Mechanisms of Interplay between Transcription Factors and the 3D Genome

Seungsoo Kim1 and Jay Shendure1,2,3,*
1Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA
2Howard Hughes Medical Institute, Seattle, WA 98195, USA
3Brotman Baty Institute for Precision Medicine, Seattle, WA 98195, USA
*Correspondence: shendure@uw.edu
https://doi.org/10.1016/j.molcel.2019.08.010

Transcription factors (TFs) bind DNA in a sequence-specific manner and thereby serve as the protein anchors and determinants of 3D genome organization. Conversely, chromatin conformation shapes TF activity, for example, by looping TF-bound enhancers to distally located target genes. Despite considerable effort, our understanding of the mechanistic relation between TFs and 3D genome organization remains limited, in large part due to this interdependency. In this review, we summarize the evidence for the diverse mechanisms by which TFs and their activity shape the 3D genome and vice versa. We further highlight outstanding questions and potential approaches for untangling the complex relation between TF activity and the 3D genome.

Transcription factors (TFs) are proteins that play a key role in regulating gene expression by recognizing and directly binding specific DNA sequences (Lambert et al., 2018). Upon binding DNA, TFs can activate transcription by directly (Chernukhin et al., 2007) or indirectly (via cofactors) bringing RNA polymerase II (Pol II) and other transcriptional machinery to the promoter sequence at the start of a gene (Allen and Taatjes, 2015). Alternatively, they can repress transcription by recruiting corepressors or interfering with the binding of other TFs (Braun and Johnson, 1997; Moody et al., 2005). Some TF binding sites occur near promoters, but in mammalian genomes, the overwhelming majority occur in clusters of binding sites called enhancers, which are typically tens to hundreds of kilobases away from the promoters they activate (Gasperini et al., 2019).

The action at a distance of enhancers is thought to be mediated by the three-dimensional (3D) looping of DNA that brings an enhancer into physical proximity with a promoter (Furlong and Levine, 2018). Toward assessing whether this is a universal phenomenon, recent technical advances have enabled increasingly comprehensive and high-resolution views of chromosome conformation (Bonev et al., 2017; Lieberman-Aiden et al., 2009; Rao et al., 2017). Although some link between 3D genome organization and TF activity seems certain, a detailed understanding of the mechanistic and causal relations between these two phenomena remains elusive. How often is physical enhancer-promoter contact frequency critical for transcriptional regulation? As DNA-binding proteins, TFs can directly and indirectly shape chromosome conformation in addition to being influenced by it. To what extent is chromosome conformation the consequence, rather than the cause, of TF activity? The diverse interactions among TFs and cofactors can also affect genome organization, further complicating the picture.

Here, we review the intricate interplay between TF activity and 3D genome organization, including the mechanisms by which TFs shape 3D genome organization, as well as the effects of chromosome conformation on TF activity. We focus primarily on mammalian genomes in interphase at the kilobase to megabase scale (i.e., beyond the known effects of TFs locally bending DNA; Kim et al., 1993), but include discussion of other organisms where relevant.

**Modes of TF Action on the 3D Genome**

TFs are defined by their ability to bind DNA, but generally function via interactions with other proteins and even RNAs (Lambert et al., 2018). These interactions have the potential to affect chromosome conformation. In this section, we outline and categorize the mechanisms by which TFs can shape the 3D genome (Figure 1). Note that for a given TF or locus, these mechanisms need not be mutually exclusive.

**Direct Oligomerization**

The most intuitive example of TF-driven DNA loops involves the direct oligomerization of TFs (Figure 1A). Some of the earliest evidence of such loops came from studies of cooperative repressor protein binding to DNA in prokaryotes (Griffith et al., 1986). Several repressors, such as LacI (Priest et al., 2014) and GalR (Qian et al., 2012) in *Escherichia coli* and the lambda phage Cl repressor (Griffith et al., 1986), bind cooperatively not only at nearby binding sites but they also can form larger oligomers that tether together distally located sets of binding sites, both in vitro and in vivo.

