



Mini-Review

Picking a nucleosome lock: Sequence- and structure-specific recognition of the nucleosome

MATTHEW M MAKOWSKI^{1,2}, GUILLAUME GAULLIER^{1,2}  and
KAROLIN LUGER^{1,2*} 

¹Department of Biochemistry, University of Colorado Boulder, Boulder, CO, USA

²Howard Hughes Medical Institute, Chevy Chase, MD, USA

*Corresponding author (Email, karolin.luger@colorado.edu)

The nucleosome presents a formidable barrier to DNA-templated transcription by the RNA polymerase II machinery. Overcoming this transcriptional barrier in a locus-specific manner requires sequence-specific recognition of nucleosomal DNA by ‘pioneer’ transcription factors (TFs). Cell fate decisions, in turn, depend on the coordinated action of pioneer TFs at cell lineage-specific gene regulatory elements. Although it is already appreciated that pioneer factors play a critical role in cell differentiation, our understanding of the structural and biochemical mechanisms by which they act is still rapidly expanding. Recent research has revealed novel insight into modes of nucleosome-TF binding and uncovered kinetic principles by which nucleosomal DNA compaction affects both TF binding and residence time. Here, we review progress and argue that these structural and kinetic studies suggest new models of gene regulation by pioneer TFs.

Keywords. Nucleosome; pioneer factors; transcription factors; transcriptional regulation

The same essential genome is replicated faithfully in nearly every one of hundreds of diverse cell types (and in ~37 trillion cells) in the human body. Differentiation into heterogeneous cell phenotypes from totipotent progenitors requires differential gene expression of the same underlying genetic information (Levine and Davidson 2005; Lee and Young 2013). During differentiation, genes might be turned ‘on’ during cell differentiation, as with lineage restricted genes, or ‘off’, for example pluripotency genes or genes specific to other lineages. Cell type classification often uses these global differences in gene expression patterns to categorize cell types based on the expression of characteristic lineage marking genes or gene sets (Regev *et al.* 2017). Many molecular mechanisms regulate differential gene expression. The nucleosome is a tertiary complex composed of 147 base pairs of DNA wrapped

around an octameric complex of histone proteins. This DNA configuration effectively compacts and distorts the nucleosomal DNA, occludes much of the histone-facing DNA surface over its entire 147 bp, and presents a serious barrier to effective transcription (Luger *et al.* 1997; Kujirai *et al.* 2018; Farnung *et al.* 2018; Ehara *et al.* 2019; Zhou *et al.* 2019). Thus, one key mechanism of gene expression regulation is the compaction of transcriptional and gene regulatory elements by nucleosomes.

Gene regulation via nucleosome occlusion of a regulatory or transcriptional element, in the absence of other regulatory factors, depends only partially on the sequence of the element being occluded (Struhl and Segal 2013). Additional factors, notably the nucleosome positioning enzymes referred to as chromatin remodelers, actively position nucleosomes via complicated ATP-dependent mechanisms (Becker and Hörz 2002; Clapier *et al.* 2017). Chromatin remodelers themselves are only sometimes weakly specific for certain DNA sequence elements (Rippe *et al.* 2007).

This article is part of the Topical Collection: Chromatin Biology and Epigenetics.

