without FoxA1 demonstrated increased hypersensitivity in the digestion patterns and indicates that FoxA1 can open H1 compacted nucleosomes.[190,192] Truncation mutants of FoxA1 missing the 174 amino acid C-terminal domain failed to open the compacted arrays,[192] underlining the importance of the disordered regions of FoxA1. Early studies also indicated that the N- and C-terminal regions of FoxA1 are crucial for binding specificity to nucleosomes over free enhancer DNA.[190] More recent work identified a short region in the C-terminus of FoxA1, conserved among FoxA pioneer factors, that interacts with core histones and contributes to chromatin opening *in vitro* (see Fig. 7).[194] A single-locus study demonstrated that FoxA1 induction caused reduction of H1 occupancy at an enhancer site during retinoic acid-mediated differentiation of embryonic stem cells.[195] In later studies, an assessment of genome-wide occupancy of linker histone H1 in mouse hepatocytes showed FoxA occupancy on nucleosomes correlates with H1 displacement, whereas the FoxA deletion mutants had a striking increase in H1 disposition. All of these results indicate that FoxA binding

displaces linker histones from the local chromatin, which could explain the subsequent increase in nucleosome accessibility and stimulation of transcription.

3.2.4 Prothymosin α

The nuclear protein prothymosin α (ProTα) is a linker histone chaperone that modulates H1 interaction with nucleosomes. Besides affecting chromatin condensation[196] and H1 mobility in the nucleus,[197] ProTα is involved in transcriptional regulation, cell proliferation, and apoptosis.[198] ProTα is fully disordered, with a highly negatively charged glutamate-rich (net charge −44) amino acid sequence and low hydrophobicity.[199,200] Borgia and co-workers used a combination of single-molecule FRET, NMR spectroscopy, and CG simulations to study the interaction between ProTα and histone H1, and showed that they form a tight complex with picomolar affinity yet remain highly disordered and dynamic in the bound state.[25] This novel interaction mode can be explained by the large opposite net charge of the two proteins which leads to complex formation through a mean-field type charge interaction without the need for defined binding sites or persistent interactions between specific individual residues. The CG simulations, which relied on a simple model involving non-specific short-range and electrostatic interactions, were able to reproduce the experimentally measured FRET efficiencies in the complex remarkably well. Later, Sottini et al. showed through an elegant set of kinetics experiments that ProTα and H1 can also form higher order but weakly interacting ternary complexes.[46] Again, integrating experiments and simulations, they showed that a second ProTα or H1 molecule can engage a preformed ProTα-H1 complex and lead to rapid exchange, keeping the system highly responsive despite the tight binding.

What is the purpose of forming such a disordered complex in the nuclear context? Heidarsson et al. addressed that question by studying the H1-ProTα interaction in the presence of reconstituted nucleosomes[17] (described also above, see Section "Linker histone H1"). Kinetic experiments using immobilized and fluorescently labeled nucleosomes showed that ProTα forms a ternary complex with H1 and the nucleosome, which accelerates the dissociation of H1 by almost two orders of magnitude through a competitive substitution mechanism. Further CG simulations confirmed the dramatic increase in dissociation rate as a function of ProTα binding and provided a molecular picture of how ProTα invades the complex by dynamically and gradually sequestering the H1 C-terminal IDR. The high negative charge

in ProTα thus competes with the electrostatic interactions between the linker DNA and the disordered regions of H1, which reduces the interaction strength of H1 with the nucleosome and leads to an opening of the nucleosome linkers. These results provide clues toward resolving long-standing issues on histone H1 including the nature of the structural ensemble of H1 on the nucleosome and the discrepancy between *in vivo* (minutes) and *in vitro* (hours) residence times of H1 on the nucleosome.[55,201] Through integrative modeling of these challenging molecules, the authors suggested that it is precisely the high degree of dynamic disorder on the H1 IDRs that allows chaperones like ProTα to invade the complex and accelerate the dissociation of H1 from the nucleosomes. For such unspecific, charge dominated binding between dynamic and disordered proteins, the formation of higher order complexes may commonly occur, providing additional functionality and enabling a sensitive concentration-dependent response during signaling. Formation of higher order oligomers and the dynamic exchange within them may be particularly important to achieve dissociation of strongly interacting polyelectrolytes,[25] and to induce formation and regulation of phase-separated condensates.[5,202]

3.3 Chromatin remodelers and histone-modifying enzymes

Chromatin remodelers dynamically modify chromatin architecture to modulate access of the transcriptional machinery to DNA, and thus regulating gene expression.[203] Remodeling pathways are largely dependent on (i) various covalent modifications of histone tails driven by ATP-independent factors[203] such as deacetylase (HDAC), methyl transferase (HMT), acetyl Transferase (HAT), (ii) ATP-dependent chromatin remodeling complexes[204] which either slide, eject or restructure nucleosomes, and (iii) chaperones that bind to histones and stimulate their transfer onto DNA or other proteins.[205] On the basis of their functions, chromatin remodelers can be roughly divided into two families: ATP-dependent enzymes that include imitation switch (ISWI), chromodomain helicase DNA binding (CHD), switch/sucrose non-fermentable (SWI/SNF) and INO80,[204] and ATP-independent enzymes including the histone methyl/acetyl transferases, kinases, and isomerases. Despite differences in mechanisms and compositions, all ATP-dependent remodelers contain a structurally similar catalytic ATPase core which converts the chemical energy of ATP hydrolysis into conformational changes. Besides actively regulating

gene expression, dynamic remodeling of chromatin imparts an epigenetic role in several key biological processes, *e.g.*, DNA replication and repair, apoptosis, and pluripotency.[206]

Chromatin remodelers have an extensive range of interacting partners. They can form multimeric complexes and interact with histones, transcription factors, nucleic acids, and various other machinery involved in the maintenance of chromatin structure.[207] Such a diverse range of interactions is difficult to explain with highly structured proteins. Predictions from amino acid sequence strongly suggest that chromatin remodelers contain substantial structural disorder,[208,209] involved in forming stable complexes and transient interactions with diverse interacting partners, potentially playing a more direct functional role than acting as simple linkers.[210,211]

3.3.1 ATP-dependent chromatin remodelers

Many ATP-dependent chromatin remodeling complexes are predicted to contain IDRs.[212] These IDRs range from relatively small regions, likely functioning as linkers, all the way to the BRG1/BRM-associated factor (BAF) complex which is made up of subunits that are predicted to contain long IDRs.[209,213] A recent study looked at the predicted disorder in BAF and found that 27 of the 30 subunits that were analyzed were predicted to be highly disordered.[209] The BAF complex is among the most frequently mutated complexes in many types of cancer, many of which are located in predicted disordered regions.[214] While the function of the predicted IDRs remains largely unknown, they are likely to assist with binding to histones, nucleic acids, and transcription factors.

ATRX (alpha thalassemia/mental retardation syndrome X-linked) belongs to the SWI/SNF family of chromatin remodeling proteins, and along with Death-associated protein 6 (DAXX), forms a complex that is necessary for H3.3 depositions into pericentric, telomeric, and ribosomal repeat sequences.[215,216] ATRX has multiple functions in the chromatin landscape, acting both as a chromatin remodeler and a histone chaperone.[217] ATRX is a large protein (2492 residues) and contains two structured domains; an N-terminal PHD-like domain and a conserved Snf2 domain.[218] The remaining ~1660 residues of ATRX sequence are predicted to be structurally disordered, with over 1300 residues in a single stretch separating the two domains.[219] The partner protein DAXX, a H3.3 histone chaperone, contains a long disordered C-terminal domain (residues 418–740).[213]

The involvement of the IDRs in ATRX and DAXX for catalyzing the deposition and remodeling of H3.3 nucleosomes, remains unclear.

Chromatin accessibility complex (CHRAC) is an evolutionarily conserved nucleosome remodeling complex that catalyzes histone octamer sliding on DNA.[220] Originally purified from *Drosophila melanogaster*, CHRAC consists of ISWI (ATPase), ACF1 and two histone fold subunits, CHRAC-14 and CHRAC-16.[221] A study looking into the function of CHRAC-14 and CHRAC-16 found unstructured N- and C-terminal domains on both proteins.[222] CHRAC-14 and CHRAC-16 form a heterodimer with a fold that resembles the geometry of histone dimer H2A-H2B,[222] which is predicted to create a surface for transient deposition of a segment of DNA as it is stripped from the core histone octamer. The C-terminal of both proteins is involved in DNA binding but with reciprocal effects; the C-terminal on CHRAC-14 increases DNA binding while the C-terminal on CHRAC-16 greatly decreases it but is still essential for sliding on DNA. It seems that the CHRAC-14/CHRAC-16 heterodimer enhances the catalysis of nucleosome sliding with weak and non-specific DNA binding. These findings were strikingly similar to the groups earlier work on the DNA chaperone HMGB1,[223] leading the authors to speculate that CHRAC-14/CHRAC-16 heterodimer serves as a built-In DNA chaperone.

3.3.2 ATP-independent chromatin remodelers

Post-translational modifications frequently occur in IDRs, as outlined above. Acetylation of the core histones enhances transcription by relaxing the condensed structure of the nucleosome, whereas deacetylation will promote chromatin condensation and transcriptional repression.[224,225] This effect is due to a charge neutralization of the acetylated lysine that weakens its interaction with the phosphate backbone of DNA. Both histone deacetylases and histone methylases are regulated by phosphorylations in predicted IDRs. Phosphorylations in HDACs 4,5,7 and 9 regulate shuttling between the nuclear and cytoplasmic compartments[226] and phosphorylations of sites flanking the nuclear localization sequence will promote chaperone protein binding and subsequent nuclear export.[227,228]

Histone methylation is a dynamic PTM central to eukaryotic transcription.[229] These modifications regulate gene expression by recruiting transcriptional cofactors that specifically recognize methylated lysine or arginine residues.[230,231] Dysregulation of histone methylation is associated with serious diseases such as cancers, developmental defects, and inflammatory bowel disease.[232,233] A recent study looked into PTMs of histone methylation

enzymes in *Saccharomyces cerevisiae*, and found that phosphorylation was strongly enriched in predicted IDRs in methyltransferases while histone demethylases were phosphorylated within ordered regions.[234] Furthermore, the authors demonstrated that a phosphorylation cluster within an IDR of methyltransferase Set2p has a major effect on levels of H3K36 methylation *in vivo*. This decrease in H3K36 methylation leads to increased cryptic transcription, which can shorten the lifespan of cells.[235]

SIRT6 is an NAD^+-dependent histone deacetylase and is highly site-specific.[236] While early experiments, using H3 peptides,[237] demonstrated that SIRT6 has an ~1000 times slower catalytic activity then other related sirtuins, the low turnover rate did not match with recent studies using whole nucleosomes as substrates that found significantly higher catalytic rates.[238] This is likely due to interactions between the intrinsically disordered C-terminal region that has a high affinity to the nucleosome[239]; with SIRT6 tethered to the nucleosome the reaction can take place with greatly enhanced activity. Interestingly, while the SIRT6 interacts with nucleosomes in a 2:1 arrangement, only a single SIRT6 molecule can occupy the high affinity site. This arrangement may be due to the asymmetry of the two acidic patches, as observed with other chromatin remodelers that have a distinct response to each acidic patch.[240]

3.3.3 Chromatin remodelers with chaperone activity

Facilitates chromatin transcription (FACT) is a histone chaperone that has a dual-role as a nucleosome remodeler and chaperone.[225,241] In gene regulation, nucleosomes must temporarily unfold and then rapidly refold after the regulatory process. FACT increases accessibility of RNA polymerase II on chromatin by unfolding the nucleosome structure (Fig. 5).[242] FACT can then act as a histone chaperone that promotes nucleosome assembly by preventing some non-productive interactions between histones and DNA.[243] Both of FACT's two subunits, SSRP1 and SPT16, contain acidic and disordered regions that are implicated in histone binding.[241,244] Unlike most other histone chaperones, FACT can bind both H2A-H2B and H3-H4 dimers simultaneously,[245] with both subunits being involved in several interactions. Cryo-EM structures of FACT or SPT16 in complex with nucleosome constructs revealed that the CTD of SPT16, that includes an acidic IDR important for H2A/H2B binding, adopts a more ordered conformation when in complex with parts of the nucleosome.[246,247] Interestingly, the CTD appears to mimic DNA by compensating for the loss of histone DNA contacts (Fig. 5).[247] In a follow-up study using NMR

