ACCEPTED MANUSCRIPT

Disordered domains in chromatin-binding proteins

Matthew Watson and Katherine Stott

Disordered domains in chromatin-binding proteins

Matthew Watson and Katherine Stott

Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, UK

Abstract

Chromatin comprises proteins, DNA and RNA, and its function is to condense and package the genome in a way that allows the necessary transactions such as transcription, replication and repair to occur in a highly organised and regulated manner. The packaging of chromatin is often thought of in a hierarchical fashion starting from the most basic unit of DNA packaging, the nucleosome, to the condensation of nucleosomal "beads on a string" by linker histones to form the 30 nm fibre and eventually large chromatin domains. However, a picture of a more heterogeneous, dynamic and liquid-like assembly is emerging, in which intrinsically disordered proteins (IDPs) and proteins containing intrinsically disordered regions (IDRs) play a central role. Disorder features at all levels of chromatin organisation, from the histone tails, which are sites of extensive post translational modification that change the fate of the underlying genomic information, right through to transcription hubs, and the recently elucidated roles of IDPs and IDRs in the condensation of large regions of the genome through liquid-liquid phase separation.

Introduction

For many decades it has generally been accepted that in order to function, proteins must adopt a defined three-dimensional structure (this has become known as the structure-function paradigm). However, it has increasingly become apparent that a large percentage of the proteome consists of proteins or protein regions that lack well-defined order (1, 2), rather they exist as heterogeneous ensembles of rapidly interconverting states for which no one set of coordinates can sufficiently describe their equilibrium conformational properties. Structured proteins are most commonly associated with catalysis, transport, biosynthesis, metabolism and membrane-spanning, while IDPs and IDRs, referred to collectively here as

IDP(R)s, are typically engaged in signalling and regulatory processes that involve protein-protein and protein-nucleic acid interactions, such as differentiation, transcriptional regulation, DNA condensation, cell cycle and RNA processing (3), i.e. many functions that are relevant to chromatin. IDP(R)s are central to chromatin organisation, from the core and linker histones (4) that are responsible for nucleosome formation and condensation, to the machinery responsible for the regulation and execution of transcriptional programmes (5). In this review, rather than focusing on the traditional hierarchical view of chromatin structure, we focus on some of the emerging concepts regarding the role of IDP(R)s in the packaging of chromatin and the regulation of nuclear processes.

The disorder continuum

Historically, disordered proteins have collectively been referred to as IDPs. However, it is important to differentiate between true IDPs, where the entire sequence lacks stable structure, and those proteins that contain a modular architecture comprising both folded domains and IDRs, where the interplay between the structured and disordered regions may be critical to understanding the protein's function (6) (Fig. 1A). It is also important to understand that intrinsic disorder represents a continuum of ensemble states between folded proteins and true random coils, that contain varying degrees of transient or stable secondary structure (Fig. 1B). Rather than being functionally inert regions of the proteome, IDPs and IDRs have been ascribed to a wide variety of functions (3), including catalysis, which was thought to require a folded 3D structure (7).

SLiMs and MoRFs

A common feature of IDP(R)s is that they make multivalent and/or promiscuous interactions. As such, hubs that form the centre of interaction networks and integrate the signals from multiple pathways are enriched in disorder (9). For example, the transcription factor and tumour suppressor p53 has more than 500 direct interaction partners in the STRING database (10). p53 is typical of transcription factors in having a modular architecture. It comprises folded DNA-binding and tetramerisation domains and several long disordered regions, which account for about 42% of the protein (11)(Figure 1C). Furthermore, the disordered regions are the sites of extensive post translational modifications (PTMs) (12). Regarding promiscuity, the conformational plasticity of IDP(R)s

makes them ideally suited to binding multiple partners. Interactions are often mediated through short linear motifs (SLiMs) or molecular recognition features (MoRFs). SLiMs comprise a short stretch of 3-10 amino acids; their surrounding context may be ordered or disordered and may also contribute to the free energy of binding (13, 14). Upon interaction of a SLiM with a structured partner, secondary structure may be induced. MoRFs are slightly longer (around 20 amino acids), are unique to IDP(R)s, and always undergo a disorder-to-order transition upon interaction (15). The compact nature of these linear motifs gives the potential for a large number of binding modules to cluster or even overlap in a given length polypeptide and in many cases multiple low affinity motifs cooperate in a multivalent interaction. The cooperativity can take different forms: multiple distinct pairwise-specific interactions (avidity); multiple epitopes along the ligand that bind a single site on a partner (allovalency); or multiple interchangeable interaction sites on both the ligand and partner ("fuzziness") (16).

Another feature of SLiMs and MoRFs is that their conformational plasticity allows multiple IDPs to bind to the same protein using different linear motifs (many-to-one interactions), or the same motif to bind to multiple partners, sometimes adopting different conformations (one-to-many interactions) (9). Promiscuity in p53, which forms a diverse range of complexes, is perhaps the most striking example of this (17, 18).

Tipping the balance by PTMs

The lack of stable structure enhances accessibility to modifying enzymes and as a result they are the predominant sites of PTM (3, 13). These modifications can have a pronounced effect on their properties. SLiMs usually interact transiently and reversibly with their partners, and therefore their interactions are easily modulated by PTMs that tune their structural and physicochemical properties. PTMs can modulate the properties of the structural ensemble, particularly if it is sensitive to charge distribution (19). Modifications may cause changes in the position of an IDP(R) in the coil-globule continuum (20) or a disorder-to-order transition (21).

The nuclear proteome is enriched in IDP(R)s

IDP(R)s typically carry out signalling and regulatory functions. It is therefore not surprising that proteomic and bioinformatic analyses indicate that the nucleus is significantly enriched

in disorder (22–24). There is a correlation between the level of protein disorder and organism complexity, at least between prokaryotes and eukaryotes (25), which coincides with the appearance of the nucleus. Moreover, organism complexity also correlates with more elaborate regulation of gene expression (26). For example, the most complex multicellular lineages (i.e., those with embryonic development, Metazoa and Embryophyta) have the most complex transcription factor (TF) repertoires (27).