Eukaryotic TFs are also capable of producing similar DNA loops via self-association in vitro (Su et al., 1991), although whether this self-association is structured and how this interaction behaves in vivo are less clear. Yin Yang 1 (YY1) is a ubiquitously expressed TF that binds both promoters and enhancers and forms homodimers, and thus could mediate enhancer-promoter looping by a similar mechanism (Weintraub et al., 2017). CCCTC-binding factor (CTCF) is another ubiquitous TF with a key architectural role in mammals and is thought to form dimers (and possibly oligomers) in vitro and in vivo (Pant et al., 2004). However, as we discuss in subsequent sections, it is unclear whether this dimerization is the means by which CTCF shapes 3D genome organization.
TFs can also form DNA loops by recruiting cofactor proteins that in turn form oligomers (Figure 1B). For example, LIM domain binding protein 1 (Ldb1) is a notable adaptor protein capable of dimerization that does not bind DNA directly, but is instead recruited to its target loci by TFs or cofactors. In mouse olfactory sensory neurons, it is recruited by the TFs Lhx2 and Ebf1 (Monahan et al., 2019), whereas in mouse erythroid cells, the cofactor Lmo2 bridges the interaction of Ldb1 with the TFs Gata1, Tal1, and E2A (Love et al., 2014). Ldb1 recruitment is also sufficient for DNA looping, as evidenced by studies in which Ldb1, fused to a designed zinc finger protein targeting either the beta or gamma globin gene promoters, induced looping with the locus control region (LCR), resulting in transcriptional activation (Deng et al., 2012, 2014).

**Cofactor Oligomerization**
TFs can also form DNA loops by recruiting cofactor proteins that in turn form oligomers (Figure 1B). For example, LIM domain binding protein 1 (Ldb1) is a notable adaptor protein capable of dimerization that does not bind DNA directly, but is instead recruited to its target loci by TFs or cofactors. In mouse olfactory sensory neurons, it is recruited by the TFs Lhx2 and Ebf1 (Monahan et al., 2019), whereas in mouse erythroid cells, the cofactor Lmo2 bridges the interaction of Ldb1 with the TFs Gata1, Tal1, and E2A (Love et al., 2014). Ldb1 recruitment is also sufficient for DNA looping, as evidenced by studies in which Ldb1, fused to a designed zinc finger protein targeting either the beta or gamma globin gene promoters, induced looping with the locus control region (LCR), resulting in transcriptional activation (Deng et al., 2012, 2014).

**Condensate Formation**
In addition to the strong protein-protein interactions at the structured interfaces that typically underlie oligomerization, weak but multivalent interactions among intrinsically disordered regions (IDRs) have been proposed to result in condensates or hubs...
that exhibit properties of liquid-liquid phase separation (Banani et al., 2017). This trending topic, which is relevant beyond the nucleus and protein-protein interactions, has recently been extensively reviewed (Alberti et al., 2017; Banani et al., 2017; Boeynaems et al., 2018); here, we focus specifically on its relevance to genome organization (Figure 1C). The hallmarks of phase separation include the formation of spherical droplets, with a composition distinct from the surrounding solution, when the components are above critical concentrations (dependent on conditions such as salt concentration or pH), and which can merge (e.g., oil droplets in water); rapid fluorescence recovery after photobleaching (which implies dynamic rather than static structures); and sometimes conversion into a gel- or solid-like state over time, in addition to the aforementioned multivalent weak interactions among disordered regions. However, the distinction between phase separation and smaller assemblies that may be below the detection limit of microscopy remains unclear. For example, what is the minimum size of a phase-separated droplet? Some have proposed the term “hub” or “mini-hub” to describe smaller assemblies that can grow into phase-separated droplets (Chong et al., 2018). Whether phase separation per se confers unique properties that are essential for genome organization remains debatable; nevertheless, growing evidence supports a role for these assemblies. Here, we use the term “condensates” to refer loosely to such assemblies, regardless of whether they phase separate.