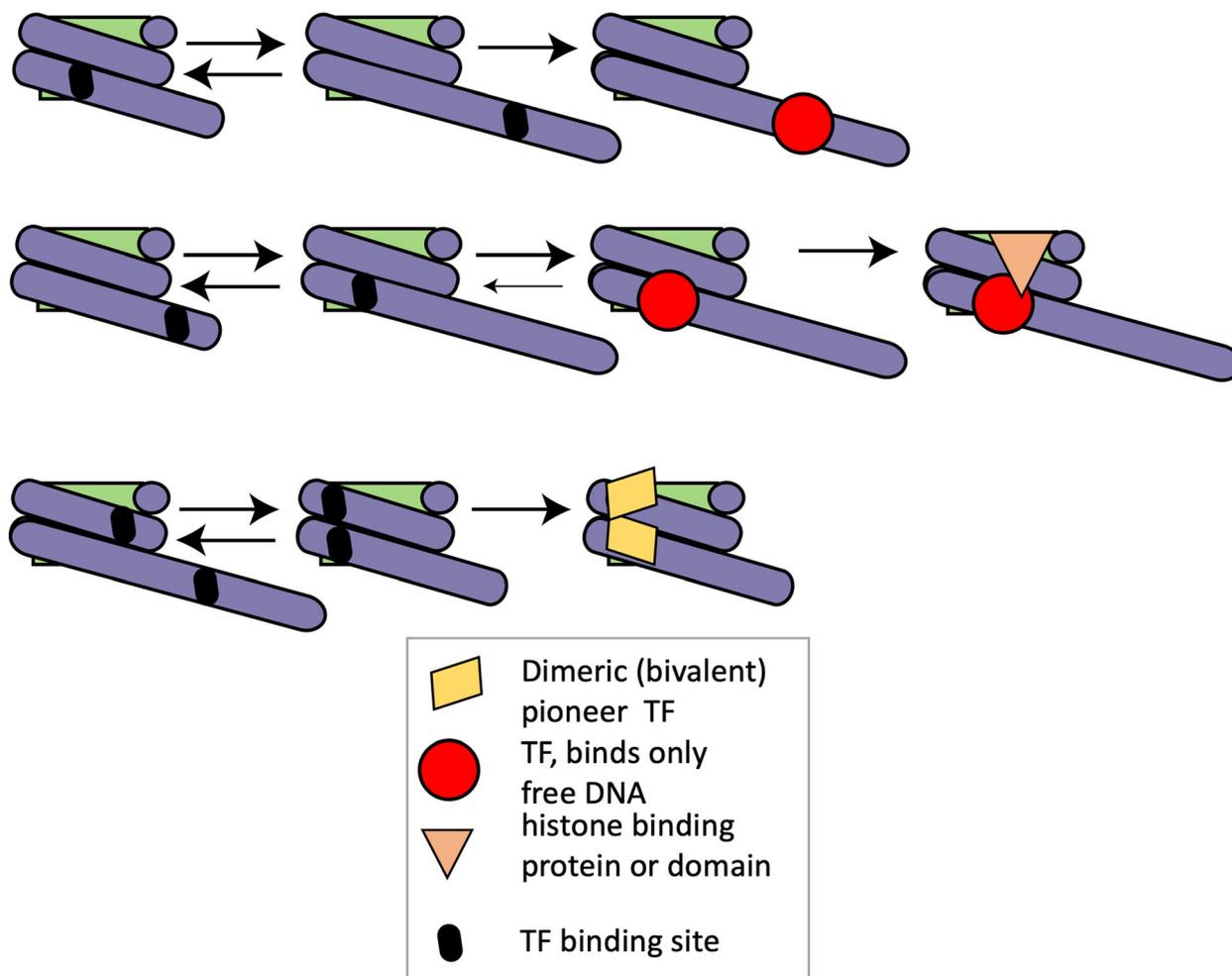


Figure 1. Nucleosome positioning by TF binding can occur through different mechanisms. Nucleosomes are repositioned either spontaneous or mediated by ATP-dependent chromatin remodelers; specific positions are stabilized by TF binding. Size of arrows denotes direction of equilibrium.

Instead, many sequence-specific transcription factors (TFs) recruit chromatin remodelers or the histone modification machinery and thus either directly or indirectly drive locus-specific chromatin remodeling *in vivo* (Lambert *et al.* 2018). Likewise, TF binding also has the potential to direct nucleosome position by restricting their mobility (figure 1).

TFs modulate nucleosome and chromatin structure using several different mechanisms (figure 2). First, TFs can directly recruit chromatin remodelers to place, position, or remove nucleosomes at a particular DNA element. Second, some TFs only recognize nucleosomal DNA when nucleosome ‘breathing’ occurs, that is when the DNA is partially and temporarily unwrapped from the nucleosome surface. This concept was pioneered by the late Jon Widom (Li and Widom 2004; Li *et al.* 2005), and it was later shown that histone post-translational modifications facilitate DNA breathing

(reviewed in Bowman and Poirier 2015). TF binding facilitates further nucleosome unwrapping by promoting the binding of additional TFs, and/or in coordination with chromatin remodelers (Swinstead *et al.* 2016b; Fierz and Poirier 2019). Some TFs can bind their cognate motifs on fully compacted nucleosomal DNA and initiate ATP-independent DNA unwrapping or even histone eviction (Perlmann and Wrangé 1988; Cirillo *et al.* 2002). However, outcomes in which TF binding stabilizes nucleosomes are also possible.