Disorder in transcription hubs

TFs integrate cellular signals and, through the recognition of cis-regulatory elements in the DNA and recruitment of coactivators and RNA polymerase, orchestrate specific patterns of gene expression (28, 29). It has long been realised that TFs are highly enriched in intrinsic disorder, with over 80% predicted to contain extensive IDRs (5). IDRs are thought to play a role in the TF binding site search process, for example the "monkey bar" model of inter-strand exchange, as described for the HOX TFs, in which the N-terminal tail can bind a neighbouring strand and facilitate transfer (30) (Fig. 2A). It has also been proposed that non-specific association of the p53 disordered C-terminal region facilitates scanning along DNA (31, 32) (Fig. 2B). Another mechanism by which IDRs can modulate DNA binding is by competition, for example in the HMG box family of architectural transcription factors, where the acidic regulatory domains interact with the DNA-binding surfaces and negatively regulate binding to DNA (33–35) (Fig. 2C). This autoinhibited state is alleviated by interactions with partners, for example with the H3 N-terminal tail, presumed to position HMGB1 at the nucleosome dyad (36). Tail binding also negatively regulates protein-protein interactions, including binding of the HMGB1 A box to a MoRF in the N-terminal p53 TAD (37).

Within TFs, the degree of disorder is much higher in transactivation domains (TADs) and their flanking regions than in DNA-binding domains, with some notable exceptions (for example SOX domain TFs, which contain an HMG box that becomes structured on binding DNA (38), and members of the HMGA family (39, 40)). Recent work on the acidic central activation domain of the yeast TF Gcn4 has suggested that the acidic residues and intrinsic disorder act to maintain the hydrophobic amino acids in a solvent-exposed state and prevent them from driving collapse (41, 42). This is supported by the finding that short sequence patterns of negatively-charged and aromatic residues in single or overlapping

SLiMs/MoRFs are a generic feature for TAD functionality (43). Furthermore, recent in-cell imaging (44) showed that TF low-complexity domains (LCDs) concentrate the transcription machinery in local high-concentration hubs, which stabilise DNA binding, recruit RNA polymerase II, and activate transcription. Mutations in the LCD of a fusion oncogene (EWS-FLI1) that removed hydrophobic residues diminished self-association and transcriptional activation, highlighting the importance of dynamic LCD-LCD interactions in hub assembly and transcription.

New lessons from IDP(R)-IDP(R) and IDP(R)-nucleic acid interactions

The prevailing view has been that IDP(R) interactions involve a disorder-to-order transition. However, an increasing number of interactions have been characterised where varying degrees of disorder are present in the complex and, due to the presence of multiple functionally-relevant states, these have been termed "fuzzy complexes" (45, 46). Recently, two extreme examples of these have been described for complexes involving the linker histone H1, the highly disordered C-terminal tail of which is responsible for condensation of the nucleosomal "beads on a string" structure in chromatin. The first involves the interaction of H1.0 with the intrinsically-disordered linker histone chaperone Prothymosin α (ProTα) (47), and the second the interaction of the intrinsically disordered H1.11L C-terminal tail with DNA (48). The complexes have picomolar and nanomolar K_d, respectively, although the H1/ProTα affinity has been debated, with a more recent study placing it in the high nanomolar range (49). What is interesting in these two studies is that the protein(s) remain highly disordered in the complex despite their high affinities. The affinity is due to opposing charges: H1 is highly basic while ProTα and DNA are highly acidic; all can essentially be thought of as polyelectrolytes. Other known partners for H1 having highly disordered acidic regions are hFACT (50), a histone chaperone critical for nucleosome reorganization during replication, transcription, and DNA repair (51), and HMGB1. H1 and HMGB1 have opposing effects on chromatin condensation (52). A direct interaction between the basic C-terminal tail of H1 and the acidic tail of HMGB1 was observed, again, without evidence of stable secondary structure (53). H1 interacts with over 100 nuclear proteins, and may play a much wider role in the nucleus than an architectural factor in chromatin structure (50, 54, 55). Approximately 1/3 of these interactions appear to be mediated by the C-terminal tail (50),

suggesting that these extreme examples of fuzzy H1 partner complexes may be a common theme.

The binding of polycations (such as H1) to polyanions (such as DNA or ProTα) is driven by a large positive entropy contribution to the free energy arising from the displacement of counterions and solvent associated with the solvent-exposed charges (56, 57). Further, unlike complexes formed by disorder-to-order transitions, entropy losses are minimised by the retention of conformational freedom. The fact that these complexes exist as an ensemble of rapidly interconverting structures in which multiple transient contacts are constantly rearranging, raises intriguing possibilities for IDP(R) interactions. It is often assumed that IDP(R)s form transient, low affinity complexes such as those involved in signalling. Conversely, high affinities are usually linked to long residence times. However, in these extreme examples of "fuzzy" complexes, one could envisage the situation whereby another binding partner could progressively peel away the dynamic interactions, *via* a ternary complex, and eventually fully exchange on a timescale that would not be possible in an analogous situation involving high-affinity interactions between structured proteins, where binding or unbinding events are usually concerted. Taken together, these observations suggest that "fuzzy" interactions between polyelectrolytes may act to catalyse the exchange of proteins that would otherwise associate too stably (Fig. 3A and B). The presence of factors such as the chaperone ProTα (58), HMGB1 (52) and hFACT (51) that could potentially engage in "fuzzy" interactions with linker histones is therefore consistent with the rapid mobility observed for H1 in the nucleus (59), despite its high affinity for mononucleosomes and nucleosome arrays (60). In the case of ProTα, this also allows H1 to be tightly sequestered in order to prevent aberrant interactions with DNA while still allowing H1 to be readily deposited at the appropriate sites in chromatin. Finally, H1 affinity is also known to be modified by phosphorylation; mutation of the five phosphorylation sites in H1 weakened interactions with linker DNA leading to a more open chromatin structure (61). Dynamic competition may also facilitate access to enzymes that post-translationally modify proteins, for example kinases that phosphorylate the consensus SPxK SLiM sites in the H1 C-terminal tail, without the need for complete dissociation of the complex (Fig. 3C).