TFs and coactivators are particularly enriched for IDRs (Staby et al., 2017). Recent studies have demonstrated several instances in which TFs and cofactors form condensates in vitro and in vivo in an IDR-dependent manner (Bojia et al., 2018; Cho et al., 2018; Chong et al., 2018; Sabari et al., 2018). Enhancers that share binding by specific TFs (e.g., estrogen receptor α, Notch) or sets of TFs and/or cofactors (e.g., Oct4– Sox2–Nanog, Ubx–Hth) have been observed to colocalize, albeit probably transiently, in “nuclear microenvironments” or “3D cliques” that are enriched for those factors (Denholtz et al., 2013; Nair et al., 2019; Petrovic et al., 2019; Tsai et al., 2017; de Wit et al., 2013). Given that these TFs are not known to oligomerize, these enhancer hubs are consistent with condensates shaping chromosome conformation. Furthermore, super-enhancers (regions with a high density of TFs and cofactor binding that drive the robust expression of target genes) display particularly frequent contacts with other enhancers bound by the same TFs (Petrovic et al., 2019), supporting the idea that the multivalency of super-enhancers helps seed condensates (Hnisz et al., 2017). Similarly, the heterochromatin protein HP1α is capable of forming condensates and of compacting DNA in vitro, which is consistent with a role in the self-association and compaction of heterochromatic regions of the genome (Larson et al., 2017; Strom et al., 2017).

Despite considerable excitement, it is worth emphasizing that relatively few experiments have been reported that would establish a causal role for condensates in shaping the 3D genome. As one encouraging example, the distance between two estradiol-induced enhancers (measured by fluorescence in situ hybridization [FISH]) was increased upon 1,6-hexanediol treatment, which disrupts condensates, compared to a 2,5-hexanediol control (Nair et al., 2019). Another study developed a light-controlled tool to trigger condensate formation at specific genomic loci; their results suggest that condensates can selectively engulf targeted regions of the genome and pull them together as the condensates coalesce due to surface tension (Shin et al., 2019). However, the effects of forming or disrupting condensates on the 3D genome have yet to be studied in a systematic, genome-wide fashion.

**Interactions with Loop Extruders**

A unique class of interactions that plays an important role in 3D genome organization is between TFs and cohesin and other structural maintenance of chromosomes (SMC) complexes implicated in loop extrusion (Figure 1D). SMC complexes can encircle DNA and actively pull a loop of DNA through the ring (Ganji et al., 2018) until they are blocked or unloaded. The loop extrusion model (Fudenberg et al., 2016; Sanborn et al., 2015) is consistent with a growing amount of experimental evidence, most notably in interphase nuclei, in which the blocking of cohesin complexes by CTCF explains the boundaries of topologically associating domains (TADs), which grow (Haahrhus et al., 2017) or disappear (Rao et al., 2017; Schwarzer et al., 2017) upon perturbations to cohesin unloading or loading, respectively. Other TFs may also be capable of altering cohesin loop positions. For example, recent experiments mutating the RNA-binding domain of CTCF led to new loop boundaries at binding sites for Oct4 and other pluripotency factors in mouse embryonic stem cells (Hansen et al., 2018).

**Interactions with Nuclear Landmarks**

Proteins localized to specific regions of the nucleus, such as nuclear pore components (Nups) and lamins, contribute to the nuclear localization of different genomic regions through direct or indirect interactions with TFs (Figure 1E). In budding yeast, most TFs are capable of recruiting DNA to nuclear pores (Brickner et al., 2019), which is associated with inducible gene expression (Brickner et al., 2012) and/or epigenetic transcriptional memory (Light et al., 2010). In metazoans, although various Nups associate with specific genomic regions in a cell type- and state-dependent manner (recently reviewed in Sun et al., 2019), the relocation of genes to nuclear pores is less prevalent; instead, mobile Nups may interact with chromatin in the nucleoplasm and at nuclear pores (Pascual-Garcia et al., 2017). Similarly, TFs can interact with the nuclear lamina, either directly or indirectly (through cofactors), to tether heterochromatin to the nuclear periphery (Zullo et al., 2012).