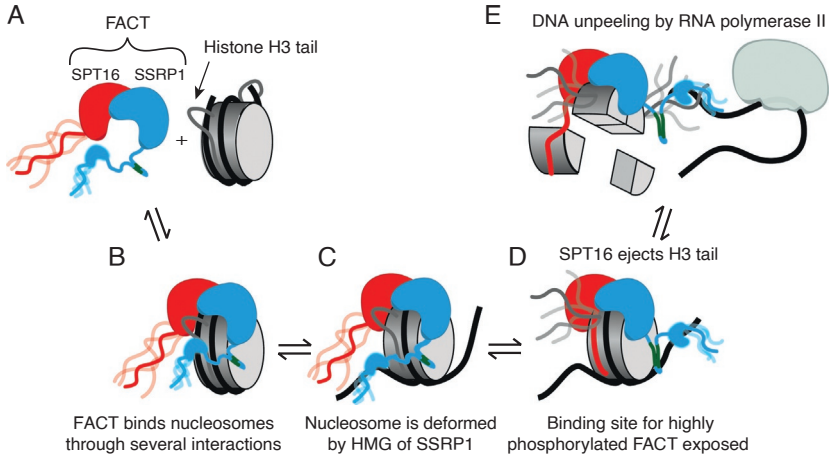

Fig. 5 Nucleosome assembly/disassembly by the histone chaperone FACT. (A) A canonical nucleosome with H3 tails (for clarity only the disordered tails of histone H3 is shown) buried in DNA gyres and the two subunits of FACT, SSRP1 and SPT16. (B) FACT binding to the nucleosome leads to deformation by the action of the HMG domain of the less phosphorylated (green area) state of SSRP1. (C) FACT with a highly phosphorylated SSRP1 has high affinity for deformed nucleosomes and replaces less phosphorylated FACT. (D) Deformation of the nucleosome exposes a binding site for the C-terminal domain of SPT16, causing increased solvent exposure of the histone H3 tail. (E) DNA is peeled off the nucleosome by RNA polymerase II (or other factors). Following transcription by RNA polymerase II, FACT can reassemble the nucleosome (not shown).

spectroscopy, it was revealed that one of the N-terminal tails of H3 adopts a different conformational ensemble when FACT is bound to the nucleosome.[248] NMR analysis of H3 tail chemical shifts indicated that it is buried in between two DNA gyres and that interaction is disrupted by the CTD of SPT16. This leads to increased solvent exposure of the tail, rendering it more susceptible to acetylation by HAT, indicating that FACT has a regulatory role in H3 acetylation. The intrinsically disordered domain (IDD) of SSRP1 has an acidic N-terminal part (AID) and a basic C-terminal part (BID). A recent study using NMR and CG molecular dynamics simulations, revealed how phosphorylation in the IDD change the intermolecular contacts between the AID and BID. These contact changes tune the affinity of SSRP1, with less phosphorylated states displaying high affinity to an intact nucleosome and highly phosphorylated states having high affinity to a deformed nucleosome, revealing an important mechanistic and regulatory role for the IDD.[210]

Another remodeler, decondensation factor 31 (Df31), is a fully disordered histone chaperone and an integral component of chromatin at all stages of *Drosophila melanogaster* lifecycle.[249,250] Df31 is suggested to have a role in the higher order structure of chromatin by promoting chromatin bridging *in vitro*.[251] Df31 binds to both histone H3 and H4 but has a higher affinity for H3.[252] Binding to H3 takes place through the intrinsically disordered H3 tail,[251] making PTMs to the H3 tail a likely modulator for binding. Recently, an RNA-dependent mechanism was discovered, where Df31 tethers chromatin-associated RNA (caRNA) to chromatin, resulting in an RNA-chromatin network which is more accessible and active.[252]

We have highlighted here how structural disorder is a prominent part of chromatin remodeling complexes but for most remodelers discussed here, detailed mechanistic insights remain hidden. FACT has, however, a well-established molecular mechanism, which was revealed with a close integration of NMR experiments and coarse-grained simulations, exemplifying the strength of such approaches.

3.4 Transcription through a nucleosomal barrier with disordered proteins

3.4.1 Transcription factors

The nucleosome represents a formidable barrier to transcription as the DNA sequence encoding a specific gene must become accessible to transcription factors in one way or another. The transcriptional machinery is rich with disorder and even the ribosomal assembly contains many disordered protein subunits.[253] The vast majority of transcription factors (TFs) (>85%) have long disordered linkers and transactivation domains (TADs) that flank their structured DBDs.[23,254] They bind cognate DNA sequences using predominantly their structured DBDs and may subsequently recruit other proteins to their binding site through their disordered TADs to initiate transcription. The TADs often contain hydrophobic residues (frequently aromatics) well interspersed with acidic residues, a feature that has been suggested to be important for keeping the region disordered and exposed in an active form allowing interactions with other proteins.[255] Nonetheless, the IDRs are not exclusively involved in protein-protein interactions: simulations have suggested that the affinity to DNA, cognate or non-specific, is tuned by disordered regions, especially those that have significant charges.[256] IDRs in TFs have also been linked to facilitating scanning for correct binding sites through non-specific interactions,[257] and to inter-strand exchange through a monkey-bar-like mechanism.[258] In fact, recent evidence points to TFs

having multiple specificity determinants encoded in their IDR sequence, helping them to identify their specific binding sites by interacting with much broader DNA regions than are recognized with only their DBD cognate sites.[259] However, in the context of our nucleosomal landscape, traditional TFs require their binding sites to be accessible for binding, *i.e.*, within "open" chromatin states.

3.4.2 Pioneer transcription factors can alter cell fate

A unique class of TFs, called pioneer-TFs (pTFs), can bind to condensed, nucleosome-rich regions of the genome and open these previously inaccessible regions to transcription (Fig. 6).[187,260] This alters the transcriptional pattern of a cell—the main determinant of its fate[260]—and can initiate cell reprogramming. Despite the ultimate change in cell fate relying on subsequent recruitment of other factors, the initial binding ability to condensed chromatin is what distinguishes pTFs from other TFs. A remarkable example of pioneer activity is the so-called Yamanaka factors; a group of four pTFs (Oct4, Sox2, Klf4, and c-Myc) that can induce a fibroblast to revert to a pluripotent stem cell (iPSC)[261]—a process that earned the discoverers the Nobel prize in 2012. Other pTFs, such as FoxA1, Ascl1, and Pu.1, have since been shown to play key roles for inducing direct reprogramming from fibroblasts to hepatocytes, neurons, and macrophage-like cells, respectively.[187] Reprogramming cell fate has immense potential for human health, with recent reports showing extraordinary examples in regenerative medicine such as sight restoration in mice, *in vitro* disease modeling, and drug discovery.[262,263] However, to fully exploit the power of pTFs for cell reprogramming, a detailed and quantitative understanding of their molecular mechanism is critically needed.[187] For example, it is largely unknown whether pTFs bind to DNA that becomes spontaneously and transiently accessible on nucleosomes or whether they actively "open" nucleosomal DNA. In other words, how pTFs can dynamically invade compacted chromatin and initiate remodeling remains unclear. Some pTFs interact with enzymes that remodel chromatin besides recruiting other TFs, and in those cases, it can be challenging to separate the actions of the two classes of proteins: are the chromatin remodelers necessary for remodeling and do the pTFs just invade chromatin to initiate binding, or can those pTFs also remodel chromatin themselves? The answers to these questions remain hidden, in part due to the highly dynamic and heterogeneous conformations of pTFs and chromatin, which render these systems notoriously difficult to assay by classical structural biology methods.

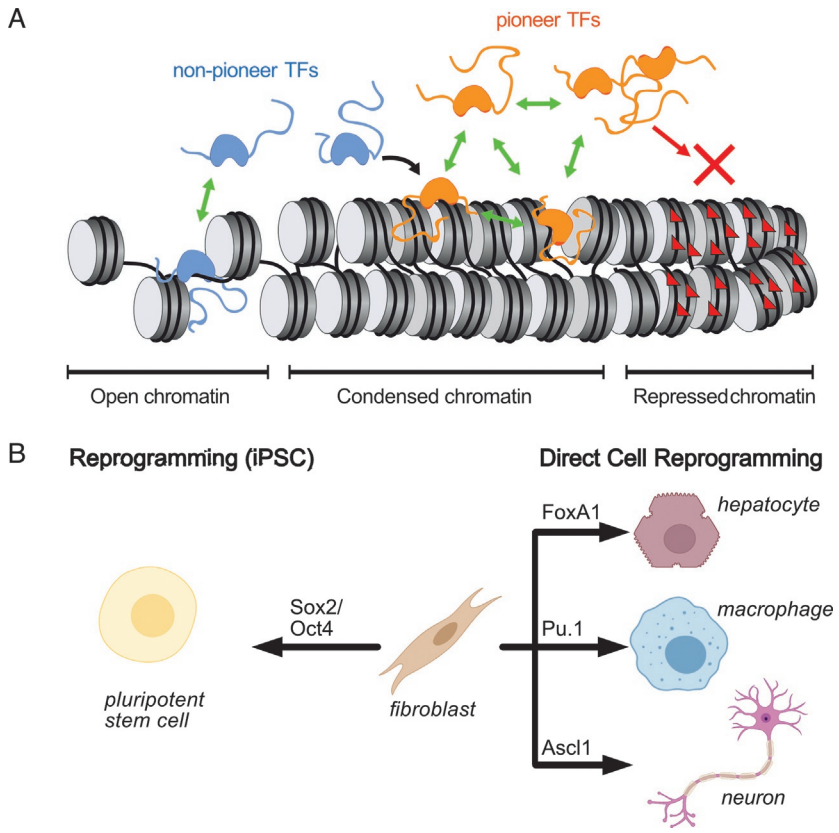

Fig. 6 Pioneer transcription factors can invade and open condensed chromatin and initiate cell-fate changes. (A) Pioneer-TFs (orange) can bind to condensed chromatin regions and render it accessible to traditional TFs (blue) or other components of the transcriptional machinery. (B) Pioneer-TFs can lead to cell-fate changes, either through reprogramming with formation of induced pluripotent stem cells (iPSCs), or through direct cell reprogramming. *Panel A based on Zaret KS, Mango SE. Pioneer transcription factors, chromatin dynamics, and cell fate control.* Curr Opin Genet Dev. *2016;37:76–81.*

Like the vast majority of TFs, pTFs are rich in disordered linkers and TADs (Fig. 7).[264] Despite their abundance in pTFs, IDRs have largely been overlooked thus far in studies of TFs, which is especially evident considering the vast number of TF DBDs in the Protein Data Bank and the total absence of 3D-structures containing entire eukaryotic TFs. Instead, intense focus has centered on the DBDs in attempts to explain pioneering activity, with impressive high-resolution structures revealing complexes between the pTF DBDs and nucleosomes.[265,266] The DBDs themselves are often disordered