Phase separation: a fuzzy perspective on chromatin higher-order structure

The nucleus contains numerous compartments that are not bound by membranes, such as nucleoli, Cajal bodies and Promyelocytic Leukemia (PML) nuclear bodies (62). It has recently become apparent that these membraneless organelles have liquid like properties and can be formed by liquid-liquid phase separation (63) in an analogous manner to that observed between oil and water, leading to a huge interest in the role of phase separation in the regulation of biological processes. Protein disorder is an important driver of phase separation, since the formation of many of these membraneless organelles is driven by the dynamic multivalent interactions between highly-disordered proteins and RNA (64). These condensates directly interact with the underlying chromatin and hence potentially play a role in its organisation. For example, recent work suggests that phase separated droplets preferentially nucleate at regions of low chromatin density and that as they grow they exclude non-targeted genomic regions, suggesting that these condensates can both sense and restructure the local chromatin environment (65). In addition to regulating spatiotemporal patterning of macromolecules, the conditions inside phase separated droplets might be significantly different from the surroundings, allowing them to partition molecules without the need for ATP-dependent processes and, through locally concentrating them, to facilitate interactions or reactions to take place that would normally be slow or thermodynamically unfavourable (66). For example, phase separated droplets containing Ddx4 were able to melt double stranded DNA (67). It is also becoming apparent that aberrant control of phase separation may play an important role in disease processes. A continuous spectrum of material properties has been described for biological condensates ranging from liquid to gel and amyloid, with mutations that perturb this balance being linked to disease (68, 69).

In addition to the well documented role in the formation of membraneless organelles, IDP(R)s and their involvement in phase separation may play a much broader role in chromatin organisation. Growing evidence is shifting our view of chromatin away from a static crystal-like structure to one in which it assumes a more dynamic liquid-like state (70). Recent live-cell super-resolution microscopy studies support this finding and suggest that nucleosomes form compact domains in which the constituent nucleosomes move coherently, and that these structures are ''liquid drops'' rather than loose bundles of fibres or extended loops (71). Newly developed electron tomography techniques for studying

chromatin also conclude that, in contrast to rigid fibres with a long persistence length, chromatin exists predominantly as a flexible disordered fibre with a diameter of 5-24 nM (72). This is supported by structural studies of reconstituted arrays that reveal an ensemble of dynamic nucleosome zigzag arrangements that are easily perturbed by salt concentration, chemical cross-linking, and crystal packing (73). One interesting view is that nucleosomes with their disordered core and linker histone tails represent "fuzzy" structures capable of bringing clusters of nucleosomes together and anchoring them to structures like the nuclear lamina *via* promiscuous multivalent interactions (74). The recent evidence that the disordered H1 tail is able to form dynamic phase separated condensates upon interaction with DNA (48) has led to the proposal that H1 might form the "liquid like molecular glue" that holds these condensed clusters of nucleosomes together (75) (Fig. 4). Interestingly, under low salt conditions, phase-separated H1/DNA complexes showed evidence of long-range order in the form of "ψ-DNA" (76, 77), which implies an ordered twisted helical packing of the DNA (48). The presence of ψ-DNA could be an indication of how H1 is able to organise the linker DNA in a nucleosome cluster or fibre, and that liquid phases formed by condensation might have some higher order organisational capability despite their dynamic properties. In the same study, phosphorylation destabilised this structure and reduced partitioning from the bulk to the condensate, suggesting how posttranslational modifications may dramatically alter the structure of nucleosome clusters.

Another observation that supports the notion of a more fluid chromatin structure comes from recent work on heterochromatin formation. Heterochromatin domains are morphologically distinct and transcriptionally silent domains, enriched in nucleosomes displaying H3K9 methylation, which is recognised by Heterochromatin Protein 1 (HP1) (78). Two recent studies have shown that HP1 from both *Drosophila* and humans is able to form phase separated condensates (79, 80). Interestingly, human HP1 requires an additional level of regulation – phosphorylation – to form an open conformation that is capable of making the polyvalent interactions necessary for phase separation. The *in vivo* observations made by Strom *et al*. (79), showed that during the early stages of *Drosophila* embryo development when heterochromatin is first formed, GFP-HP1a formed spherical foci that displayed characteristics of phase separated liquids. Intriguingly, they also found that as heterochromatin formation became more established an increasing fraction of the HP1

became immobile and the droplets less spherical, suggesting a transition to a more gel-like structure that may represent stable repressed heterochromatin.

In conclusion, despite the absence of discrete membrane-bound compartments, a remarkable number of complex processes are organised both spatially and temporally within the nucleus. Phase separation driven by IDP(R)/nucleic acid interactions provides an attractive way to explain how this organisation can be achieved. The increasing evidence that phase transitions are ubiquitous in nuclear architecture and transcriptional regulation does not necessarily exclude previous models, but suggests that our understanding of these processes may need to be revised to take it into account.

Summary points

1) Intrinsic disorder is a feature of many nuclear proteins that organise and regulate chromatin.

2) IDP(R)s impart dynamic features, observations of which are challenging the portrayal of the highly organised hierarchical chromatin structure by crystallography and electron microscopy.

3) IDP(R)s often mediate fluid networks of multivalent interactions, which retain a high degree of disorder, enabling molecular exchange processes or driving liquid-liquid phase separation.

4) These dynamic features of chromatin are highly responsive to regulatory signals, reflecting the need to carry out complex processes involved with genomic maintenance, and the execution of specific transcriptional programmes.

5) Elucidating how these dynamic interaction networks and phase transitions orchestrate chromatin structure and function will require a multidisciplinary approach and promises to be an exciting area for many years to come.

Funding

Abbreviations

DBD, DNA-binding domain; HP1, heterochromatin protein 1; IDP, intrinsically disordered protein; IDR, intrinsically disordered region; LCD, low complexity domain; MoRF, molecular recognition feature; PML, Promyelocytic Leukemia; ProTα, prothymosin α; PRR, proline-rich region; PTM, post-translational modification; RD, regulatory domain; SLiM, short linear motifs; TAD, transactivation domain; TD, tetramerisation domain; TF, transcription factor.