Most TFs are unable to bind preferred motifs if those motifs are embedded within nucleosomal DNA. In fact, early models proposed a competitive mechanism whereby TF binding precluded nucleosome formation in transcriptionally active regions (Workman 2006). On the other hand, TFs that have the ability to recognize their preferred motifs efficiently in the context of fully compacted nucleosomal DNA are referred to as

for binding while rejecting other sites (figure 2, top row). The role of histone post-translational modifications in either directly or indirectly affecting TF binding (pioneer or not) is also under investigation (figure 3).

High-resolution structural details such as the rotational position of pioneer factor motifs around the nucleosome can substantially impact pioneer factor binding (Sekiya *et al.* 2009; Sahu *et al.* 2010; Cui and Zhurkin 2014; Soufi *et al.* 2015). Similarly, the location of the motif with respect to the histone octamer core affects pioneer TF binding. For example, FoxA1 is a canonical pioneer

factor that prefers near-dyad binding (McPherson *et al.* 1993). In contrast, the widely studied TF and tumor suppressor p53 binds near the nucleosomal DNA entry and exit sites (Yu and Buck 2018). As yet another variation on this theme, the Yamanaka factor Sox2 does not exhibit a preference for a precise location of its site around the nucleosome, but rather whether the site faces inward or outward and whether or not the abundant nuclear enzyme poly(ADP-ribose) polymerase 1 (PARP1) cooperatively binds with Sox2 (Liu and Kraus 2017). In addition to the motif itself, DNA shape features and motif density can also play a role in guiding