**Protein-RNA Interactions**

In addition to protein-protein interactions, protein-RNA binding plays a role in structuring the 3D genome (Figure 1F). Both YY1 and CTCF, ubiquitously expressed TFs with structural roles, are known to bind noncoding RNAs (ncRNAs) (Kung et al., 2015; Sigova et al., 2015). Specifically depleting these interactions, either by adding RNase in vitro or mutating the RNA-binding domain, weakens the ability of YY1 to form loops (Weintraub et al., 2017) and of CTCF to form TAD boundaries by blocking cohesin (Hansen et al., 2018). Other TF-ncRNA interactions contribute to the formation of condensates, as knockdown of enhancer RNAs (eRNAs) can inhibit TF recruitment to the enhancers, and consequently enhancer-enhancer looping (Hnisz et al., 2017; Nair et al., 2019). These protein-RNA interactions, in addition to the potential roles of other RNA-binding proteins
and splicing factors (Bertero et al., 2019; Xiao et al., 2019), provide plausible mechanisms by which noncoding RNAs such as eRNAs (Sigova et al., 2015) or long noncoding RNAs such as Xist (Kung et al., 2015) and Firre (Yang et al., 2015) may shape 3D genome organization.

**Chromatin Modification**

Finally, TFs can affect 3D genome organization by modifying their chromatin context, and thus the recruitment of other TFs or proteins (Figure 1G). Many TFs recruit cofactors that modify DNA (e.g., methylation), histone tails (e.g., lysine methylation, acetylation), or nucleosome positioning, which can affect the binding of other TFs or cofactors (O’Malley et al., 2016; Yin et al., 2017; Zhu et al., 2018). TFs or proteins that either specifically bind (or cannot bind) to modified DNA (e.g., methyl-CpG-binding protein 2 [MeCP2]) or histones (e.g., the acetyl-lysine-binding BRD4) can then shape chromosome conformation through protein-protein and protein-RNA interactions. Chromatin states are associated with multiple scales of genome organization: DNA methylation can affect CTCF binding and thereby enhancer-promoter looping and TAD boundaries (Flavahan et al., 2016; Wiehle et al., 2019), while larger genomic regions enriched in active or inactive chromatin marks self-associate in A or B compartments, respectively (Di Pierro et al., 2017; Lieberman-Aiden et al., 2009). Even without covalent modifications, some TFs known as pioneer factors can establish open chromatin by binding nucleosomal DNA (Zhu et al., 2018) and either displacing nucleosomes directly or recruiting chromatin remodelers (Spitz and Furlong, 2012). Quantitative trait locus (QTL) analysis suggests that the resulting changes in chromatin accessibility contribute to the influence of TFs on long-range 3D genome organization (Tehranchi et al., 2019). These indirect, chromatin-mediated effects may provide a means by which TFs with only transient binding to DNA can trigger more enduring changes to chromosome conformation and ultimately gene expression.

**A Mechanistic View of 3D Genome Structures**

We next apply the principles outlined above to review the state of our mechanistic understanding of the various features of 3D genome organization (Figure 2). However, direct evidence for specific mechanisms is limited; thus, we also highlight key questions that remain.
A-B Compartments