References

1. Dunker,A.K., Lawson,J.D., Brown,C.J., Romero,P., Oh,J.S., Oldfield,C.J., Campen,A.M., Catherine,M., Hipps,K.W., Ausio,J., *et al.* (2001) Intrinsically Disordered Protein. *J. Mol. Graph. Model.*, **19**, 26–59.

2. Wright,P.E. and Dyson,H.J. (1999) Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.*, **293**, 321–331.

3. Xie,H., Vucetic,S., Iakoucheva,L.M., Oldfield,C.J., Dunker,A.K., Uversky,V.N. and Obradovic,Z. (2007) Functional anthology of intrinsic disorder. 1. Biological processes and functions of proteins with long disordered regions. *J. Proteome Res.*, **6**, 1882–1898.

4. Hansen,J.C., Lu,X., Ross,E.D. and Woody,R.W. (2006) Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J. Biol. Chem.*, **281**, 1853–1856.

5. Liu,J., Perumal,N.B., Oldfield,C.J., Su,E.W., Uversky,V.N. and Dunker,A.K. (2006) Intrinsic disorder in transcription factors. *Biochemistry*, **45**, 6873–6888.

6. Dunker,A.K., Babu,M.M., Barbar,E., Blackledge,M., Bondos,S.E., Dosztányi,Z., Dyson,H.J., Forman-Kay,J., Fuxreiter,M., Gsponer,J., *et al.* (2013) What's in a name? Why these proteins are intrinsically disordered. *Intrinsically Disord. Proteins*, **1**, e241571-5.

7. Chakrabortee,S., Meersman,F., Kaminski Schierle,G.S., Bertoncini,C.W., McGee,B., Kaminski,C.F. and Tunnacliffe,A. (2010) Catalytic and chaperone-like functions in an intrinsically disordered protein associated with desiccation tolerance. *Proc. Natl. Acad. Sci. U. S. A.*, **107**, 16084–16089.

8. Guex,N. and Peitsch,M.C. (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*, **18**, 2714–2723.

9. Hu,G., Wu,Z., Uversky,V.N. and Kurgan,L. (2017) Functional analysis of human hub proteins and their interactors involved in the intrinsic disorder-enriched interactions. *Int. J. Mol. Sci.*, **18**, 1–40.

10. Szklarczyk,D., Franceschini,A., Wyder,S., Forslund,K., Heller,D., Huerta-Cepas,J., Simonovic,M., Roth,A., Santos,A., Tsafou,K.P., *et al.* (2015) STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.*, **43**, D447–D452.

11. Joerger,A.C. and Fersht,A.R. (2008) Structural biology of the tumor suppressor p53. *Annu. Rev. Biochem.*, **77**, 557–82.

12. Dai,C. and Gu,W. (2010) p53 post-translational modification: deregulated in tumorigenesis. *Trends Mol. Med.*, **16**, 528–536.

13. Van Roey,K., Uyar,B., Weatheritt,R.J., Dinkel,H., Seiler,M., Budd,A., Gibson,T.J. and Davey,N.E. (2014) Short linear motifs: Ubiquitous and functionally diverse protein interaction modules directing cell regulation. *Chem. Rev.*, **114**, 6733–6778.

14. Davey,N.E., Van Roey,K., Weatheritt,R.J., Toedt,G., Uyar,B., Altenberg,B., Budd,A., Diella,F., Dinkel,H. and Gibson,T.J. (2012) Attributes of short linear motifs. *Mol. Biosyst.*, **8**, 268–281.

15. Vacic,V., Oldfield,C.J., Mohan,A., Radivojac,P., Cortese,M.S., Uversky,V.N. and Dunker,A.K. (2007) Characterization of molecular recognition features, MoRFs, and their binding partners. *J. Proteome Res.*, **6**, 2351–2366.

16. Olsen,J.G., Teilum,K. and Kragelund,B.B. (2017) Behaviour of intrinsically disordered proteins in protein–protein complexes with an emphasis on fuzziness. *Cell. Mol. Life Sci.*, **74**, 3175–3183.

17. Uversky,V.N. and Dunker,A.K. (2010) Understanding protein non-folding. *Biochim. Biophys. Acta - Proteins Proteomics*, **1804**, 1231–1264.

18. Hsu,W.L., Oldfield,C.J., Xue,B., Meng,J., Huang,F., Romero,P., Uversky,V.N. and Dunker,A.K. (2013) Exploring the binding diversity of intrinsically disordered proteins involved in one-to-many binding. *Protein Sci.*, **22**, 258–273.

19. Das,R.K. and Pappu,R. V. (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, 13392–13397.

20. Firman,T. and Ghosh,K. (2018) Sequence charge decoration dictates coil-globule transition in intrinsically disordered proteins. *J. Chem. Phys.*, **148**, 1–11.

21. Bah,A., Vernon,R.M., Siddiqui,Z., Krzeminski,M., Muhandiram,R., Zhao,C., Sonenberg,N., Kay,L.E. and Forman-Kay,J.D. (2015) Folding of an intrinsically disordered protein by

phosphorylation as a regulatory switch. *Nature*, **519**, 106–109.

22. Frege,T. and Uversky,V.N. (2015) Intrinsically disordered proteins in the nucleus of human cells. *Biochem. Biophys. Reports*, **1**, 33–51.

23. Skupien-Rabian,B., Jankowska,U., Swiderska,B., Lukasiewicz,S., Ryszawy,D., Dziedzicka-Wasylewska,M. and Kedracka-Krok,S. (2015) Proteomic and bioinformatic analysis of a nuclear intrinsically disordered proteome. *J. Proteomics*, **130**, 76–84.

24. Wang,C., Uversky,V.N. and Kurgan,L. (2016) Disordered nucleiome: abundance of intrinsic disorder in the DNA- and RNA-binding proteins in 1121 species from Eukaryota, Bacteria and Archaea. *Proteomics*, **16**, 1486–1498.

25. Schad,E., Tompa,P. and Hegyi,H. (2011) The relationship between proteome size , structural disorder and organism complexity. *Genome Biol.*, **12**, R120.

26. Levine,M. and Tjian,R. (2003) Transcription regulation and animal diversity. *Nature*, **424**, 147–151.

27. Mendoza,A. De and Sestak,M.S. (2013) Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages. *Proc. Natl. Acad. Sci. U. S. A.*, **110**, E4858–E4866.