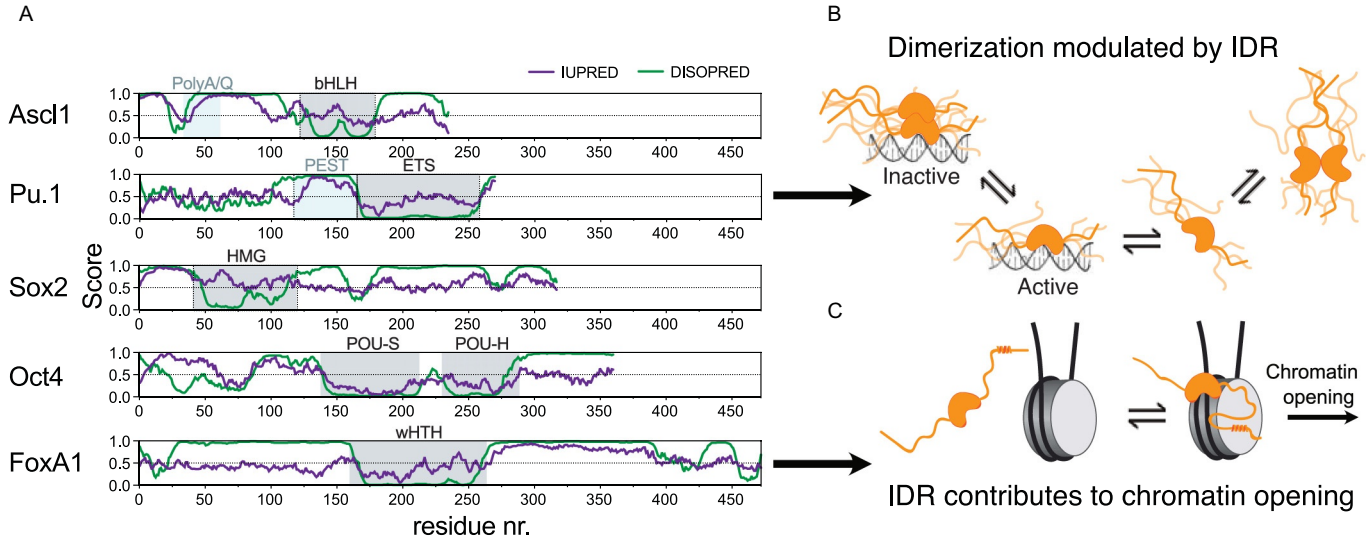

Fig. 7 Intrinsically disordered regions in pioneer transcription factors. (A) Disorder predictions for five pTFs based on two different predictors. Domains are highlighted. (B) The intrinsically disordered and acidic PEST domain in the pTF Pu.1 modulates the formation of a dimer on and off DNA. A 1:1 Pu.1-DNA complex activates transcription and dimerization negatively regulates the activity. The dimer in the absence of DNA is furthermore thermodynamically destabilized compared to the monomer. (C) FoxA1 interacts with the core histones in a nucleosome through a short motif in its C-terminal IDR. This interaction contributes to chromatin opening and thus its pioneering functions. The schematics of the conformations of structured domains and IDRs in panels (B) and (C) are purely for illustrative purposes. *Panel B reaction scheme is based on Xhani S, Lee S, Kim HM, et al. Intrinsic disorder controls two functionally distinct dimers of the master transcription factor PU.1.* Sci Adv. *2020;6(8): eaay3178, Panel C based on results from Iwafuchi M, Cuesta I, Donahue G, et al. Gene network transitions in embryos depend upon interactions between a pioneer transcription factor and core histones.* Nat Genet. *2020;52(4):418–427.*

before binding to their cognate DNA sequence, followed by a disorder-to-order transition upon complex formation.[267] The DBDs are also often the major contributors to DNA affinity and in some cases such as for Sox2, the IDRs seemingly weaken affinity for the cognate sequence.[268] A recent computational study also implicated rotational and sliding dynamics of the DNA on the nucleosome to be important for binding of pTFs. Using CG models and simulations, Tan and Takada showed that Sox2 recognizes a certain rotational phase of its binding site which induced sliding, affecting allosterically the binding of Oct4 or another Sox2 molecule.[269] Clearly, the DBDs are critical for pTF function but what possible role do IDRs play in chromatin opening and subsequent reprogramming pathways?

3.4.3 Roles of intrinsic disorder in pioneer transcription factors

Disordered regions are frequently involved in protein-protein interactions[270] and in TFs the TADs often recruit components necessary for transcription. Moreover, interplay between ordered and disordered regions is poorly understood but expected as IDRs have usually co-evolved with ordered regions and the conformational propensities of IDRs may therefore be modulated by folded domains and *vice versa*. It is possible that, after scanning and binding recognition sites, the DBDs act as anchors to allow the disordered regions to inflict interactions that disrupt internucleosome contacts in chromatin, leading to opening of the chromatin fiber. In that way, IDRs could be involved in actively opening chromatin, modulating oligomerization regulating pioneer activity (see below for Pu.1), or involved in recruiting chromatin remodeling enzymes, followed by opening of chromatin, and subsequent binding of other transcription factors to the exposed DNA. Ultimately, IDRs may have multiple, context-dependent roles regulated by cell-type, chromatin modifications, and local sequence determinants. Nevertheless, the role of IDRs has been glimpsed recently for many pTFs, suggesting a function in chromatin opening and a large-scale impact on gene expression networks.

Strong evidence of IDR involvement in chromatin opening comes from recent work from the Zaret lab, which revealed a role of IDRs in the prototypical pTFs FoxA1 and FoxA2 for modulating interactions with core histone proteins in a nucleosome.[194] Using a combination of sequence analysis, cross-linking, and mass spectrometry, the authors discovered a conserved 9 amino-acid sequence in the disordered C-terminal, which is critical for chromatin opening functions through an interaction with the core histones

in a nucleosome. This short region likely forms a transient α-helix, as helix formation could be induced by addition of helix-promoting trifluoroethanol in a short peptide when monitored by circular dichroism spectroscopy. When this region was deleted, the chromatin opening ability, measured by DNase cleavage sensitivity, was severely reduced, as well as the ability to activate certain target genes. Using mouse embryos, the authors further went to show that deletion of the short a-helix led to a 60% reduction in target gene activation, severely impairing embryonic development by affecting gene expression and chromatin accessibility. Clearly, this disordered region plays a crucial role in the pioneering function of FoxA1. Beyond pioneering functions, the FoxA proteins also have heavy ties to cancer biology through their direct interaction with both the estrogen and androgen receptors,[271] and FoxA1 is currently hailed as a very promising therapeutic target. The interaction of FoxA1 with both receptors is influenced by PTMs in the disordered regions, including SUMOylation that has a negative effect on transcriptional activity and on association with the androgen receptor.[272]

The key Yamanaka factor Sox2 has a short N-terminal and a long, ~200-residue C-terminal IDR flanking an HMG-box DBD.[273] The Sox2 HMG-box cooperates with the Oct4 POU-domain, and this interaction is critical for producing iPSC and maintaining pluripotency but the efficiency of reprogramming is conferred by the extreme C-terminal IDR[274] through a currently unclear mechanism. Recent studies have shown how Sox2 and Oct4 act in concerted fashion to invoke structural changes in the core nucleosome structure ranging from subtle local distortion to fully removing DNA from one side, depending on the cognate binding site location.[265] However, the dynamic events of scanning and binding that finally lead to chromatin opening are still mostly unknown. The IDR region immediately flanking the C-terminal side of the DBD (120–160) has recently been implicated in RNA binding, even concurrently with the DBD being DNA-bound.[273] The authors went on to show that deletion of the RNA binding domain severely reduced the efficiency of iPSC generation, demonstrating a clear link between the IDR and cell reprogramming.

Pu.1 is a hematopoietic master regulator pTF that contains an N-terminal TAD, a disordered anionic PEST domain (rich in prolines, glutamic acids, serines and threonines), and a structured DBD called ETS (Erythroblast transformation specific) domain. Xhani et al. showed that

Pu.1 dimerizes through its DBD and gene expression is regulated by two distinct dimeric states: a transcriptionally active 1:1 complex and an inactive ternary complex involving two Pu.1 molecules bound to a single DNA recognition site (Fig. 7),[275] forming a negative feedback mechanism that the authors confirmed *in vivo*. Using NMR spectroscopy and tryptophan fluorescence experiments, the authors showed that the intrinsically disordered PEST domain reduced the binding affinity of the second Pu.1 molecule to form a ternary complex. Interestingly, however, the PEST domain also promotes homodimerization in the absence of DNA. The two dimeric forms were found to be non-equivalent, with an asymmetric DNA-bound Pu.1 dimer and a symmetric homodimer in the DNA-free state. A legion of serines in the PEST domain is phosphorylated *in vivo*, which prompted the authors to introduce phosphomimetic substitutions in that region. Indeed, the degree of negative feedback was reduced with phosphomimetic substitutions which promoted the formation of a transcriptionally active 1:1 complex with DNA. It remains to be determined whether a similar regulatory dimerization mechanism would be observed on nucleosomes but the positively charged histone tails may provide an additional interaction interface for the negative charges in the PEST domain. There may furthermore be other complicating factors, as binding of Pu.1 to nucleosomes has been reported to be context-specific, suggesting a non-classical pioneering role for Pu.1.[276]

Yet another example of a disordered pTF is the achaete-scute homolog 1 (Ascl1), which drives the conversion of fibroblasts to neurons.[277] Ascl1 is a relatively small transcription factor that has a characteristic polyA/polyQ region in the N-terminal and a basic helix-loop-helix DBD in the C-terminal. In a clever, fragment-based approach, Baronti et al. were able to use NMR spectroscopy to dissect the highly aggregation-prone Ascl1[278] and found an extended and dynamic structure with transient helix formation yet no persistent tertiary interactions—a classical characteristic of an IDP. Little mechanistic information is available on the interactions between Ascl1 and DNA or nucleosomes but a genome-wide analysis showed that it is one of only a handful of TFs that binds strongly to both DNA and nucleosomes albeit likely as a heterodimer.[279]

We have highlighted a subset of pTFs that have been studied by biophysical approaches but many other established pTFs are predicted to contain long IDRs.[264] Molecular biology has over the years been extraordinarily powerful at identifying the key players in transcriptional regulation networks

during cell development. Yet, the link between molecular properties of pTFs, especially the role of their IDRs, and cell reprogramming is still largely missing. Integrative modeling approaches, using available structural information in concert with biophysical studies and simulations, might be a potent strategy to understand the physical principles of cell-identity pathways, leading us closer to controlling cell fate.

4. Common sequence features of disordered nucleosome-binding proteins

In the disordered interactions and their regulation reviewed above, charge emerges as a recurring theme. Charge is a principal component of chromatin and is often utilized by IDPs to elicit a specific cellular response. While the DNA backbone is highly acidic, the linker and core histone tails are highly basic, creating an electrostatic balance in the NCP.[280] Opposite charges in the DNA and histone tails have been implicated in a number of inter- and intra-nucleosomal interactions, which act to either condense or decondense chromatin. Moreover, PTMs that alter charge in the disordered histone tails have been shown to affect nucleosome stability.[281] For instance, neutralization of positive charge by acetylation or introduction of negative charge by phosphorylation of basic residues in the histone H3/H4 tail regions, weakens the histone-DNA interactions by reducing electrostatic attraction.[281] Consequently, chromatin takes on an open structure, increasing nucleosome accessibility to modifying enzymes. Charge has an especially clear role for the highly disordered H1 competitors (protamines, HMG proteins, ProTα). A common feature among these proteins may be that the unspecific nature of charge interactions and the high fraction of charges allows these proteins to interact in complexes beyond a basic 1:1 stoichiometry, exchange rapidly in a concentration-dependent manner, and keep regulatory systems highly responsive despite high affinity binding. Those molecular parameters would in turn be finely regulated by PTMs that affect charge.