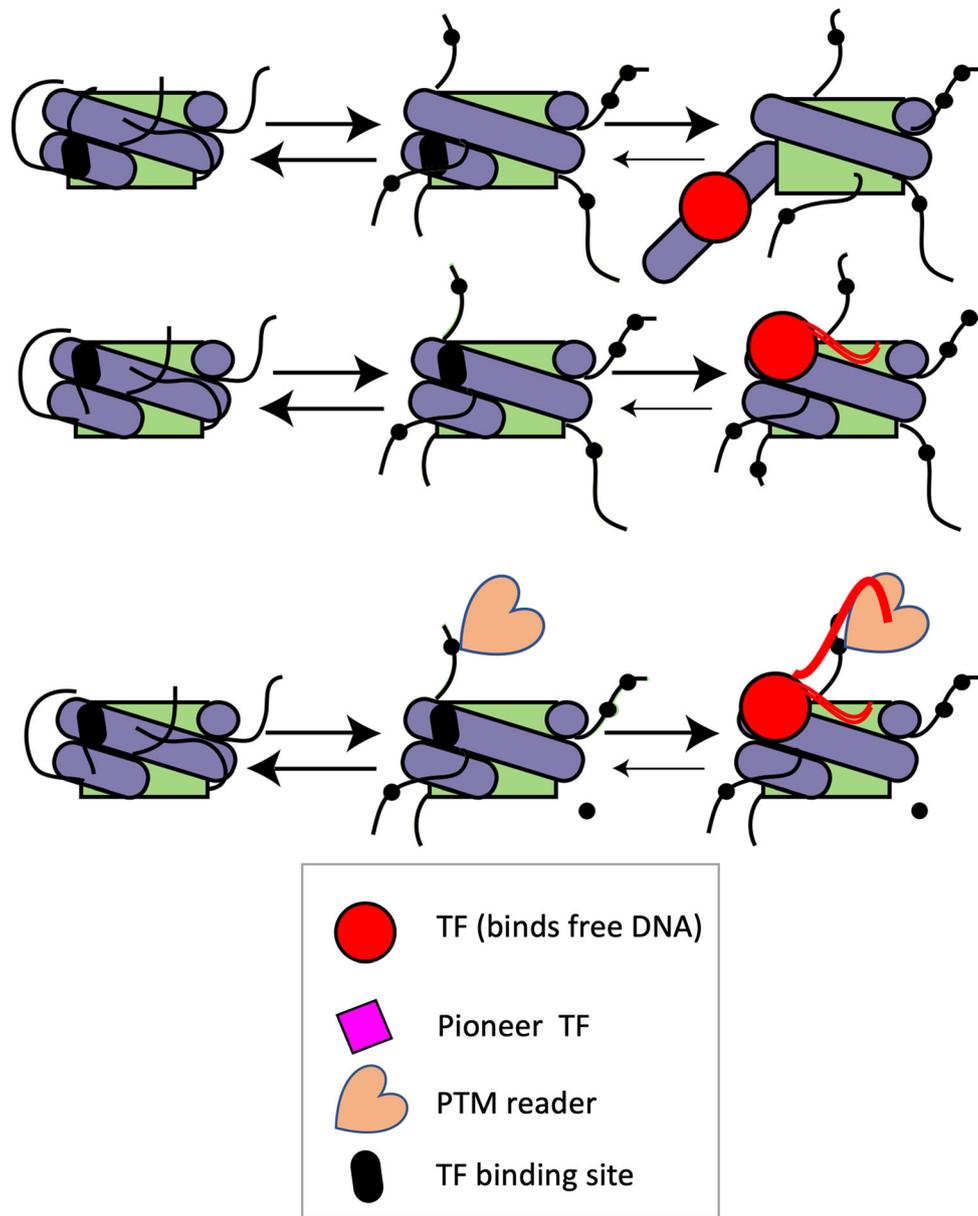


Figure 3. Histone post-translational modifications can regulate DNA accessibility. PTMs on histone tails (black lines) can promote TF binding either by increasing DNA breathing, increasing accessibility of nucleosomal DNA, or recruitment of PTM readers that in turn recruit TF.

nucleosome recognition by pioneer TFs (Ghosh *et al.* 2018). Furthermore, non-specific TF binding in the absence of a preferred motif can still induce strong pioneer factor–nucleosome interactions that are only a few fold weaker than in the presence of a strong motif (Sekiya *et al.* 2009; Yu and Buck 2018). With Satb1, a multimeric ‘genome organizer’ transcription factor that binds to nucleosomal DNA, a sequence-specific DNA binding domain confers sequence specificity but is not required for high-affinity nucleosome binding interactions (Ghosh *et al.* 2018). Some pioneer factors also possess direct histone interacting domains that regulate nucleosome recognition, for example the H3-H4 interacting C-terminal domain of FoxA1 (figure 1), although it is still unclear how common these domains are in sequence-specific TFs (Cirillo *et al.* 2002). There are currently no high-resolution cryo-EM or crystallographic structures of pioneer TFs bound to nucleosomes (reviewed in Zhou *et al.* 2019) despite one documented attempt (Takizawa *et al.* 2018), but structural studies fueled by the cryo-EM ‘resolution revolution’ will likely soon offer exciting insights into the atomistic details of TF–nucleosome interactions.

Apart from conventional biochemical and structural approaches, a recent *tour de force* study utilized a SELEX experimental design with high-throughput sequencing to identify nucleosome binding patterns for hundreds of transcription factors or DNA binding domains (Zhu *et al.* 2018). In addition to entry/exit site and dyad binding, this study uncovered a number of novel TF–nucleosome binding modes facilitated by the unique topography of the nucleosome. Of particular note was a cross-gyre or supergroove (Edayathuman-galam *et al.* 2004) spanning binding mode. In this binding mode, two T-box family DNA binding domains (T and Tbx2) bound as a presumptive dimer to motifs separated by approximately 80 base pairs in sequence but spatially close to each other within the major supergroove of the nucleosome (figure 1). A number of T-box family DNA binding domains dimerize on DNA (Coll *et al.* 2002; El Omari *et al.* 2012; Newman *et al.* 2017) or in solution (Liu *et al.* 2016; Pradhan *et al.* 2016). Additionally, T-box supergroove binding reportedly stabilized the wrapped nucleosome (Zhu *et al.* 2018), as was reported much earlier for a bivalent small molecule (Edayathuman-galam *et al.* 2004). Therefore, it seems reasonable to hypothesize that some T-box DNA binding domains can dimerize across the nucleosomal gyres and, by stably homodimerizing, prevent nucleosome dissociation (figure 2, top row). This hypothesis envisions a type of ‘reverse’ pioneer factor, a heretofore unreported

means of repressing gene expression through stabilizing the nucleosome. How significant and common this mechanism of gene expression regulation is *in vivo* merits further study.