28. Spitz,F. and Furlong,E.E.M. (2012) Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.*, **13**, 613–626.

29. Reiter,F., Wienerroither,S. and Stark,A. (2017) Combinatorial function of transcription factors and cofactors. *Curr. Opin. Genet. Dev.*, **43**, 73–81.

30. Vuzman,D., Azia,A. and Levy,Y. (2010) Searching DNA via a 'Monkey Bar' Mechanism: The Significance of Disordered Tails. *J. Mol. Biol.*, **396**, 674–684.

31. Melero,R., Rajagopalan,S., Lazaro,M., Joerger,A.C., Brandt,T., Veprintsev,D.B., Lasso,G., Gil,D., Scheres,S.H.W., Carazo,J.M., *et al.* (2011) Electron microscopy studies on the quaternary structure of p53 reveal different binding modes for p53 tetramers in complex with DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 557–562.

32. Tafvizi,A., Huang,F., Fersht,A.R., Mirny,L.A. and van Oijen,A.M. (2011) A single-molecule characterization of p53 search on DNA. *Proc. Natl. Acad. Sci. U. S. A.*, **108**, 563–568.

33. Watson,M., Stott,K. and Thomas,J.O. (2007) Mapping intramolecular interactions between domains in HMGB1 using a tail-truncation approach. *J. Mol. Biol.*, **374**, 1286–97.

34. Stott,K., Watson,M., Howe,F.S., Grossmann,J.G. and Thomas,J.O. (2010) 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.*, **403**, 706–22.

35. Stott,K., Watson,M., Bostock,M.J., Mortensen,S.A., Travers,A., Grasser,K.D. and Thomas,J.O. (2014) Structural insights into the mechanism of negative regulation of single-box high mobility group proteins by the acidic tail domain. *J. Biol. Chem.*, **289**, 29817–29826.

36. Watson,M., Stott,K., Fischl,H., Cato,L. and Thomas,J.O. (2014) Characterization of the interaction between HMGB1 and H3-a possible means of positioning HMGB1 in chromatin. *Nucleic Acids Res.*, **42**, 848–859.

37. Rowell,J.P., Simpson,K.L., Stott,K., Watson,M. and Thomas,J.O. (2012) HMGB1-Facilitated p53 DNA Binding Occurs via HMG-Box/p53 Transactivation Domain Interaction, Regulated by the Acidic Tail. *Structure*, **20**, 2014–2024.

38. Weiss,M.A. (2001) Floppy SOX: mutual induced fit in HMG (High-Mobility Group) Box-DNA recognition. *Mol. Endocrinol.*, **15**, 353–362.

39. Reeves,R. (2001) Molecular biology of HMGA proteins: hubs of nuclear function. *Gene*, **277**, 63–81.

40. Garabedian,A., Bolufer,A., Leng,F. and Fernandez-Lima,F. (2018) Peptide Sequence Influence on the Conformational Dynamics and DNA binding of the Intrinsically Disordered AT-Hook 3 Peptide. *Sci. Rep.*, **8**, 1–8.

41. Warfield,L., Tuttle,L.M., Pacheco,D., Klevit,R.E. and Hahn,S. (2014) A sequence-specific transcription activator motif and powerful synthetic variants that bind Mediator using a fuzzy protein interface. *Proc. Natl. Acad. Sci. U. S. A.*, **111**, E3506–E3513.

42. Staller,M. V, Holehouse,A.S., Swain-Lenz,D., Das,R.K., Pappu,R. V and Cohen,B.A. (2018) A High-Throughput Mutational Scan of an Intrinsically Disordered Acidic Transcriptional Activation Domain. *Cell Syst.*, **6**, 444-455.e6.

43. Ravarani,C.N.J., Erkina,T.Y., Baets,G. De, Dudman,D.C., Erkine,A.M. and Babu,M.M. (2018) High-throughput discovery of functional disordered regions: investigation of transactivation domains. *Mol. Syst. Biol.*, **14**, 1–14.

44. Chong,S., Dugast-Darzacq,C., Liu,Z., Dong,P., Dailey,G.M., Cattoglio,C., Heckert,A., Banala,S., Lavis,L., Darzacq,X., *et al.* (2018) Imaging dynamic and selective low-complexity domain interactions that control gene transcription. *Science*, **2555**, 1–16.

45. Tompa,P. and Fuxreiter,M. (2008) Fuzzy complexes: polymorphism and structural

disorder in protein-protein interactions. *Trends Biochem. Sci.*, **33**, 2–8.

46. Sharma,R., Raduly,Z., Miskei,M. and Fuxreiter,M. (2015) Fuzzy complexes: Specific binding without complete folding. *FEBS Lett.*, **589**, 2533–2542.

47. Borgia,A., Borgia,M.B., Bugge,K., Kissling,V.M., Heidarsson,P.O., Fernandes,C.B., Sottini,A., Soranno,A., Buholzer,K.J., Nettels,D., *et al.* (2018) Extreme disorder in an ultra-high-affinity protein complex. *Nature*, **555**, 61–66.

48. Turner,A.L., Watson,M., Wilkins,O.G., Cato,L., Travers,A., Thomas,J.O. and Stott,K. (2018) Highly disordered histone H1-DNA model complexes and their condensates. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, 11964–11969.

49. Feng,H., Zhou,B.-R. and Bai,Y. (2018) Binding Affinity and Function of the Extremely Disordered Protein Complex Containing Human Linker Histone H1.0 and Its Chaperone ProTα. *Biochemistry*, **57**, 6645–6648.

50. Kalashnikova,A.A., Winkler,D.D., McBryant,S.J., Henderson,R.K., Herman,J.A., Deluca,J.G., Luger,K., Prenni,J.E. and Hansen,J.C. (2013) Linker histone H1.0 interacts with an extensive network of proteins found in the nucleolus. *Nucleic Acids Res.*, **41**, 4026–4035.

51. Winkler,D.D. and Luger,K. (2011) The histone chaperone FACT: Structural insights and mechanisms for nucleosome reorganization. *J. Biol. Chem.*, **286**, 18369–18374.