In the cell, several transcription factors, chromatin remodelers and architectural proteins function in a dynamic balance, ultimately controlling gene expression. Understanding the effects of charge in IDPs that interact with chromatin and chromatin-binding proteins may provide insight into their specific cellular mechanisms. To better understand charge properties, we calculated kappa (κ) values for the IDRs of proteins discussed in this review (Fig. 8). κ is a patterning parameter used to describe strong and weak

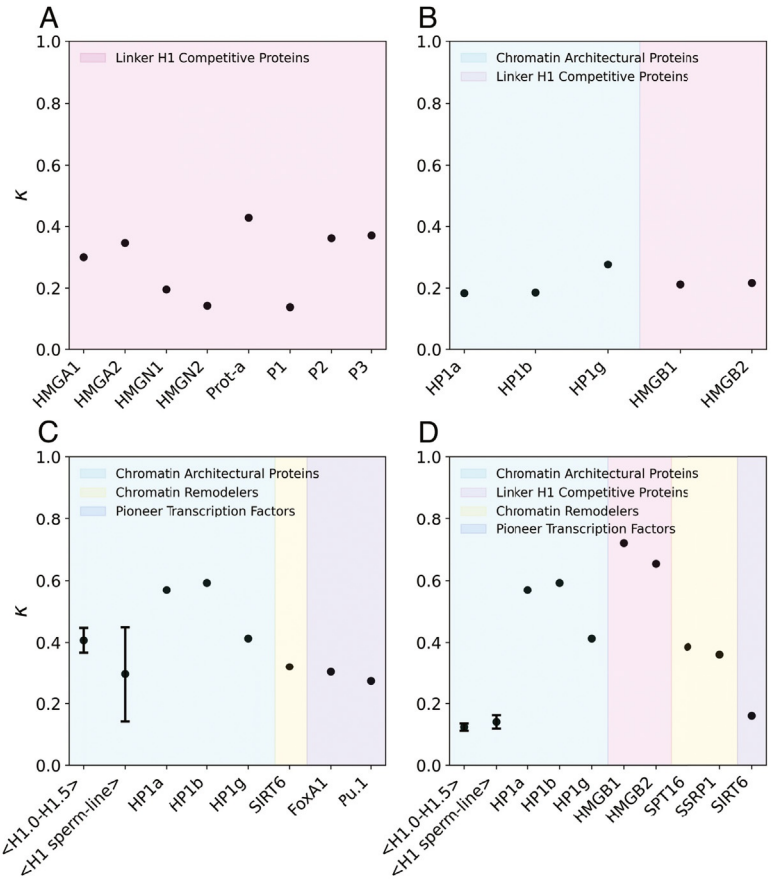

Fig. 8 Overview of the charge distribution within the intrinsically disordered regions of nucleosome-binding proteins. Understanding the role of charge in nucleosome-binding proteins may prove important to frame the functional space of different intrinsically disordered proteins within the context of transcriptional regulation. The parameter kappa (κ), which can take any value between 0 and 1 and has been formulated to link sequence properties to conformational behavior of intrinsically disordered regions,[26] describes the overall charge asymmetry in an amino-acid sequence. A low κ indicates more evenly distributed positive and negative charge, whereas κ increasingly close to 1 indicates blocks of positive and negative charge. The need to strongly coordinate DNA may render charge distribution an important factor to finely tune protein-DNA interactions. While for fully disordered proteins (shown in A) κ is found to vary considerably, for intrinsically disordered hinge regions linking structured domains (shown in B) κ is low and similar across all proteins. A clear difference can be noted between the κ values of the N-terminal (C) domains of linker histones that are involved in the physiological homeostasis of chromatin and those expressed by sperm-line cells, which are involved in extreme chromatin compaction and that show a lower κ. While no other trend can be clearly seen for the other proteins or for their C-terminal domains (shown in D), it is important to acknowledge that our classification is arbitrary and founded on the current understanding of the role that these proteins have within the nucleosomal landscape.

polyampholytes. A low κ value is indicative of well distributed negative and positive charges along an intrinsically disordered domain which generate extended ensembles, where intramolecular electrostatic attractions and repulsions are counterbalanced.[26] On the contrary, a κ value close to 1.0 indicates blocks of opposite charges that strongly interact leading to globule-like conformations with low radius of gyration.

HMG proteins compete with histone H1 to bind chromatin and thus HMG-nucleosome interactions often result in chromatin decondensation. For instance, HMGA1 competes with linker H1 on SARs of nucleosomal DNA, displacing H1 to non-SAR DNA and inhibiting chromatin compaction.[184] Interestingly, the κ value for HMGA1 is similar to that of the CTD of linker H1.1, H1.2 and H1.5 ($0.3 < \kappa < 0.4$). Linker histone H1.5 has a long CTD tail, containing more than two S/T-P-X-K sites, resulting in a high affinity for heterochromatin. In contrast, H1.1 and H1.2 have shorter CTD tails, with fewer S/T-P-X-K sites, and are enriched at euchromatic regions.[78] Therefore, charge distribution, in addition to net positive charge and disorder, may also impact affinity.

Like HMGA1/A2, HMGB1/B2 also contain an acidic tail and displace linker H1 from the nucleosomal dyad.[173] However, unlike HMGA1/A2, HMGB1/B2 are not completely disordered, although the disordered CTD is required for correct HMGB function.[175] HMGB1 is involved in the regulation of p53; a tumor suppressor that binds to DNA which acts by protect cells from malignant transformation.[282,283] HMGB1 has been shown to stimulate the linear DNA-p53 interaction *in vitro* and, *in vivo*, p53 activity is increased.[284] Additionally, HMGB1 and p53 have been shown to directly interact *via* the PXXPXP motif in the disordered NTD of p53 and HMG boxes in HMGB1. Moreover, the disordered acidic tail in HMGB1 is a direct determinant of this interaction, as it shields the positive charge in the HMG box decreasing p53-HMGB1 affinity and linking disorder to protein function.[285] In contrast to HMGB1, the interaction between H1.2 and p53 induces p53 repression in DNA damage response. Moreover, this interaction is negatively regulated by acetylation in the p53 CTD and phosphorylation in the H1.2 CTD. In both cases, PTM acts to disrupt the p53-H1.2 interactions, directly implicating charge and disorder in protein functionality.[286]

The connection between charge and disorder is prominent when considering the interplay of HMGN1/N2 and H1T2, H1oo and HILS1 variants. HMGN1/N2 promote chromatin decompaction by interacting with

nucleosomal DNA at the major grooves flanking the dyad and competing with linker H1 for binding sites.[155] Furthermore, HMGN1/N2 has a low kappa value ($0.1 < \kappa < 0.2$), that is similar to that of the H1T2, HILS1 and H1oo CTD and characteristic of disordered proteins. Therefore, HMGN1/N2 may use its disorder to compete with these H1 variants for binding sites. Interestingly, while most H1 variants have few arginine residues, H1T2 and HILS1 have an almost equal fraction of lysine and arginine residues in the CTD. Because arginine forms stronger interactions with the DNA phosphate backbone, the testis specific variants are likely to be harder to displace. For instance, during spermatogenesis, inactivation of the gene for H1T2 leads to defects in DNA condensation and chromatin packing; effects that are not favorable in cell development.[287]

5. Concluding remarks

The nucleus is enriched in proteins that are disordered and thus highly dynamic. These proteins play key roles in maintaining the genome and regulating its read-out. Despite decades of active research on IDPs and their well-recognized importance in ensuring the homeostasis of the nucleus, we still lack an exhaustive description of the interactions between chromatin components and IDPs, especially with respect to how they translate to biological function and regulation. New methodological paradigms are needed to tackle intrinsic disorder in the nucleus, because of both the intrinsic dynamic character and the physico-chemical properties of the interacting molecular partners, which frequently feature extremely strong electrostatics. Consequently, in recent years there has been a considerable upsurge in methodological development, especially for single-molecule techniques which can discriminate distinct conformational sub-populations and sequences of molecular events. Computational approaches that directly integrate single-molecule data and simulations, featuring customized potential energy functions tuned on the basis of experimental findings, have provided an unprecedented view of the ensemble of some key disordered interactions in the nucleus. Remarkably, simple potential energy functions that dominantly account for electrostatic contributions to binding, have been able to exhaustively reproduce experimental findings and provide a mechanistic understanding of protein-protein and protein-DNA interactions.[17,25,288] In the future, such simple customized potentials may evolve into more complex

combinations of potential energy terms that might take into account, explicitly, the effects of post-translational modifications, such as methylation and acetylation, in specific sites along intrinsically disordered domains. Additionally, the modeling and parameterization of explicit ionic species, especially for coarse-grained simulations, would be a considerable advancement for a more accurate estimation of the energetics involved in the nucleosomal landscape, especially considering the primary role of ions in defining the association of strong disordered polyampholyte chains that interact with chromatin. Overall, access to integrative modeling approaches is still a challenge, as it requires strong collaborative efforts between different research groups. Nevertheless, creating synergy between experiments and simulations is key to refining our view of the disordered nuclear milieu.

Acknowledgments

This work was funded by the Icelandic Research Fund (Rannís grant nr. 206591-052 and 217392-051, to P.O.H.) and the Centre for eResearch (http://www.eresearch.auckland.ac.nz) at the University of Auckland for their help in facilitating this research. We apologize to the authors of other excellent work that we could not cite due to space limitations.

References

1. Wright PE, Dyson HJ. Intrinsically unstructured proteins: re-assessing the protein structure-function paradigm. *J Mol Biol*. 1999;293(2):321–331.
2. Schad E, Tompa P, Hegyi H. The relationship between proteome size, structural disorder and organism complexity. *Genome Biol*. 2011;12(12):R120.
3. van der Lee R, Buljan M, Lang B, et al. Classification of intrinsically disordered regions and proteins. *Chem Rev*. 2014;114(13):6589–6631.
4. Frege T, Uversky VN. Intrinsically disordered proteins in the nucleus of human cells. *Biochem Biophys Rep*. 2015;1:33–51.
5. Schuler B, Borgia A, Borgia MB, et al. Binding without folding—the biomolecular function of disordered polyelectrolyte complexes. *Curr Opin Struct Biol*. 2020;60:66–76.
6. Wright PE, Dyson HJ. Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol*. 2015;16(1):18–29.
7. Theillet FX, Binolfi A, Frembgen-Kesner T, et al. Physicochemical properties of cells and their effects on intrinsically disordered proteins (IDPs). *Chem Rev*. 2014;114(13):6661–6714.
8. Habchi J, Tompa P, Longhi S, Uversky VN. Introducing protein intrinsic disorder. *Chem Rev*. 2014;114(13):6561–6588.
9. Hsu WL, Oldfield CJ, Xue B, et al. Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. *Protein Sci*. 2013;22(3):258–273.
10. Kulkarni P, Uversky VN. Intrinsically disordered proteins: the dark horse of the dark proteome. *Proteomics*. 2018;18(21 – 22):e1800061.

11. Fyodorov DV, B-r Z, Skoultchi AI, Bai Y. Emerging roles of linker histones in regulating chromatin structure and function. *Nat Rev Mol Cell Biol*. 2017;19:192–206. Nature Publishing Group.
12. McGinty R, Tan S. Nucleosome structure and function. *Chem Rev*. 2014;115(6): 2255–2273.
13. Bednar J, Garcia-Saez I, Boopathi R, et al. Structure and dynamics of a 197 bp nucleosome in complex with linker histone H1. *Mol Cell*. 2017;66(3):384–397.e8.
14. Peng Z, Mizianty MJ, Xue B, Kurgan L, Uversky VN. More than just tails: Intrinsic disorder in histone proteins. *Mol Biosyst*. 2012;8(7):1886–1901.
15. Martire S, Banaszynski LA. The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biol*. 2020;21(9):522–541.
16. Lever MA, Th'ng JP, Sun X, Hendzel MJ. Rapid exchange of histone H1.1 on chromatin in living human cells. *Nature*. 2000;408(6814):873–876.
17. Heidarsson PO, Mercadante D, Sottini A, et al. Disordered proteins enable histone chaperoning on the nucleosome. *bioRxiv*. 2020;2020.04.17.046243.
18. Kale S, Goncearenco A, Markov Y, Landsman D, Panchenko AR. Molecular recognition of nucleosomes by binding partners. *Curr Opin Struct Biol*. 2019;56: 164–170.
19. Walsh CT, Garneau-Tsodikova S, Gatto Jr GJ. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angew Chem Int Ed Engl*. 2005;44(45): 7342–7372.
20. Bowman GD, Poirier MG. Post-translational modifications of histones that influence nucleosome dynamics. *Chem Rev*. 2015;115(6):2274–2295.
21. Darling AL, Uversky VN. Intrinsic disorder and posttranslational modifications: the darker side of the biological dark matter. *Front Genet*. 2018;9:158.
22. Luo C, Hajkova P, Ecker JR. Dynamic DNA methylation: in the right place at the right time. *Science*. 2018;361(6409):1336–1340.
23. Liu J, Perumal NB, Oldfield CJ, Su EW, Uversky VN, Dunker AK. Intrinsic disorder in transcription factors. *Biochemistry*. 2006;45(22):6873–6888.
24. Tompa P, Fuxreiter M. Fuzzy complexes: polymorphism and structural disorder in protein-protein interactions. *Trends Biochem Sci*. 2008;33:2–8.
25. Borgia A, Borgia MB, Bugge K, et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature*. 2018;555(7694):61–66.
26. Das RK, Pappu RV. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc Natl Acad Sci U S A*. 2013;110(33):13392–13397.
27. Cuvier O, Fierz B. Dynamic chromatin technologies: from individual molecules to epigenomic regulation in cells. *Nat Rev Genet*. 2017;18(8):457–472.
28. Hellman LM, Fried MG. Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat Protoc*. 2007;2(8):1849–1861.
29. Jain SS, Tullius TD. Footprinting protein-DNA complexes using the hydroxyl radical. *Nat Protoc*. 2008;3(6):1092–1100.
30. Gontier A, Varela PF, Nemoz C, et al. Measurements of protein-DNA complexes interactions by isothermal titration calorimetry (ITC) and microscale thermophoresis (MST). *Methods Mol Biol*. 2021;2247:125–143.
31. Nakato R, Sakata T. Methods for ChIP-seq analysis: a practical workflow and advanced applications. *Methods*. 2021;187:44–53.
32. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods*. 2013;10(12): 1213–1218.