Each chromosome can be segmented into active A compartments and inactive B compartments, roughly 1–10 Mb, each of which preferentially associates with other compartments of the same identity (Figure 2A) (Lieberman-Aiden et al., 2009). This self-association is consistent with condensate formation and/or oligomerization by proteins specifically binding to either active or inactive chromatin states (Rowley et al., 2017), such as Pol II or HP1α, respectively. Note that both condensate formation and oligomerization can produce such self-association; distinguishing the two mechanisms will require further experiments (Erdel and Rippe, 2018). A and B compartments can be further divided into self-associating subcompartments A1–A2 and B1–B4, suggesting that this self-association is not binary, but instead is more quantitative and in line with the diversity of chromatin states (Rao et al., 2014). In support of this, modeling the genome as a series of subcompartments, each with a propensity to self-aggregate, predicts experimental chromatin conformation data from chromatin state data (Di Pierro et al., 2017). Although the A-B compartment pattern is generally similar across cell types, reflecting the alternating pattern of gene-rich and gene-poor regions of the genome, the pattern also shifts upon differentiation (Stadhouders et al., 2018), as pioneer TFs can initiate the opening or closing of chromatin (Zaret and Mango, 2016). However, due to the multitude of molecular changes that occur during differentiation, the relative contributions of specific factors remain unclear.

Furthermore, phase separation of active and inactive chromatin is antagonized by cohesin-mediated loop extrusion (Nuebler et al., 2018). Depleting the cohesin unloader, thereby increasing the processivity of cohesin, leads to weaker self-association of the A and B compartments (Haarhuis et al., 2017), while depleting cohesin itself leads to stronger and finer compartments (Schwarzer et al., 2017). This suggests that loop extrusion by cohesin both disaggregates compartments (i.e., reduces contacts between distant compartments) and merges adjacent compartments (i.e., creates compartments that span greater genomic distances).

TADs and Loop Domains

Compartments can be subdivided into TADs, roughly 0.1–1 Mb regions that preferentially self-associate and are usually bounded by CTCF binding sites and/or housekeeping gene promoters (Figure 2B) (Dixon et al., 2012; Nora et al., 2012). A growing body of evidence supports the role of loop extrusion by cohesin coupled with CTCF blocking of cohesin in the establishment of TADs (Fudenberg et al., 2016; Nora et al., 2017; Nuebler et al., 2018; Rao et al., 2017; Sanborn et al., 2015). However, the molecular mechanism by which CTCF blocks cohesin remains unclear. The bias toward convergent CTCF motifs at TAD boundaries and experiments inverting CTCF motifs suggest that CTCF may block cohesin in a directional manner (Guo et al., 2015; Rao et al., 2014; Sanborn et al., 2015). However, the structural basis of any such directional blocking remains unknown. CTCF is thought to dimerize, but is this simply a side effect of two molecules each blocking the progression of a single cohesin complex? Furthermore, despite its constitutive expression, CTCF binding varies across cell types; what features of chromatin state drive this differential binding? DNA methylation is known to affect CTCF binding (Flavahan et al., 2016; Wiehle et al., 2019), but other features such as eRNA transcription may also contribute, given the ability of CTCF to bind RNA (Hansen et al., 2018).

Recent studies have also highlighted the role of other factors in TAD establishment. A systematic study of the Dppa2/4 locus in mouse embryonic stem cells revealed that deletion of TAD boundaries is not sufficient to disrupt TAD structure, and internal early replication control elements (ERCEs) instead determine TAD structure (Simà et al., 2019). This is corroborated by other similar observations at other loci (Lupiáñez et al., 2015) and recent microscopy evidence that even upon cohesin depletion, TAD-like structures form in single cells but with random boundary locations (Bintu et al., 2018). Many ERCEs overlap enhancers and promoters, but do all enhancers and promoters contribute to TADs? If not, then what features determine their relative contributions? If condensate formation is responsible, one prediction would be that high densities of binding or recruitment sites for the driving factors would be correlated with greater roles for TAD establishment (Hnisz et al., 2017).