In addition to these recent developments in our understanding of the structural characteristics of TF and pioneer factor nucleosome binding, a recent kinetic study offered equally important insights into the molecular mechanisms by which pioneer factors act (Donovan *et al.* 2019). Slow off-rates and long residence times, as well as low mobility on chromatin (Sekiya *et al.* 2009), are proposed kinetic properties of pioneer factors, although there are some conflicting results in the case of FoxA1 (Sekiya *et al.* 2009; Swinstead *et al.* 2016a). These properties are justified under the assumption that pioneering action – unwinding or even evicting histones from a stable nucleosome – is expected to be a slow kinetic step. However, by definition, the binding affinity of a pioneer factor to a nucleosome should be the same, or nearly the same, as to free linear DNA (Zaret and Carroll 2011). Therefore, slower off-rates on chromatin and nucleosomes would necessarily be compensated by slower on-rates to maintain a similar equilibrium binding constant. And, for two yeast pioneer factors (Reb1 and Cbf1), a recent report demonstrates that this mechanism, referred to by the authors as dissociation rate compensation, is correct (Donovan *et al.* 2019). For Reb1, its slower dissociation rate on nucleosomes correlates with slow exchange in FRAP experiments *in vivo*. It remains unclear if this dissociation rate compensation mechanism is unique to a few select families or types of TFs, including Reb1 and Cbf1, or if it is instead common among most TFs with pioneer factor activity. In contrast, Rap1, another yeast pioneer TF, has shorter residence times on nucleosome arrays compared to linear DNA but still unwraps the chromatin fiber by inhibiting inter-nucleosome contacts and recruiting the RSC chromatin remodeler (Mivelaz *et al.* 2019).

What does this accumulated structural and kinetic information teach us about pioneer factor activity? In general, it seems increasingly clear that most TFs are repelled, often strongly, by nucleosomes, while the TFs that do bind nucleosomes tend to unwrap them (Zhu *et al.* 2018). One kinetic mechanism for nucleosome unwinding is dissociation rate compensation (Donovan *et al.* 2019); yet other kinetic mechanisms such as the one employed by Rap1 clearly exist (Mivelaz *et al.* 2019). In a broader sense, how do different structural modes of sequence-specific nucleosome recognition correlate with kinetic parameters of binding to and unwrapping of nucleosomal DNA? Does more severe dissociation rate compensation occur upon binding at some nucleosomal locations compared to others? How

do cooperative interactions among TFs (Adams and Workman 1995; Polach and Widom 1996) change the kinetic landscape? On the other hand, only a few TFs bind to and stabilize nucleosomes (Zhu et al. 2018), and it is uncertain how this nucleosome-stabilizing binding mode will relate to TF binding on free DNA. Do nucleosome-stabilizing TFs utilize a dissociation rate compensation mechanism? It is logical to suppose that nucleosome stabilization by a gyre-spanning dimer (or even a bivalent monomer) would co-occur with slower off-rates to maximize the duration of the stabilizing effect, yet this remains to be demonstrated.

In sum, the nucleosome represents an important means of regulating transcription in cells. But beyond acting as a simple transcriptional barrier, recent findings indicate that we should consider the nucleosome as a unique structural template that is recognized in a specific manner by transcriptional regulators. In the context of the nucleosome, DNA sequences can adopt new regulatory features that are not present in the strictly linear form, and the properties of nucleosomal DNA must be understood in terms of structure as well as in a time-dependent manner. Understanding the structural and mechanistic details of these nucleosome-specific TF functions will, in turn, inform a more sophisticated and nuanced picture of gene regulation and chromatin structure in the nucleus that incorporates activities that impinge on nucleosome structure and stability, such as ATP-dependent chromatin remodelers and the post-translational modification of histones.

Acknowledgements

This work was supported by the Howard Hughes Medical Institute and NIH R01CA218255.