33. Nakane T, Kotecha A, Sente A, et al. Single-particle cryo-EM at atomic resolution. *Nature*. 2020;587(7832):152–156.
34. Shi Y. A glimpse of structural biology through X-ray crystallography. *Cell*. 2014;159(5):995–1014.
35. Wu M, Lander GC. Present and emerging methodologies in Cryo-EM single-particle analysis. *Biophys J*. 2020;119(7):1281–1289.
36. Loveland AB, Demo G, Korostelev AA. Cryo-EM of elongating ribosome with EF-Tu*GTP elucidates tRNA proofreading. *Nature*. 2020;584(7822):640–645.
37. 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(6648):251–260.
38. Adhireksan Z, Sharma D, Lee PL, Davey CA. Near-atomic resolution structures of interdigitated nucleosome fibres. *Nat Commun*. 2020;11(1):4747.
39. Song F, Chen P, Sun D, et al. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science*. 2014;344:376–380.
40. Garcia-Saez I, Menoni H, Boopathi R, et al. Structure of an H1-bound 6-nucleosome array reveals an untwisted two-start chromatin fiber conformation. *Mol Cell*. 2018; 72(5):902–915.e7.
41. Bharath MS, Chandra NR, Rao M. Molecular modeling of the chromatosome particle. *Nucleic Acids Res*. 2003;31(14):4264–4274.
42. Wong H, Victor J-M, Mozziconacci J. An all-atom model of the chromatin fiber containing linker histones reveals a versatile structure tuned by the nucleosomal repeat length. *PLoS One*. 2007;2(9):e877.
43. Stehr R, Kepper N, Rippe K, Wedemann G. The effect of internucleosomal interaction on folding of the chromatin fiber. *Biophys J*. 2008;95(8):3677–3691.
44. Kepper N, Foethke D, Stehr R, Wedemann G, Rippe K. Nucleosome geometry and internucleosomal interactions control the chromatin fiber conformation. *Biophys J*. 2008;95(8):3692–3705.
45. Öztürk MA, Cojocaru V, Wade RC. Dependence of chromatosome structure on linker histone sequence and posttranslational modification. *Biophys J*. 2018;114(10): 2363–2375.
46. Sottini A, Borgia A, Borgia MB, et al. Polyelectrolyte interactions enable rapid association and dissociation in high-affinity disordered protein complexes. *Nat Commun*. 2020;11(1):5736.
47. Milles S, Jensen MR, Lazert C, et al. An ultraweak interaction in the intrinsically disordered replication machinery is essential for measles virus function. *Sci Adv*. 2018;4(8): eaat7778.
48. Theillet FX, Binolfi A, Bekei B, et al. Structural disorder of monomeric alpha-synuclein persists in mammalian cells. *Nature*. 2016;530(7588):45–50.
49. Kay LE. NMR studies of protein structure and dynamics—a look backwards and forwards. *J Magn Reson*. 2011;213(2):492–494.
50. van Emmerik CL, van Ingen H. Unspinning chromatin: revealing the dynamic nucleosome landscape by NMR. *Prog Nucl Magn Reson Spectrosc*. 2019;110:1–19.
51. Zhou BR, Jiang J, Feng H, Ghirlando R, Xiao TS, Bai Y. Structural mechanisms of nucleosome recognition by linker histones. *Mol Cell*. 2015;59(4):628–638.
52. Abramov G, Velyvis A, Rennella E, Wong LE, Kay LE. A methyl-TROSY approach for NMR studies of high-molecular-weight DNA with application to the nucleosome core particle. *Proc Natl Acad Sci U S A*. 2020;117(23):12836–12846.
53. Bernado P, Svergun DI. Analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Methods Mol Biol*. 2012;896:107–122.
54. Henriques J, Arleth L, Lindorff-Larsen K, Skepo M. On the calculation of SAXS profiles of folded and intrinsically disordered proteins from computer simulations. *J Mol Biol*. 2018;430(16):2521–2539.

55. Misteli T, Gunjan A, Hock R, Bustin M, Brown DT. Dynamic binding of histone H1 to chromatin in living cells. *Nature*. 2000;408(6814):877–881.
56. Brucale M, Schuler B, Samori B. Single-molecule studies of intrinsically disordered proteins. *Chem Rev*. 2014;114(6):3281–3317.
57. Asher WB, Geggier P, Holsey MD, et al. Single-molecule FRET imaging of GPCR dimers in living cells. *Nat Methods*. 2021;18(4):397–405.
58. Lerner E, Barth A, Hendrix J, et al. FRET-based dynamic structural biology: challenges, perspectives and an appeal for open-science practices. *Elife*. 2021;10.
59. Schuler B, Hofmann H. Single-molecule spectroscopy of protein folding dynamics—expanding scope and timescales. *Curr Opin Struct Biol*. 2013;23:36–47.
60. Gunther E, Klauss A, Toro-Nahuelpan M, Schuler D, Hille C, Faivre D. The in vivo mechanics of the magnetotactic backbone as revealed by correlative FLIM-FRET and STED microscopy. *Sci Rep*. 2019;9(1):19615.
61. Lerner E, Cordes T, Ingargiola A, et al. Toward dynamic structural biology: two decades of single-molecule Förster resonance energy transfer. *Science*. 2018;359: eaan1133.
62. Heidarsson PO, Cecconi C. From folding to function: complex macromolecular reactions unraveled one-by-one with optical tweezers. *Essays Biochem*. 2021;65(1): 129–142.
63. Rudnizky S, Khamis H, Malik O, Melamed P, Kaplan A. The base pair-scale diffusion of nucleosomes modulates binding of transcription factors. *Proc Natl Acad Sci U S A*. 2019;116(25):12161–12166.
64. Gomes GW, Krzeminski M, Namini A, et al. Conformational ensembles of an intrinsically disordered protein consistent with NMR, SAXS, and single-molecule FRET. *J Am Chem Soc*. 2020;142(37):15697–15710.
65. Vendruscolo M, Dobson CM. Protein dynamics: Moore's law in molecular biology. *Curr Biol*. 2011;21(2):R68–R70.
66. Fan L, Roberts VA. Complex of linker histone H5 with the nucleosome and its implications for chromatin packing. *Proc Natl Acad Sci U S A*. 2006;103(22):8384–8389.
67. Collepardo-Guevara R, Schlick T. Chromatin fiber polymorphism triggered by variations of DNA linker lengths. *Proc Natl Acad Sci U S A*. 2014;111(22):8061–8066.
68. Teif VB, Kepper N, Yserentant K, Wedemann G, Rippe K. Affinity, stoichiometry and cooperativity of heterochromatin protein 1 (HP1) binding to nucleosomal arrays. *J Phys Condens Matter*. 2015;27(6):064110.
69. Watanabe S, Mishima Y, Shimizu M, Suetake I, Takada S. Interactions of HP1 bound to H3K9me3 dinucleosome by molecular simulations and biochemical assays. *Biophys J*. 2018;114(10):2336–2351.
70. Alston JJ, Soranno A, Holehouse AS. Integrating single-molecule spectroscopy and simulations for the study of intrinsically disordered proteins. *Methods*. 2021.
71. Sanders JC, Holmstrom ED. Integrating single-molecule FRET and biomolecular simulations to study diverse interactions between nucleic acids and proteins. *Essays Biochem*. 2021;65(1):37–49.
72. Rhodes D, Robinson PJJ. Structure of the '30 nm' chromatin fibre: a key role for the linker histone. *Curr Opin Struct Biol*. 2006;16:336–343.
73. Hergeth SP, Schneider R. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep*. 2015;16:1439–1453.
74. Hansen JC, Lu X, Ross ED, Woody RW. Intrinsic protein disorder, amino acid composition, and histone terminal domains*. *J Biol Chem*. 2006;281(4):1853–1856.
75. Goytisolo FA, Gerchman SE, Yu X, et al. Identification of two DNA-binding sites on the globular domain of histone H5. *EMBO J*. 1996;15:3421–3429.
76. Gibbs EB, Kriwacki RW. Linker histones as liquid-like glue for chromatin. *Proc Natl Acad Sci U S A*. 2018;115(47):11868–11870.

77. Izzo A, Kamieniarz K, Schneider R. The histone H1 family: specific members, specific functions? *Biol Chem*. 2008;389(4):333–343.
78. Th'ng JP, Sung R, Ye M, Hendzel MJ. H1 family histones in the nucleus: control of binding and localization by the C-terminal domain. *J Biol Chem*. 2005;280(30): 27809–27814.
79. Lorch Y, LaPointe JW, Kornberg RD. Nucleosomes inhibit the initiation of transcription but allow chain elongation with the displacement of histones. *Cell*. 1987;49(2): 203–210.
80. Roque A, Ponte I, Suau P. Post-translational modifications of the intrinsically disordered terminal domains of histone H1: effects on secondary structure and chromatin dynamics. *Chromosoma*. 2017;126(1):83–91.
81. Talasz H, Helliger W, Puschendorf B, Lindner H. In vivo phosphorylation of histone H1 variants during the cell cycle. *Biochemistry*. 1996;35(6):1761–1767.
82. Hergeth SP, Dundr M, Tropberger P, et al. Isoform-specific phosphorylation of human linker histone H1. 4 in mitosis by the kinase Aurora B. *J Cell Sci*. 2011;124(10): 1623–1628.
83. Chu C-S, Hsu P-H, Lo P-W, et al. Protein kinase A-mediated serine 35 phosphorylation dissociates histone H1. 4 from mitotic chromosome. *J Biol Chem*. 2011;286(41): 35843–35851.
84. Clausell J, Happel N, Hale TK, Doenecke D, Beato M. Histone H1 subtypes differentially modulate chromatin condensation without preventing ATP-dependent remodeling by SWI/SNF or NURF. *PLoS One*. 2009;4(10):e0007243.
85. Daujat S, Zeissler U, Waldmann T, Happel N, Schneider R. HP1 binds specifically to Lys26-methylated histone H1.4, whereas simultaneous Ser27 phosphorylation blocks HP1 binding. *J Biol Chem*. 2005;280(45):38090–38095.
86. Alexandrow MG, Hamlin JL. Chromatin decondensation in S-phase involves recruitment of Cdk2 by Cdc45 and histone H1 phosphorylation. *J Cell Biol*. 2005; 168(6):875–886.
87. Meergans T, Albig W, Doenecke D. Varied expression patterns of human H1 histone genes in different cell lines. *DNA Cell Biol*. 1997;16(9):1041–1049.
88. Weiss T, Hergeth S, Zeissler U, et al. Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. *Epigenetics Chromatin*. 2010;3(1):1–13.
89. Vaquero A, Scher M, Lee D, Erdjument-Bromage H, Tempst P, Reinberg D. Human SirT1 interacts with histone H1 and promotes formation of facultative heterochromatin. *Mol Cell*. 2004;16(1):93–105.
90. Terme J-M, Millán-Ariño L, Mayor R, et al. Dynamics and dispensability of variant-specific histone H1 Lys-26/Ser-27 and Thr-165 post-translational modifications. *FEBS Lett*. 2014;588(14):2353–2362.
91. Kamieniarz K, Izzo A, Dundr M, et al. A dual role of linker histone H1. 4 Lys 34 acetylation in transcriptional activation. *Genes Dev*. 2012;26(8):797–802.
92. Arents G, Burlingame RW, Wang B-C, Love WE, Moudrianakis EN. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A*. 1991;88(22):10148–10152.
93. Kurumizaka H, Kujirai T, Takizawa Y. Contributions of histone variants in nucleosome structure and function. *J Mol Biol*. 2021;433(6):166678.
94. Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol*. 2002;319(5):1097–1113.
95. Zentner GE, Henikoff S. Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol*. 2013;20(3):259–266.
96. Allan J, Harborne N, Rau DC, Gould H. Participation of core histone" tails" in the stabilization of the chromatin solenoid. *J Cell Biol*. 1982;93(2):285–297.