Architectural Stripes

Some TADs contain “architectural stripes” or frequently interacting regions (FIREs), regions that form many contacts throughout their domains (Figure 2C). Stripes are associated with tissue-specific super-enhancers and are thought to be formed by loop extrusion, through robust loading of cohesin accompanied by a cluster of unidirectional CTCF binding sites (a stripe anchor) that blocks loop extrusion from progressing in that direction (Kraft et al., 2019; Schmitt et al., 2016; Vian et al., 2018). As the DNA in the other direction is extruded through the cohesin complexes, each segment successively forms a contact with the stripe anchor, resulting in a stripe when contact maps are generated on a heterogeneous population of cells. What then determines the genomic positioning of cohesin loading? Do the condensate-forming properties of super-enhancers play a role? In mouse embryonic stem cells (mESCs), cohesin and its loader Nipbl are found at promoters and enhancers, in addition to CTCF sites (Kagey et al., 2010). In yeast, the RSC chromatin remodeler recruits cohesin loading (Muñoz et al., 2019), while in Dro sophila, DNA replication complexes have been implicated (Pherson et al., 2019). Another potential contributor is that the recruitment of Pol II to active enhancers increases the mobility of those loci, thereby increasing the frequency of long-range contacts (Gu et al., 2018).

Enhancer-Promoter Loops

Not all active enhancers are near stripe anchors, and yet enhancers presumably act on distal promoters by DNA looping. Do all active enhancers preferentially form specific contacts with their cognate promoters, or is the “random” background contact frequency from being in the same TAD sufficient for function (Figure 2)? The relative frequency of enhancer-promoter pairs with architectural stripes at either the enhancer or promoter (Figure 2C), focal enhancer-promoter contacts (Figure 2D), or TAD background contact frequency (Figure 2E) remains difficult to test systematically, due to the paucity of functionally validated enhancer-promoter pairs. In at least some cases, however, enhancers appear to form focal contacts with their target genes (Bonev et al., 2017; Gasperini et al., 2019).
What are the mechanisms of these contacts, and are all contacts functionally equivalent? In the canonical example of enhancer-promoter (E-P) looping at the beta globin locus, loops are mediated by dimerization of the adaptor protein Ldb1 (Love et al., 2014). A small hairpin RNA (shRNA) screen in mESCs implicated the Mediator and cohesin complexes (Kagey et al., 2010), but how these shape enhancer-promoter loops is unclear. Mediator can form condensates (Cho et al., 2018), along with BRD4 and other coactivators (Sabari et al., 2018), that can bridge enhancers and promoters, but it is unclear whether this provides specificity among enhancers and promoters bound by specific TFs (Tsai et al., 2017). A key remaining question is what is the logic of condensate formation—how many distinct types of condensates are there, and what drives their specificity? IDRs have been classified into categories by their amino acid composition (e.g., acidic, proline rich, glutamine rich). Do these IDR classes segregate into distinct condensates? Mediator and cohesin also interact in vitro (Kagey et al., 2010), but how do they interact in vivo? Do Mediator and other transcriptional complexes block loop extrusion by cohesin? Recent studies using the Micro-C variant of Hi-C to obtain nucleosome-resolution contact maps revealed that many promoters and enhancers produce boundaries of “microTADs” with stripes and loops reminiscent of loop extrusion processes (Heisz et al., 2019; Krietenstein et al., 2019). YY1 dimerization may also contribute to enhancer-promoter loop formations by binding both the enhancer and promoter and tethering them together (Weintraub et al., 2017), with an additional role for enhancer RNAs that aid in YY1 recruitment (Sigova et al., 2015).