References

- Adams CC and Workman JL 1995 Binding of disparate transcriptional activators to nucleosomal DNA is inherently cooperative. *Mol. Cell. Biol.* **15** 1405–1421
- Becker PB and Hörz W 2002 ATP-Dependent nucleosome remodeling. *Annu. Rev. Biochem.* **71** 247–273
- Bowman GD and Poirier MG 2015 Post-translational modifications of histones that influence nucleosome dynamics. *Chem. Rev.* **115** 2274–2295
- Cirillo LA, Lin FR, Cuesta I, Friedman D, Jarnik M and Zaret KS 2002 Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* **9** 279–289
- Clapier CR, Iwasa J, Cairns BR and Peterson CL 2017 Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18** 407–422
- Coll M, Seidman JG and Müller CW 2002 Structure of the DNA-bound T-Box domain of human TBX3, a transcription factor responsible for Ulnar-Mammary syndrome. *Structure* **10** 343–356
- Cui F and Zhurkin VB 2014 Rotational positioning of nucleosomes facilitates selective binding of p53 to response elements associated with cell cycle arrest. *Nucleic Acids Res.* **42** 836–847
- Donovan BT, Chen H, Jipa C, Bai L and Poirier MG 2019 Dissociation rate compensation mechanism for budding yeast pioneer transcription factors. *eLife* **8** e43008
- Edayathumangalam RS, Weyermann P, Gottesfeld JM, Dervan PB and Luger K 2004 Molecular recognition of the nucleosomal ‘super groove’. *Proc. Natl. Acad. Sci. USA* **101** 6864–6869
- Ehara H, Kujirai T, Fujino Y, Shirouzu M, Kurumizaka H and Sekine S 2019 Structural insight into nucleosome transcription by RNA polymerase II with elongation factors. *Science* **363** 744–747
- El Omari K, Mesmaeker JD, Karia D, Ginn H, Bhattacharya S and Mancini EJ 2012 Structure of the DNA-bound T-box domain of human TBX1, a transcription factor associated with the DiGeorge syndrome. *Proteins Struct. Funct. Bioinforma.* **80** 655–660
- Farnung L, Vos SM and Cramer P 2018 Structure of transcribing RNA polymerase II-nucleosome complex. *Nat. Commun.* **9** 5432
- Fierz B and Poirier MG 2019 Biophysics of Chromatin Dynamics. *Annu. Rev. Biophys.* **48** null
- Ghosh RP, Shi Q, Yang L, Reddick MP, Nikitina T, Zhurkin VB, Fordyce P, Stasevich TJ, Chang HY, Greenleaf WJ and Liphardt JT 2018 Satb1 integrates DNA sequence, shape, motif density and torsional stress to differentially bind targets in nucleosome-dense regions. *bioRxiv* 450262
- Iwafuchi-Doi M 2019 The mechanistic basis for chromatin regulation by pioneer transcription factors. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **11** e1427
- Iwafuchi-Doi M and Zaret KS 2014 Pioneer transcription factors in cell reprogramming. *Genes Dev.* **28** 2679–2692
- Jolma A, Yin Y, Nitta KR, Dave K, Popov A, Taipale M, Enge M, Kivioja T, Morgunova E and Taipale J 2015 DNA-dependent formation of transcription factor pairs alters their binding specificity. *Nature* **527** 384–388
- Kujirai T, Ehara H, Fujino Y, Shirouzu M, Sekine S and Kurumizaka H 2018 Structural basis of the nucleosome transition during RNA polymerase II passage. *Science* **362** 595–598
- Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR and Weirauch MT 2018 The human transcription factors. *Cell* **172** 650–665