97. Zheng C, Hayes JJ. Structures and interactions of the core histone tail domains. *Biopolymers*. 2003;68(4):539–546.
98. Li G, Levitus M, Bustamante C, Widom J. Rapid spontaneous accessibility of nucleosomal DNA. *Nat Struct Mol Biol*. 2005;12(1):46–53.
99. Andresen K, Jimenez-Useche I, Howell SC, Yuan C, Qiu X. Solution scattering and FRET studies on nucleosomes reveal DNA unwrapping effects of H3 and H4 tail removal. *PLoS One*. 2013;8(11):e78587.
100. Dorigo B, Schalch T, Bystricky K, Richmond TJ. Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J Mol Biol*. 2003;327(1):85–96.
101. Shogren-Knaak M, Ishii H, Sun J-M, Pazin MJ, Davie JR, Peterson CL. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science*. 2006;311(5762):844–847.
102. Brower-Toland B, Wacker DA, Fulbright RM, Lis JT, Kraus WL, Wang MD. Specific contributions of histone tails and their acetylation to the mechanical stability of nucleosomes. *J Mol Biol*. 2005;346(1):135–146.
103. Anderson J, Lowary P, Widom J. Effects of histone acetylation on the equilibrium accessibility of nucleosomal DNA target sites. *J Mol Biol*. 2001;307(4):977–985.
104. Kurdistani SK, Tavazoie S, Grunstein M. Mapping global histone acetylation patterns to gene expression. *Cell*. 2004;117(6):721–733.
105. Martin BJE, Brind'Amour J, Kuzmin A, et al. Transcription shapes genome-wide histone acetylation patterns. *Nat Commun*. 2021;12(1):210.
106. Bannister AJ, Zegerman P, Partridge JF, et al. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*. 2001;410(6824):120–124.
107. Sugiyama K, Sugiura K, Hara T, et al. Aurora-B associated protein phosphatases as negative regulators of kinase activation. *Oncogene*. 2002;21(20):3103–3111.
108. Goto H, Yasui Y, Nigg EA, Inagaki M. Aurora-B phosphorylates histone H3 at serine28 with regard to the mitotic chromosome condensation. *Genes Cells*. 2002;7(1):11–17.
109. Rossetto D, Avvakumov N, Côté J. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics*. 2012;7(10):1098–1108.
110. Tatsuka M, Katayama H, Ota T, et al. Multinuclearity and increased ploidy caused by overexpression of the aurora-and Ipl1-like midbody-associated protein mitotic kinase in human cancer cells. *Cancer Res*. 1998;58(21):4811–4816.
111. Wang H, Huang Z-Q, Xia L, et al. Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. *Science*. 2001;293(5531):853–857.
112. Wysocka J, Swigut T, Milne TA, et al. WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell*. 2005;121(6):859–872.
113. Lomberk G, Bensi D, Fernandez-Zapico ME, Urrutia R. Evidence for the existence of an HP1-mediated subcode within the histone code. *Nat Cell Biol*. 2006;8(4):407–415.
114. Azzaz AM, Vitalini MW, Thomas AS, et al. Human heterochromatin protein 1α promotes nucleosome associations that drive chromatin condensation. *J Biol Chem*. 2014;289(10):6850–6861.
115. Kilic S, Bachmann AL, Bryan LC, Fierz B. Multivalency governs HP1alpha association dynamics with the silent chromatin state. *Nat Commun*. 2015;6:7313.
116. Jacobs SA, Khorasanizadeh S. Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science*. 2002;295(5562):2080–2083.
117. Hiragami-Hamada K, Shinmyozu K, Hamada D, et al. N-terminal phosphorylation of HP1α promotes its chromatin binding. *Mol Cell Biol*. 2011;31(6):1186–1200.
118. Bryan LC, Weilandt DR, Bachmann AL, et al. Single-molecule kinetic analysis of HP1-chromatin binding reveals a dynamic network of histone modification and DNA interactions. *Nucleic Acids Res*. 2017;45(18):10504–10517.

119. Hayakawa T, Haraguchi T, Masumoto H, Hiraoka Y. Cell cycle behavior of human HP1 subtypes: distinct molecular domains of HP1 are required for their centromeric localization during interphase and metaphase. *J Cell Sci*. 2003;116(16):3327–3338.
120. Bosch-Presegué L, Raurell-Vila H, Thackray JK, et al. Mammalian HP1 isoforms have specific roles in heterochromatin structure and organization. *Cell Rep*. 2017;21(8): 2048–2057.
121. Canzio D, Larson A, Narlikar GJ. Mechanisms of functional promiscuity by HP1 proteins. *Trends Cell Biol*. 2014;24(6):377–386.
122. Mishima Y, Watanabe M, Kawakami T, et al. Hinge and chromoshadow of HP1α participate in recognition of K9 methylated histone H3 in nucleosomes. *J Mol Biol*. 2013;425(1):54–70.
123. Muchardt C, Guillemé M, Seeler JS, Trouche D, Dejean A, Yaniv M. Coordinated methyl and RNA binding is required for heterochromatin localization of mammalian HP1α. *EMBO Rep*. 2002;3(10):975–981.
124. Badugu R, Yoo Y, Singh PB, Kellum R. Mutations in the heterochromatin protein 1 (HP1) hinge domain affect HP1 protein interactions and chromosomal distribution. *Chromosoma*. 2005;113(7):370–384.
125. Lachner M, O'Carroll D, Rea S, Mechtler K, Jenuwein T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*. 2001;410(6824):116–120.
126. LeRoy G, Weston JT, Zee BM, Young NL, Plazas-Mayorca MD, Garcia BA. Heterochromatin protein 1 is extensively decorated with histone code-like post-translational modifications. *Mol Cell Proteomics*. 2009;8(11):2432–2442.
127. Nishibuchi G, Machida S, Osakabe A, et al. N-terminal phosphorylation of HP1α increases its nucleosome-binding specificity. *Nucleic Acids Res*. 2014;42(20): 12498–12511.
128. Shimojo H, Kawaguchi A, Oda T, et al. Extended string-like binding of the phosphorylated HP1alpha N-terminal tail to the lysine 9-methylated histone H3 tail. *Sci Rep*. 2016;6:22527.
129. Liang JF, Yang VC, Vaynshteyn Y. The minimal functional sequence of protamine. *Biochem Biophys Res Commun*. 2005;336:653–659.
130. Ward WS, Coffey DS. DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biol Reprod*. 1991;44:569–574.
131. D'Ippolito RA, Minamino N, Rivera-Casas C, et al. Protamines from liverwort are produced by post-translational cleavage and C-terminal di-aminopropanelation of several male germ-specific H1 histones. *J Biol Chem*. 2019;294:16364–16373.
132. Rathke C, Baarends WM, Awe S, Renkawitz-Pohl R. Chromatin dynamics during spermiogenesis. *Biochim Biophys Acta*. 2014;1839(3):155–168.
133. Domenjoud L, Kremling H, Burfeind P, Maier WM, Engel W. On the expression of protamine genes in the testis of man and other mammals. *Andrologia*. 1991;23:333–337.
134. Carrell DT, Emery BR, Hammoud S. Altered protamine expression and diminished spermatogenesis: what is the link? *Hum Reprod Update*. 2007;13(3):313–327.
135. Balhorn R. The protamine family of sperm nuclear proteins. *Genome Biol*. 2007; 8(9):227.
136. Bode J, Willmitzer L, Opatz K. On the competition between protamines and histones: studies directed towards the understanding of spermiogenesis. *Eur J Biochem*. 1977;72(2):393–403.
137. Mao AH, Crick SL, Vitalis A, Chicoine CL, Pappu RV. Net charge per residue modulates conformational ensembles of intrinsically disordered proteins. *Proc Natl Acad Sci U S A*. 2010;107:8183–8188.
138. Balhorn R, Cosman M, Thornton K, et al. Protamine mediated condensation of DNA in mammalian sperm. In: *Male Gamete*. Cache River Press; 1999:55–70.

139. Bench GS, Friz AM, Corzett MH, Morse DH, Balhorn R. DNA and total protamine masses in individual sperm from fertile semen of selected mammals. *Cytometry*. 1996;23:263–271.
140. Balhorn R, Corzett M, Mazrimas J, Watkins B. Identification of bull protamine disulfides. *Biochemistry*. 1991;30:175–181.
141. Vilfan ID, Conwell CC, Hud NV. Formation of native-like mammalian sperm cell chromatin with folded bull protamine. *J Biol Chem*. 2004;279:20088–20095.
142. Ukogu OA, Smith AD, Devenica LM, et al. Protamine loops DNA in multiple steps. *Nucleic Acids Res*. 2020;48(11):6108–6119.
143. Hsiang MW, Cole RD. Structure of histone H1-DNA complex: effect of histone H1 on DNA condensation. *Proc Natl Acad Sci U S A*. 1977;74(11):4852–4856.
144. Brunner A, Nanni P, Mansuy I. Epigenetic marking of sperm by post-translational modification of histones and protamines. *Epigenetics Chromatin*. 2014;7:2.
145. Bustin M. High mobility group proteins. *Biochim Biophys Acta, Gene Regul Mech*. 1799;2010:1–2.
146. Nicolas RH, Goodwin GH. Isolation and analysis. In: Johns EW, ed. *The Chromosomal Proteins*. Academic Press; 1982:41–68.
147. Bustin M. Chromatin unfolding and activation by HMGN* chromosomal protiens. *Trends Biochem Sci*. 2001;26:431–437.
148. Phair RD, Misteli T. High mobility of proteins in the mammalian cell nucleus. *Nature*. 2000;404:604–609.
149. Alfonso PJ, Crippa MP, Hayes JJ, Bustin M. The footprint of chromosomal proteins HMG-14 and HMG-17 on chromatin subunits. *J Mol Biol*. 1994;236(1):189–198.
150. Postnikov YV, Trieschmann L, Rickers A, Bustin M. Homodimers of chromosomal proteins HMG-14 and HMG-17 in nucleosome cores. *J Mol Biol*. 1995;252:423–432.
151. Ueda T, Postnikov YV, Bustin M. Distinct domains in high mobility group N variants modulate specific chromatin modifications. *J Biol Chem*. 2006;281(15): 10182–10187.
152. Postnikov YV, Bustin M. Functional interplay between histone H1 and HMG proteins in chromatin. *Biochim Biophys Acta, Gene Regul Mech*. 1859;2016:462–467.
153. Ueda T, Catez F, Gerlitz G, Bustin M. Delineation of the protein module that anchors HMGN proteins to nucleosomes in the chromatin of living cells. *Mol Cell Biol*. 2008;28:2872–2883.
154. Trieschmann L, Martin B, Bustin M. The chromatin unfolding domain of chromosomal protein HMG-14 targets the N-terminal tail of histone H3 in nucleosomes. *Proc Natl Acad Sci U S A*. 1998;95:5468–5473.
155. Kato H, Van Ingen H, Zhou BR, et al. Architecture of the high mobility group nucleosomal protein 2-nucleosome complex as revealed by methyl-based NMR. *Proc Natl Acad Sci U S A*. 2011;108:12283–12288.
156. Lim JH, Catez F, Birger Y, et al. Chromosomal protein HMGN1 modulates histone H3 phosphorylation. *Mol Cell*. 2004;15:573–584.
157. Zhang Q, Wang Y. High mobility group proteins and their post-translational modifications. *Biochim Biophys Acta*. 2008;1784(9):1159–1166.
158. Bergel M, Herrera JE, Thatcher BJ, et al. Acetylation of novel sites in the nucleosomal binding domain of chromosomal protein HMG-14 by p300 alters its interaction with nucleosomes. *J Biol Chem*. 2000;275(15):11514–11520.
159. Murphy KJ, Cutter AR, Fang H, Postnikov YV, Bustin M, Hayes JJ. HMGN1 and 2 remodel core and linker histone tail domains within chromatin. *Nucleic Acids Res*. 2017;45:9917–9930.
160. Albright SC, Wiseman JM, Lange RA, Garrard WT. Subunit structures of different electrophoretic forms of nucleosomes. *J Biol Chem*. 1980;255(8):3673–3684.