52. Thomas,J.O. and Stott,K. (2012) H1 and HMGB1: modulators of chromatin structure. *Biochem. Soc. Trans.*, **40**, 341–6.

53. Cato,L., Stott,K., Watson,M. and Thomas,J.O. (2008) The interaction of HMGB1 and linker histones occurs through their acidic and basic tails. *J. Mol. Biol.*, **384**, 1262–72.

54. Kalashnikova,A.A., Rogge,R.A. and Hansen,J.C. (2016) Linker histone H1 and protein–protein interactions. *Biochim. Biophys. Acta*, **1859**, 455–461.

55. Fyodorov,D. V, Zhou,B.R., Skoultchi,A.I. and Bai,Y. (2018) Emerging roles of linker histones in regulating chromatin structure and function. *Nat. Rev. Mol. Cell Biol.*, **19**, 192–206.

56. Mascotti,D.P. and Lohman,T.M. (1990) Thermodynamic extent of counterion release upon binding oligolysines to single-stranded nucleic acids. *Proc. Natl. Acad. Sci. U. S. A.*, **87**, 3142–3146.

57. Iwahara,J., Esadze,A. and Zandarashvili,L. (2015) Physicochemical Properties of Ion Pairs of Biological Macromolecules. *Biomolecules*, **5**, 2435–2463.

58. George,E.M. and Brown,D.T. (2010) Prothymosin α is a component of a linker histone chaperone. *FEBS Lett.*, **584**, 2833–2836.

59. Catez,F., Ueda,T. and Bustin,M. (2006) Determinants of histone H1 mobility and chromatin binding in living cells. *Nat. Struct. Mol. Biol.*, **13**, 305–310.

60. White,A.E., Hieb,A.R. and Luger,K. (2016) A quantitative investigation of linker histone interactions with nucleosomes and chromatin. *Sci. Rep.*, **6**, 19122.

61. Contreras,A., Hale,T.K., Stenoien,D.L., Rosen,J.M., Mancini,M. a and Herrera,R.E. (2003) The dynamic mobility of histone H1 is regulated by cyclin/CDK phosphorylation. *Mol. Cell. Biol.*, **23**, 8626–8636.

62. Mao,Y.S., Zhang,B. and Spector,D.L. (2011) Biogenesis and function of nuclear bodies. *Trends Genet.*, **27**, 295–306.

63. Hyman,A.A., Weber,C.A. and Jülicher,F. (2014) Liquid-Liquid Phase Separation in Biology. *Annu. Rev. Cell Dev. Biol.*, **30**, 39–58.

64. Zhu,L. and Brangwynne,C.P. (2015) Nuclear bodies: the emerging biophysics of nucleoplasmic phases. *Curr. Opin. Cell Biol.*, **34**, 23–30.

65. Shin,Y., Chang,Y., Lee,D.S.W., Berry,J., Sanders,D.W., Ronceray,P., Wingreen,N.S., Haataja,M. and Brangwynne,C.P. (2018) Liquid nuclear condensates mechanically sense and restructure the genome. *Cell*, **175**, 1481–1491.

66. Stroberg,W. and Schnell,S. (2018) Do cellular condensates accelerate biochemical reactions? Lessons from microdroplet chemistry. *Biophys. J.*, **115**, 3–8.

67. Nott,T.J., Craggs,T.D. and Baldwin,A.J. (2016) Membraneless organelles can melt nucleic acid duplexes and act as biomolecular filters. *Nat. Chem.*, **8**, 569–575.

68. Aguzzi,A. and Altmeyer,M. (2016) Phase Separation : Linking Cellular Compartmentalization to Disease. *Trends Cell Biol.*, **26**, 547–558.

69. Shin,Y. and Brangwynne,C.P. (2017) Liquid phase condensation in cell physiology and disease. *Science*, **357**.

70. Maeshima,K., Ide,S., Hibino,K. and Sasai,M. (2016) Liquid-like behavior of chromatin. *Curr. Opin. Genet. Dev.*, **37**, 36–45.

71. Nozaki,T., Imai,R., Tanbo,M., Nagashima,R., Tamura,S., Tani,T., Joti,Y., Tomita,M., Hibino,K., Kanemaki,M.T., *et al.* (2017) Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. *Mol. Cell*, **67**, 282–293.

72. Ou,H.D., Phan,S., Deerinck,T.J., Thor,A., Ellisman,M.H. and O'Shea,C.C. (2017)

ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science*, **357**.

73. Zhou,B.R., Jiang,J., Ghirlando,R., Norouzi,D., Sathish Yadav,K.N., Feng,H., Wang,R., Zhang,P., Zhurkin,V. and Bai,Y. (2018) Revisit of Reconstituted 30-nm Nucleosome Arrays Reveals an Ensemble of Dynamic Structures. *J. Mol. Biol.*, **430**, 3093–3110.

74. Olins,D.E. and Olins,A.L. (2018) Epichromatin and chromomeres: A 'fuzzy' perspective. *Open Biol.*, **8**.

75. Gibbs,E.B. and Kriwacki,R.W. (2018) Linker histones as liquid-like glue for chromatin. *Proc. Natl. Acad. Sci. U. S. A.*, **115**, 11868–11870.

76. Maniatis,T., Venable,J.H. and Lerman,L.S. (1974) The Structure of Psi DNA. *J. Mol. Biol.*, **84**, 37.

77. Maestre,M.F. and Reich,C. (1980) Contribution of Light Scattering to the Circular Dichroism of Deoxyribonucleic Acid Films, Deoxyribonucleic Acid-Polylysine Complexes, and Deoxyribonucleic Acid Particles in Ethanolic Buffers. *Biochemistry*, **19**, 5214–5223.

78. Wang,J., Jia,S.T. and Jia,S. (2016) New Insights into the Regulation of Heterochromatin. *Trends Genet.*, **32**, 284–294.

79. Strom,A.R., Emelyanov,A. V., Mir,M., Fyodorov,D. V., Darzacq,X. and Karpen,G.H. (2017) Phase separation drives heterochromatin domain formation. *Nature*, **547**, 241–245.