- Lee TI and Young RA 2013 Transcriptional regulation and its misregulation in disease. *Cell* **152** 1237–1251
- Levine M and Davidson EH 2005 Gene regulatory networks for development. *Proc. Natl. Acad. Sci.* **102** 4936–4942
- Li G, Levitus M, Bustamante C and Widom J 2005 Rapid spontaneous accessibility of nucleosomal DNA. *Nat. Struct. Mol. Biol.* **12** 46–53
- Li G and Widom J 2004 Nucleosomes facilitate their own invasion. *Nat. Struct. Mol. Biol.* **11** 763–769
- Liu CF, Brandt GS, Hoang QQ, Naumova N, Lazarevic V, Hwang ES, Dekker J, Glimcher LH, Ringe D and Petsko GA 2016 Crystal structure of the DNA binding domain of the transcription factor T-bet suggests simultaneous recognition of distant genome sites. *Proc. Natl. Acad. Sci.* **113** E6572–E6581
- Liu Z and Kraus WL 2017 Catalytic-independent functions of PARP-1 determine Sox2 pioneer activity at intractable genomic loci. *Mol. Cell* **65** 589–603
- Luger K, Mäder AW, Richmond RK, Sargent DF and Richmond TJ 1997 Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389** 251–260
- Mayran A, Sochodolsky K, Khetchoumian K, Harris J, Gauthier Y, Bemmo A, Balsalobre A and Drouin J 2019 Pioneer and nonpioneer cooperation drives lineage specific chromatin opening. *bioRxiv* 472647
- McPherson CE, Shim E-Y, Friedman DS and Zaret KS 1993 An active tissue-specific enhancer and bound transcription factors existing in a precisely positioned nucleosomal array. *Cell* **75** 387–398
- Mirny LA 2010 Nucleosome-mediated cooperativity between transcription factors. *Proc. Natl. Acad. Sci.* **107** 22534–22539
- Mivelaz M, Cao A-M, Kubik S, Zencir S, Hovius R, Boichenko I, Stachowicz AM, Kurat CF, Shore D and Fierz B 2019 The mechanistic basis for chromatin invasion and remodeling by the yeast pioneer transcription factor Rap1. *bioRxiv* 541284
- Newman JA, Gavard AE, Krojer T, Shrestha L, Burgess-Brown NA, von Delft F, Arrowsmith CH, Edwards A, Bountra C and Gileadi O 2017 Crystal structure of human Brachyury (T) in complex with DNA. (<https://www.rcsb.org/structure/6F58>)
- Perlmann T and Wrangé O 1988 Specific glucocorticoid receptor binding to DNA reconstituted in a nucleosome. *EMBO J.* **7** 3073–3079
- Polach KJ and Widom J 1996 A Model for the Cooperative Binding of Eukaryotic Regulatory Proteins to Nucleosomal Target Sites. *J. Mol. Biol.* **258** 800–812
- Pradhan L, Gopal S, Li S, Ashur S, Suryanarayanan S, Kasahara H and Nam H-J 2016 Intermolecular interactions of cardiac transcription factors NKX2.5 and TBX5. *Biochemistry* **55** 1702–1710
- Regev A, Teichmann SA, Lander ES, Amit I, Benoist C, Birney E, Bodenmiller B, Campbell P, *et al.* 2017 The human cell atlas. *eLife* **6** e27041
- Reiter F, Wienerroither S and Stark A 2017 Combinatorial function of transcription factors and cofactors. *Curr. Opin. Genet. Dev.* **43** 73–81
- Rippe K, Schrader A, Riede P, Strohner R, Lehmann E and Längst G 2007 DNA sequence- and conformation-directed positioning of nucleosomes by chromatin-remodeling complexes. *Proc. Natl. Acad. Sci.* **104** 15635–15640
- Sahu G, Wang D, Chen CB, Zhurkin VB, Harrington RE, Appella E, Hager GL and Nagaich AK 2010 p53 binding to nucleosomal DNA depends on the rotational positioning of DNA response element. *J. Biol. Chem.* **285** 1321–1332
- Sekiya T, Muthurajan UM, Luger K, Tulin AV and Zaret KS 2009 Nucleosome-binding affinity as a primary determinant of the nuclear mobility of the pioneer transcription factor FoxA. *Genes Dev.* **23** 804–809
- Soufi A, Garcia MF, Jaroszewicz A, Osman N, Pellegrini M and Zaret KS 2015 Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161** 555–568
- Struhl K and Segal E 2013 Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* **20** 267–273
- Swinstead EE, Miranda TB, Paakinaho V, Baek S, Goldstein I, Hawkins M, Karpova TS, Ball D, Mazza D, Lavis LD, Grimm JB, Morisaki T, Grøntved L, Presman DM and Hager GL 2016a Steroid receptors reprogram FoxA1 occupancy through dynamic chromatin transitions. *Cell* **165** 593–605
- Swinstead EE, Paakinaho V, Presman DM and Hager GL 2016b Pioneer factors and ATP-dependent chromatin remodeling factors interact dynamically A new perspective. *BioEssays* **38** 1150–1157
- Takizawa Y, Tanaka H, Machida S, Koyama M, Maehara K, Ohkawa Y, Wade PA, Wolf M and Kurumizaka H 2018 Cryo-EM structure of the nucleosome containing the ALB1 enhancer DNA sequence. *Open Biol.* **8** 170255
- Workman JL 2006 Nucleosome displacement in transcription. *Genes Dev.* **20** 2009–2017
- Yu X and Buck MJ 2018 Defining TP53 pioneering capabilities with competitive nucleosome binding assays. *Genome Res.* **29** 107–115
- Zaret KS 2018 Pioneering the chromatin landscape. *Nat. Genet.* **50** 167–169
- Zaret KS and Carroll JS 2011 Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* **25** 2227–2241
- Zhou K, Gaullier G and Luger K 2019 Nucleosome structure and dynamics are coming of age. *Nat. Struct. Mol. Biol.* **26** 3–13
- Zhu F, Farnung L, Kaasinen E, Sahu B, Yin Y, Wei B, Dodonova SO, Nitta KR, Morgunova E, Taipale M, Cramer P and Taipale J 2018 The interaction landscape between transcription factors and the nucleosome. *Nature* **562** 76–81