161. van Houte LP, Chuprina VP, van der Wetering M, Boelens R, Kaptein R, Clevers H. Solution structure of the sequence-specific HMG box of the lymphocyte transcriptional activator Sox-4. *J Biol Chem*. 1995;270(51):30516–30524.
162. Sheflin LG, Spaulding SW. High mobility group protein 1 preferentially conserves torsion in negatively supercoiled DNA. *Biochemistry*. 1989;28:5658–5664.
163. Štros M, Štokrová J, Thomas JO. Dna looping by the HMG-box domains of HMG1 and modulation of DNA binding by the acidic C-terminal domain. *Nucleic Acids Res*. 1994;22:1044–1051.
164. Travers A. Recognition of distorted DNA structures by HMG domains. *Curr Opin Struct Biol*. 2000;10:102–109.
165. Thomas JO, Travers AA. HMG1 and 2, and related 'architectural' DNA-binding proteins. *Trends Biochem Sci*. 2001;26:167–174.
166. Jaouen S, De Koning L, Gaillard C, Muselíková-Polanská E, Štros M, Strauss F. Determinants of specific binding of HMGB1 protein to hemicatenated DNA loops. *J Mol Biol*. 2005;353:822–837.
167. Ugrinova I, Pasheva EA, Armengaud J, Pashev IG. In vivo acetylation of HMG1 protein enhances its binding affinity to distorted DNA structures. *Biochemistry*. 2001; 40(48):14655–14660.
168. Knapp S, Müller S, Digilio G, Bonaldi T, Bianchi ME, Musco G. The long acidic tail of high mobility group box 1 (HMGB1) protein forms an extended and flexible structure that interacts with specific residues within and between the HMG boxes. *Biochemistry*. 2004;43:11992–11997.
169. Stott K, Watson M, Howe FS, Grossmann JG, Thomas JO. Tail-mediated collapse of HMGB1 is dynamic and occurs via differential binding of the acidic tail to the A and B domains. *J Mol Biol*. 2010;403:706–722.
170. Kawase T, Sato K, Ueda T, Yoshida M. Distinct domains in HMGB1 are involved in specific intramolecular and nucleosomal interactions. *Biochemistry*. 2008;47: 13991–13996.
171. An W, Van Holde K, Zlatanova J. The non-histone chromatin protein HMG1 protects linker DNA on the side opposite to that protected by linker histones. *J Biol Chem*. 1998;273:26289–26291.
172. Jackson JB, Rill RL. Circular dichroism, thermal denaturation, and deoxyribonuclease I digestion studies of nucleosomes highly enriched in high mobility group proteins HMG 1 and HMG 2. *Biochemistry*. 1981;20:1042–1046.
173. Catez F, Yang H, Tracey KJ, Reeves R, Misteli T, Bustin M. Network of dynamic interactions between histone H1 and high-mobility-group proteins in chromatin. *Mol Cell Biol*. 2004;24:4321–4328.
174. Cato L, Stott K, Watson M, Thomas JO. The interaction of HMGB1 and linker histones occurs through their acidic and basic tails. *J Mol Biol*. 2008;384:1262–1272.
175. Štros M. HMGB proteins: interactions with DNA and chromatin. *Biochim Biophys Acta, Gene Regul Mech*. 1799;2010:101–113.
176. Polanska E, Pospisilova S, Stros M. Binding of histone H1 to DNA is differentially modulated by redox state of HMGB1. *PLoS One*. 2014;9(2):e89070.
177. Lehn DA, Elton TS, Johnson KR, Reeves R. A conformational study of the sequence specific binding of HMG-I (Y) with the bovine interleukin-2 cDNA. *Biochem Int*. 1988;16:963–971.
178. Evans JNS, Zajicek J, Nissen MS, Munske G, Smith V, Reeves R. 1H and 13C NMR assignments and molecular modelling of a minor groove DNA-binding peptide from the HMG-I protein. *Int J Pept Protein Res*. 1995;45:554–560.
179. Huth JR, Bewley CA, Nissen MS, et al. The solution structure of an HMG-I(Y)-DNA complex defines a new architectural minor groove binding motif. *Nat Struct Biol*. 1997;4:657–665.

180. Cui T, Leng F. Specific recognition of AT-rich DNA sequences by the mammalian high mobility group protein AT-hook 2: a SELEX study. *Biochemistry*. 2007;46:13059–13066.
181. Reeves R, Nissen MS. Interaction of high mobility group-I(Y) nonhistone proteins with nucleosome core particles. *J Biol Chem*. 1993;268:21137–21146.
182. Li O, Vasudevan D, Davey CA, Dröge P. High-level expression of DNA architectural factor HMGA2 and its association with nucleosomes in human embryonic stem cells. *Genesis*. 2006;44:523–529.
183. Zhao K, Kas E, Gonzalez E, Laemmli UK. SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. *EMBO J*. 1993;12:3237–3247.
184. Saitoh Y, Laemmli UK. Metaphase chromosome structure: bands arise from a differential folding path of the highly AT-rich scaffold. *Cell*. 1994;76:609–622.
185. Kishi Y, Fujii Y, Hirabayashi Y, Gotoh Y. HMGA regulates the global chromatin state and neurogenic potential in neocortical precursor cells. *Nat Neurosci*. 2012;15:1127–1133.
186. Wei T, Liu H, Chu B, et al. Phosphorylation-regulated HMGA1a-P53 interaction unveils the function of HMGA1a acidic tail phosphorylations via synthetic proteins. *Cell Chem Biol*. 2021;28:722–732.e8.
187. Iwafuchi-Doi M, Zaret KS. Pioneer transcription factors in cell reprogramming. *Genes Dev*. 2014;28(24):2679–2692.
188. Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M, Zaret KS. Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell*. 2015;161:555–568.
189. Clark KL, Halay ED, Lai E, Burley SK. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*. 1993;364:412–420.
190. Cirillo LA, McPherson CE, Bossard P, et al. Binding of the winged-helix transcription factor HNF3 to a linker histone site on the nucleosome. *EMBO J*. 1998;17:244–254.
191. Chaya D, Hayamizu T, Bustin M, Zaret KS. Transcription factor FoxA (HNF3) on a nucleosome at an enhancer complex in liver chromatin. *J Biol Chem*. 2001;276:44385–44389.
192. 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. *Mol Cell*. 2002;9:279–289.
193. Iwafuchi-Doi M, Donahue G, Kakumanu A, et al. The Pioneer transcription factor FoxA maintains an accessible nucleosome configuration at enhancers for tissue-specific gene activation. *Mol Cell*. 2016;62:79–91.
194. Iwafuchi M, Cuesta I, Donahue G, et al. Gene network transitions in embryos depend upon interactions between a pioneer transcription factor and core histones. *Nat Genet*. 2020;52(4):418–427.
195. Taube JH, Allton K, Duncan SA, Shen L, Barton MC. Foxa1 functions as a pioneer transcription factor at transposable elements to activate Afp during differentiation of embryonic stem cells. *J Biol Chem*. 2010;285:16135–16144.
196. Gómez-Márquez J, Rodríguez P. Prothymosin α is a chromatin-remodelling protein in mammalian cells. *Biochem J*. 1998;333:1–3.
197. George EM, Brown DT. Prothymosin α is a component of a linker histone chaperone. *FEBS Lett*. 2010;584:2833–2836.
198. Mosoian A. Intracellular and extracellular cytokine-like functions of prothymosin α: Implications for the development of immunotherapies. *Future Med Chem*. 2011;3:1199–1208.
199. Cast K, Damaschun H, Eckert K, et al. Prothymosin α: a biologically active protein with random coil conformation. *Biochemistry*. 1995;34:13211–13218.

200. Uversky VN, Gillespie JR, Millett IS, et al. Natively unfolded human prothymosin α adopts partially folded collapsed conformation at acidic pH. *Biochemistry*. 1999;38:15009–15016.
201. Bednar J, Hamiche A, Dimitrov S. H1-nucleosome interactions and their functional implications. *Biochim Biophys Acta*. 2016;1859(3):436–443.
202. Turner AL, Watson M, Wilkins OG, et al. Highly disordered histone H1-DNA model complexes and their condensates. *Proc Natl Acad Sci U S A*. 2018;115(47):11964–11969.
203. Kouzarides T. Chromatin modifications and their function. *Cell*. 2007;128(4):693–705.
204. Gangaraju VK, Bartholomew B. Mechanisms of ATP dependent chromatin remodeling. *Mutat Res*. 2007;618(1–2):3–17.
205. Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell*. 2004;116(1):51–61.
206. Mossink B, Negwer M, Schubert D, Nadif KN. The emerging role of chromatin remodelers in neurodevelopmental disorders: a developmental perspective. *Cell Mol Life Sci*. 2021;78(6):2517–2563.
207. Längst G, Manelyte L. Chromatin remodelers: from function to dysfunction. *Gene*. 2015;6:299–324.
208. Sandhu KS. Intrinsic disorder explains diverse nuclear roles of chromatin remodeling proteins. *J Mol Recognit*. 2009;22(1):1–8.
209. El Hadidy N, Uversky VN. Intrinsic disorder of the BAF complex: roles in chromatin remodeling and disease development. *Int J Mol Sci*. 2019;20(21):5260.
210. Aoki D, Awazu A, Fujii M, et al. Ultrasensitive change in nucleosome binding by multiple phosphorylations to the intrinsically disordered region of the histone chaperone FACT. *J Mol Biol*. 2020;432(16):4637–4657.
211. Santofimia-Castano P, Rizzuti B, Pey AL, et al. Intrinsically disordered chromatin protein NUPR1 binds to the C-terminal region of polycomb RING1B. *Proc Natl Acad Sci U S A*. 2017;114(31):E6332–E6341.
212. Yan L, Wu H, Li X, Gao N, Chen Z. Structures of the ISWI–nucleosome complex reveal a conserved mechanism of chromatin remodeling. *Nat Struct Mol Biol*. 2019;26(4):258–266.
213. Warren C, Shechter D. Fly fishing for histones: catch and release by histone chaperone intrinsically disordered regions and acidic stretches. *J Mol Biol*. 2017;429:2401–2426.
214. Hodges C, Kirkland JG, Crabtree GR. The many roles of BAF (mSWI/SNF) and PBAF complexes in cancer. *Cold Spring Harb Perspect Med*. 2016;6(8):a026930.
215. Goldberg AD, Banaszynski LA, Noh K-M, et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell*. 2010;140(5):678–691.
216. Drané P, Ouararhni K, Depaux A, Shuaib M, Hamiche A. The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. *Genes Dev*. 2010;24(12):1253–1265.
217. Dyer MA, Qadeer ZA, Valle-Garcia D, Bernstein E. ATRX and DAXX: mechanisms and mutations. *Cold Spring Harb Perspect Med*. 2017;7(3).
218. Gibbons RJ, Wada T, Fisher CA, et al. Mutations in the chromatin-associated protein ATRX. *Hum Mutat*. 2008;29(6):796–802.
219. Mitson M, Kelley LA, Sternberg MJE, Higgs DR, Gibbons RJ. Functional significance of mutations in the Snf2 domain of ATRX. *Hum Mol Genet*. 2011;20(13):2603–2610.
220. Varga-Weisz PD, Wilm M, Bonte E, Dumas K, Mann M, Becker PB. Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature*. 1997;388(6642):598–602.
221. Scacchetti A, Brueckner L, Jain D, et al. CHRAC/ACF contribute to the repressive ground state of chromatin. *Life Sci Alliance*. 2018;1(1):e201800024.