80. Larson,A.G., Elnatan,D., Keenen,M.M., Trnka,M.J., Johnston,J.B., Burlingame,A.L., Agard,D.A., Redding,S. and Narlikar,G.J. (2017) Liquid droplet formation by HP1α suggests a role for phase separation in heterochromatin. *Nature*, **547**, 236–240.
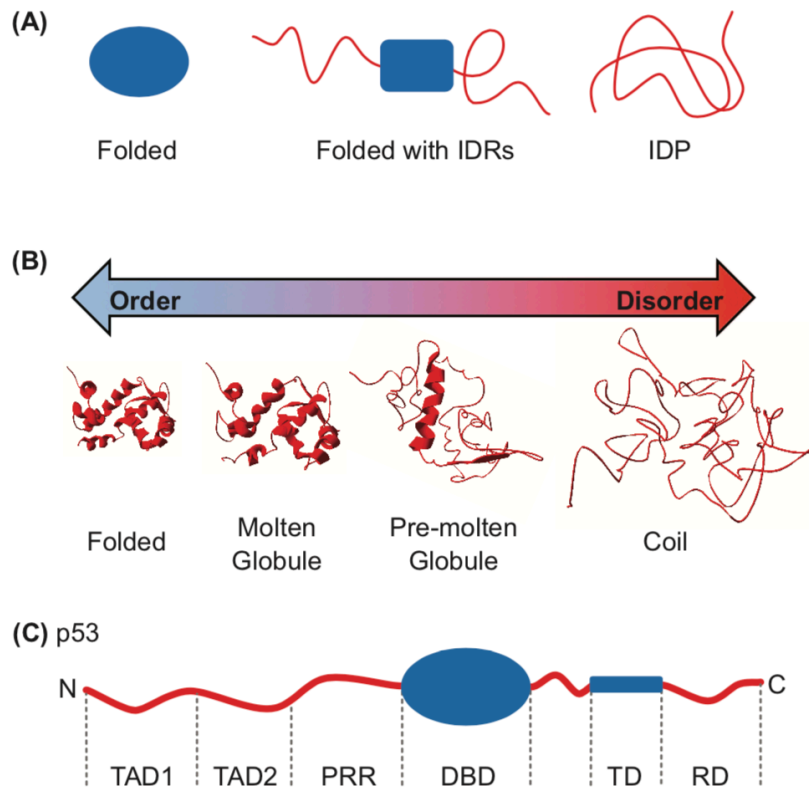
Figure 1: The continuum of intrinsic disorder.

(A) Proteins can be described either as folded, as having a modular architecture with both folded (blue) and intrinsically-disordered (red) regions (IDRs), or as true IDPs in which the whole polypeptide contains little or no stable structure. (B) The structures populated by IDP(R)s sit on a continuum between stably folded (left) and random coil (right), ranging from collapsed molten globule-like structures, in which a significant amount of secondary structure is present through to expanded coil-like structures with little or no secondary structure. Protein schematics were generated from PDB: 4ULW using Swiss PDB viewer (8). (C) The modular architecture of p53, in which the DNA-binding domain (DBD) and tetramerisation domain (TD) are stably folded in the free protein, while the transactivation domains (TADs), proline-rich region (PRR) and regulatory domain (RD) are all IDRs.
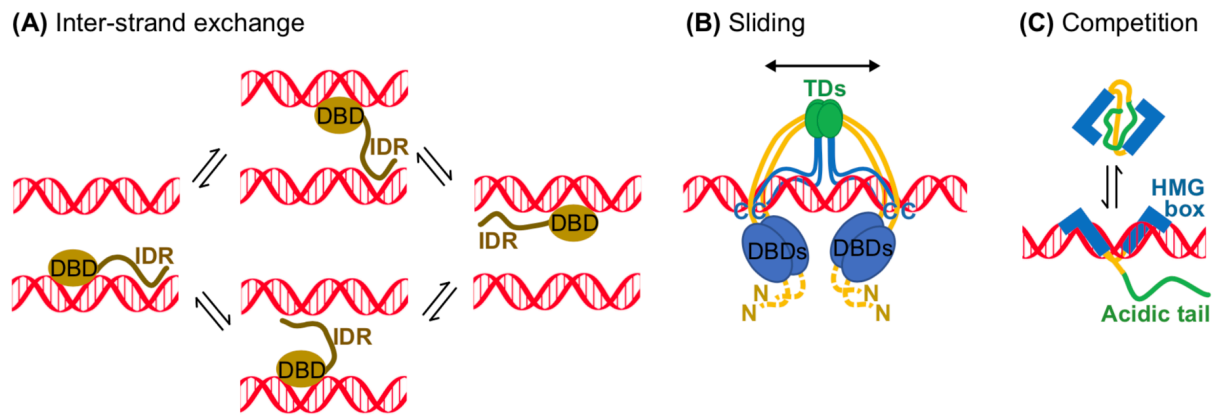
Figure 2: Role of IDRs in DNA binding.

(A) "Monkey bar" inter-strand exchange mechanism of DNA binding site search (30). Either the DNA-binding domain (DBD) or the flanking disordered region (IDR) can dissociate from one strand and bind to another. Subsequent dissociation and transfer of the remaining region results in exchange of the protein to a new strand without ever becoming fully dissociated from DNA. (B) Model for p53 tetramer scanning (32) (TD = tetramerisation domain and DBD = DNA-binding domain). The C-terminal IDRs (blue) bind DNA in a non-sequence specific manner and mediate fast sliding of p53 along DNA while the core domains frequently sample the DNA by repeated association and dissociation. (C) Competitive inhibition of DNA binding. The intrinsically disordered acidic tails of HMGB proteins (green) make extensive contacts with the concave DNA-binding surface of the HMG boxes (blue), negatively regulating binding of the protein to low-affinity DNA sites (35).
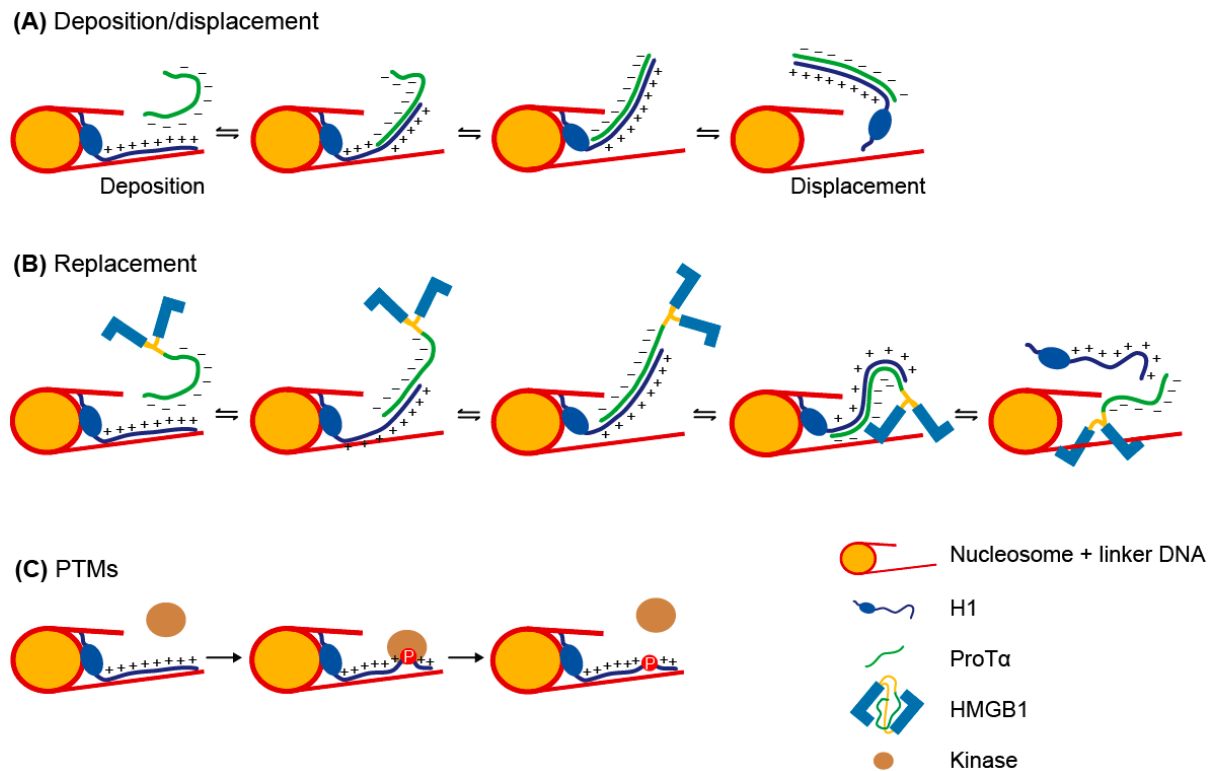
Figure 3: Potential role of "fuzzy interactions" in the deposition, displacement and replacement of H1 in chromatin.

(A) An acidic IDP such as ProTα (green), can either displace (left to right), or deposit (right to left), H1 (blue) in chromatin. It is proposed here that "fuzzy interactions" between the H1 C-terminal tail and ProTα or DNA (red) progressively unpeel interactions allowing rapid exchange despite the high affinities of the H1 for both the nucleosome and ProTα. (B) A similar mechanism could be used by an architectural protein such as HMGB1, to replace H1 by gradually unpeeling the H1-tail/DNA interactions, concomitantly decreasing the affinity of H1 for the nucleosome and increasing the affinity of the HMG box(es) for linker DNA (left to right). Conversely, in the reverse process (right to left) H1 could displace HMGB1 from the linker DNA. (C) "Fuzzy interactions" facilitate post-translational modifications. The ability of these "fuzzy interactions" to be partially unpeeled without complete dissociation might allow enzymes (brown) to modify the H1 tail while it is tightly bound to chromatin.
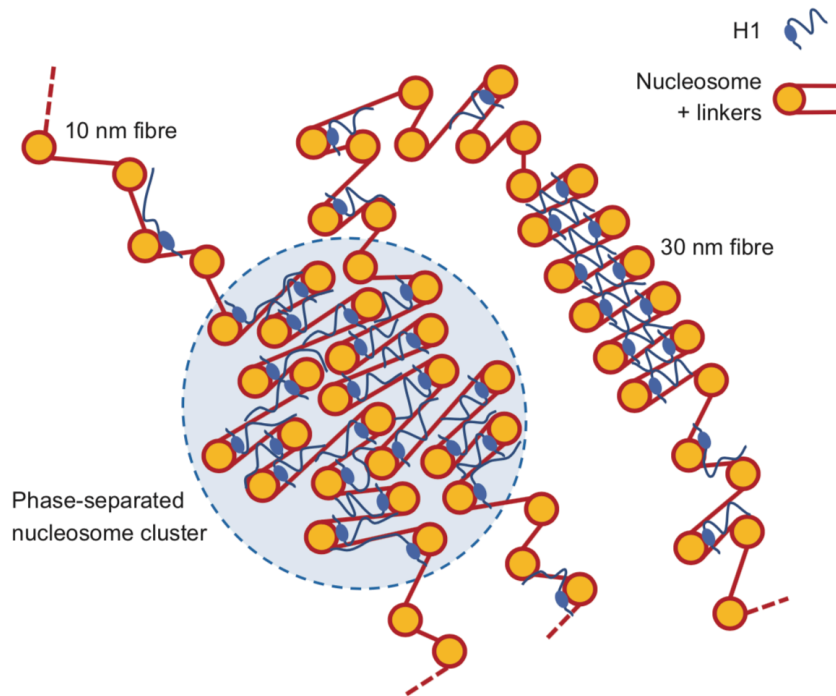
Figure 4: The role of H1 in condensing chromatin into different states.

The highly basic H1 C-terminal tail binds predominantly to the linker DNA between nucleosomes in the 10 nm "beads-on-a-string" fibre. This neutralises the charge on the DNA and allows the tight packing of chromatin into the 30 nm fibre in which the linker DNA and H1 are at the core of the fibre. Alternatively, the observation that the H1 tail phase separates upon interaction with DNA (48) has led to the suggestion that the clusters of nucleosomes observed by microscopy might be brought together and partitioned by H1 into a phase separated droplet or microdomain forming a more dynamic yet still highly condensed chromatin domain (75).