222. Hartlepp KF, Fernández-Tornero C, Eberharter A, Grüne T, Müller CW, Becker PB. The histone fold subunits of drosophila CHRAC facilitate nucleosome sliding through dynamic DNA interactions. *Mol Cell Biol*. 2005;25(22):9886–9896.
223. Bonaldi T, Längst G, Strohner R, Becker PB, Bianchi ME. The DNA chaperone HMGB1 facilitates ACF/CHRAC-dependent nucleosome sliding. *EMBO J*. 2002; 21(24):6865–6873.
224. Asfaha Y, Schrenk C, Alves Avelar LA, et al. Recent advances in class IIa histone deacetylases research. *Bioorg Med Chem*. 2019;27(22):115087.
225. Gurova K, Chang H-W, Valieva ME, Sandlesh P, Studitsky VM. Structure and function of the histone chaperone FACT—resolving FACTual issues. *Biochim Biophys Acta Gene Regul Mech*. 2018;1861(9):892–904.
226. Di Giorgio E, Brancolini C. Regulation of class IIa HDAC activities: it is not only matter of subcellular localization. *Epigenomics*. 2016;8(2):251–269.
227. Zhang CL, McKinsey TA, Olson EN. The transcriptional corepressor MITR is a signal-responsive inhibitor of myogenesis. *Proc Natl Acad Sci U S A*. 2001;98(13):7354.
228. Wang AH, Yang XJ. Histone deacetylase 4 possesses intrinsic nuclear import and export signals. *Mol Cell Biol*. 2001;21(17):5992–6005.
229. Black Joshua C, Van Rechem C, Whetstine JR. Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell*. 2012;48(4):491–507.
230. Martin C, Zhang Y. The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol*. 2005;6(11):838–849.
231. Bedford MT, Richard S. Arginine methylation: an emerging regulator of protein function. *Mol Cell*. 2005;18(3):263–272.
232. Sarmento OF, Svingen PA, Xiong Y, et al. The role of the histone methyltransferase enhancer of Zeste homolog 2 (EZH2) in the pathobiological mechanisms underlying inflammatory bowel disease (IBD)*. *J Biol Chem*. 2017;292(2):706–722.
233. Greer EL, Shi Y. Histone methylation: a dynamic mark in health, disease and inheritance. *Nat Rev Genet*. 2012;13(5):343–357.
234. Separovich RJ, Wong MWM, Chapman TR, Slavich E, Hamey JJ, Wilkins MR. Post-translational modification analysis of Saccharomyces cerevisiae histone methylation enzymes reveals phosphorylation sites of regulatory potential. *J Biol Chem*. 2021;296:100192.
235. Sen P, Dang W, Donahue G, et al. H3K36 methylation promotes longevity by enhancing transcriptional fidelity. *Genes Dev*. 2015;29(13):1362–1376.
236. Tennen RI, Berber E, Chua KF. Functional dissection of SIRT6: Identification of domains that regulate histone deacetylase activity and chromatin localization. *Mech Ageing Dev*. 2010;131(3):185–192.
237. Pan PW, Feldman JL, Devries MK, Dong A, Edwards AM, Denu JM. Structure and biochemical functions of SIRT6*. *J Biol Chem*. 2011;286(16):14575–14587.
238. Gil R, Barth S, Kanfi Y, Cohen HY. SIRT6 exhibits nucleosome-dependent deacetylase activity. *Nucleic Acids Res*. 2013;41(18):8537–8545.
239. Liu WH, Zheng J, Feldman JL, et al. Multivalent interactions drive nucleosome binding and efficient chromatin deacetylation by SIRT6. *Nat Commun*. 2020;11(1):5244.
240. Levendosky RF, Bowman GD. Asymmetry between the two acidic patches dictates the direction of nucleosome sliding by the ISWI chromatin remodeler. *Elife*. 2019;8: e45472.
241. Zhou K, Liu Y, Luger K. Histone chaperone FACT FAcilitates chromatin transcription: mechanistic and structural insights. *Curr Opin Struct Biol*. 2020;65:26–32.
242. Saunders A, Werner J, Andrulis ED, et al. Tracking FACT and the RNA polymerase II elongation complex through chromatin in vivo. *Science*. 2003;301(5636):1094–1096.
243. Gurard-Levin ZA, Quivy J-P, Almouzni G. Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annu Rev Biochem*. 2014;83(1):487–517.

244. Winkler DD, Muthurajan UM, Hieb AR, Luger K. Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events*. *J Biol Chem*. 2011;286(48):41883–41892.
245. Wang T, Liu Y, Edwards G, Krzizike D, Scherman H, Luger K. The histone chaperone FACT modulates nucleosome structure by tethering its components. *Life Sci Alliance*. 2018;1(4):e201800107.
246. Liu Y, Zhou K, Zhang N, et al. FACT caught in the act of manipulating the nucleosome. *Nature*. 2020;577(7790):426–431.
247. Mayanagi K, Saikusa K, Miyazaki N, et al. Structural visualization of key steps in nucleosome reorganization by human FACT. *Sci Rep*. 2019;9(1):10183.
248. Tsunaka Y, Ohtomo H, Morikawa K, Nishimura Y. Partial replacement of nucleosomal DNA with human FACT induces dynamic exposure and acetylation of histone H3 N-terminal tails. *iScience*. 2020;23(10):101641.
249. Szőllősi E, Bokor M, Bodor A, et al. Intrinsic structural disorder of DF31, a drosophila protein of chromatin decondensation and Remodeling activities. *J Proteome Res*. 2008;7(6):2291–2299.
250. Crevel G, Huikeshoven H, Cotterill S. Df31 is a novel nuclear protein involved in chromatin structure in Drosophila melanogaster. *J Cell Sci*. 2001;114(1):37.
251. Guillebault D, Cotterill S. The drosophila Df31 protein interacts with histone H3 tails and promotes chromatin bridging in vitro. *J Mol Biol*. 2007;373(4):903–912.
252. Schubert T, Pusch MC, Diermeier S, et al. Df31 protein and snoRNAs maintain accessible higher-order structures of chromatin. *Mol Cell*. 2012;48(3):434–444.
253. Fuxreiter M, Tompa P, Simon I, Uversky VN, Hansen JC, Asturias FJ. Malleable machines take shape in eukaryotic transcriptional regulation. *Nat Chem Biol*. 2008;4:728–737.
254. Staby L, O'Shea C, Willemoes M, Theisen F, Kragelund BB, Skriver K. Eukaryotic transcription factors: paradigms of protein intrinsic disorder. *Biochem J*. 2017;474(15):2509–2532.
255. Staller MV, Holehouse AS, Swain-Lenz D, Das RK, Pappu RV, Cohen BA. A high-throughput mutational scan of an intrinsically disordered acidic transcriptional activation domain. *Cell Syst*. 2018;6(4):444–455. [e6].
256. Vuzman D, Levy Y. Intrinsically disordered regions as affinity tuners in protein–DNA interactions. *Mol Biosyst*. 2012;8:47.
257. Tafvizi A, Huang F, Fersht AR, Mirny LA, van Oijen AM. A single-molecule characterization of p53 search on DNA. *Proc Natl Acad Sci U S A*. 2011;108(2):563–568.
258. Vuzman D, Azia A, Levy Y. Searching DNA via a "monkey Bar" mechanism: the significance of disordered tails. *J Mol Biol*. 2010;396:674–684.
259. Brodsky S, Jana T, Mittelman K, et al. Intrinsically disordered regions direct transcription factor in vivo binding specificity. *Mol Cell*. 2020;79(3):459–471.e4.
260. Zaret KS, Mango SE. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr Opin Genet Dev*. 2016;37:76–81.
261. Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006;126(4):663–676.
262. Lu Y, Brommer B, Tian X, et al. Reprogramming to recover youthful epigenetic information and restore vision. *Nature*. 2020;588(7836):124–129.
263. Jang J, Yoo J-E, Lee J-A, et al. Disease-specific induced pluripotent stem cells: a platform for human disease modeling and drug discovery. *Exp Mol Med*. 2012;44(3):202–213.
264. Xue B, Oldfield CJ, Van YY, Dunker AK, Uversky VN. Protein intrinsic disorder and induced pluripotent stem cells. *Mol Biosyst*. 2012;8(1):134–150.

265. Michael AK, Grand RS, Isbel L, et al. Mechanisms of OCT4-SOX2 motif readout on nucleosomes. *Science*. 2020;368(6498):1460–1465.
266. Dodonova SO, Zhu F, Dienemann C, Taipale J, Cramer P. Nucleosome-bound SOX2 and SOX11 structures elucidate pioneer factor function. *Nature*. 2020;580(7805):669–672.
267. Weiss MA. Floppy SOX: mutual induced fit in hmg (high-mobility group) box-DNA recognition. *Mol Endocrinol*. 2001;15(3):353–362.
268. Holmes ZE, Hamilton DJ, Hwang T, et al. The Sox2 transcription factor binds RNA. *Nat Commun*. 2020;11(1):1805.
269. Tan C, Takada S. Nucleosome allostery in pioneer transcription factor binding. *Proc Natl Acad Sci U S A*. 2020;117(34):20586–20596.
270. Shammas SL. Mechanistic roles of protein disorder within transcription. *Curr Opin Struct Biol*. 2017;42:155–161.
271. Robinson JL, Carroll JS. FoxA1 is a key mediator of hormonal response in breast and prostate cancer. *Front Endocrinol (Lausanne)*. 2012;3:68.
272. Sutinen P, Rahkama V, Rytinki M, Palvimo JJ. Nuclear mobility and activity of FOXA1 with androgen receptor are regulated by SUMOylation. *Mol Endocrinol*. 2014;28(10):1719–1728.
273. Hou L, Wei Y, Lin Y, et al. Concurrent binding to DNA and RNA facilitates the pluripotency reprogramming activity of Sox2. *Nucleic Acids Res*. 2020;48(7):3869–3887.
274. Aksoy I, Jauch R, Eras V, et al. Sox transcription factors require selective interactions with Oct4 and specific transactivation functions to mediate reprogramming. *Stem Cells*. 2013;31(12):2632–2646.
275. Xhani S, Lee S, Kim HM, et al. Intrinsic disorder controls two functionally distinct dimers of the master transcription factor PU.1. *Sci Adv*. 2020;6(8):eaay3178.
276. Minderjahn J, Schmidt A, Fuchs A, et al. Mechanisms governing the pioneering and redistribution capabilities of the non-classical pioneer PU.1. *Nat Commun*. 2020;11(1):402.
277. Chanda S, Ang CE, Davila J, et al. Generation of induced neuronal cells by the single reprogramming factor ASCL1. *Stem Cell Rep*. 2014;3(2):282–296.
278. Baronti L, Hosek T, Gil-Caballero S, et al. Fragment-based NMR study of the conformational dynamics in the bHLH transcription factor Ascl1. *Biophys J*. 2017;112(7):1366–1373.
279. Fernandez Garcia M, Moore CD, Schulz KN, et al. Structural features of transcription factors associating with nucleosome binding. *Mol Cell*. 2019;75(5):921–932.e6.
280. Hiragami-Hamada K, Nakayama J-I. Do the charges matter?—Balancing the charges of the chromodomain proteins on the nucleosome. *J Biochem*. 2019;165(6):455–458.
281. Pepenella S, Murphy KJ, Hayes JJ. Intra-and inter-nucleosome interactions of the core histone tail domains in higher-order chromatin structure. *Chromosoma*. 2014;123(1–2):3–13.
282. Rowell JP, Simpson KL, Stott K, Watson M, Thomas JO. HMGB1-facilitated p53 DNA binding occurs via HMG-box/p53 transactivation domain interaction, regulated by the acidic tail. *Structure*. 2012;20(12):2014–2024.
283. Štros M, Kučírek M, Sani SA, Polanská E. HMGB1-mediated DNA bending: Distinct roles in increasing p53 binding to DNA and the transactivation of p53-responsive gene promoters. *Biochim Biophys Acta, Gene Regul Mech*. 1861;2018:200–210.
284. Jayaraman L, Moorthy NC, Murthy KG, Manley JL, Bustin M, Prives C. High mobility group protein-1 (HMG-1) is a unique activator of p53. *Genes Dev*. 1998;12(4):462–472.

285. Thomas JO, Stott K. H1 and HMGB1: modulators of chromatin structure. *Biochem Soc Trans*. 2012;40:341–346.
286. Kim K, Jeong KW, Kim H, et al. Functional interplay between p53 acetylation and H1.2 phosphorylation in p53-regulated transcription. *Oncogene*. 2012;31(39):4290–4301.
287. Tanaka H, Iguchi N, Isotani A, et al. HANP1/H1T2, a novel histone H1-like protein involved in nuclear formation and sperm fertility. *Mol Cell Biol*. 2005;25(16):7107–7119.
288. Holmstrom ED, Liu Z, Nettels D, Best RB, Schuler B. Disordered RNA chaperones can enhance nucleic acid folding via local charge screening. *Nat Commun*. 2019;10(1):2453.