**Trans Contacts between TADs and Chromosomes**

Although recent attention has been focused on contacts within TADs, some of the first clues that aggregative processes such as condensate formation may shape genome conformation came from observations of colocalization among co-regulated genes across multiple chromosomes (Figure 2F) (Schoenfelder et al., 2010). In mouse erythroid cells, genes regulated by Klf1 are enriched for colocalization with one another and with clusters of Pol II in “transcription factories” (Schoenfelder et al., 2010). In vivo microscopy of Drosophila embryos confirmed a similar association of homologous and ectopic svb alleles (Tsai et al., 2017). Condensate formation by TFs and associated cofactors is a potential mechanism of such contacts (Boja et al., 2018; Chong et al., 2018), but mutations and drugs that disrupt condensate formation have yet to be tested for effects on trans contacts. Oligomerizing TFs and cofactors also contribute to interchromosomal contacts. The cofactor Ldb1 mediates interchromosomal contacts among olfactory gene clusters in mouse olfactory sensory neurons (Monahan et al., 2019). However, in most cases, whether oligomerization per se shapes 3D genome organization is unclear. In heat-shocked yeast, the heat shock response protein Hsf1 mediates the clustering of Hsf1 target genes (Chowdhary et al., 2019), but whether the ability of Hsf1 to form trimers mediates these contacts remains uncertain. Similarly, YY1 and CTCF are implicated in the transient homolog pairing of the X inactivation centers in differentiating mESCs, in which the role of CTCF is presumably unrelated to loop extrusion, which is thought to act on individual chromosomes rather than on pairs of chromosomes (Xu et al., 2007). Oligomerizing factors do not form all possible contacts among their binding sites, suggesting that additional mechanisms must contribute to trans contacts, particularly for highly restricted contacts such as focal homolog pairing (Hogan et al., 2015; Kim et al., 2017; Xu et al., 2006). For example, Rgt1 mediates homologous pairing between TAD1 promoters and HXT3 promoters in saturated yeast, but not contacts between the two distinct promoters (Kim et al., 2019b). Future work testing perturbations of cofactors, specific domains, and catalytic positions will help distinguish whether these TFs mediate pairing through direct oligomerization, chromatin state, or other effects.

**Nuclear Localization**

Many DNA-DNA contacts, including those in cis but particularly those in trans, reflect shared localization at a nuclear subcompartment (i.e., a region of the nucleus with distinct nucleic acid and protein composition) (Figure 2G). For example, gene-dense and highly transcribed regions form hubs of contacts at nuclear speckles (Chen et al., 2018b; Quinodoz et al., 2018), whereas silenced heterochromatonic regions colocalize at nucleoli (Quinodoz et al., 2018) or the nuclear lamina (Guelen et al., 2008). Regions specifically targeted by the Polycomb repressive complex form contacts at Polycomb bodies (Denhardtz et al., 2013).

In each case, the colocalization of specific genomic regions at these membraneless compartments is consistent with condensates, whether they are formed by TFs, cofactors, transcription machinery, and/or splicing machinery. Recent studies have begun to demonstrate specific examples of factors, such as HP1α, that could drive phase separation. However, each sub-compartment often contains many enriched factors (Fong et al., 2013), and their contributions to condensate formation are unclear. Can a factor be colocalized with such condensates but be dispensable for their formation? Which factors interact with the nuclear landmarks and thereby position these condensates within the nucleus? These compartments are also each associated with either active or inactive transcription, and the role of transcription or lack thereof in genomic localization is complex and will require additional investigation (van Steensel and Furlong, 2019). For example, using an artificial TF to temporarily decondense chromatin at a locus without affecting transcription was sufficient to reposition the locus away from the nuclear periphery, suggesting that some of the effects of TF binding and transcription on locus localization may be indirect (Therizols et al., 2014).

Many of these nuclear subcompartments are also associated with specific chromatin marks, such as H3K9me3 for heterochromatin. Are chromatin marks sufficient for, or simply correlated with, nuclear localization? In the case of H3K9me3, recruiting SUV39H1 to a lacO array was sufficient to reposition the locus to heterochromatin, but a chromodomain mutant capable of depositing H3K9me3 but unable to bind it was not (Wijchers et al., 2016). Thus, protein binding to specific histone modifications is likely to play a key role in nuclear organization.

Also, what role does random genomic colocalization play in establishing nuclear subcompartments? The local abundance of binding sites for condensate-forming factors is a key parameter in determining their formation (Alberti, 2017; Chong et al., 2018; Hnisz et al., 2017), and once formed, a transient contact can be stabilized. In some cases, recruiting a nucleating factor to a single locus can be sufficient to create a condensate (e.g.,
for Cajal bodies) (Wang et al., 2018), but this may not be true for other types of subcompartments.

The Implications of the 3D Genome for TF Function

Thus far, we have discussed the mechanisms by which TFs can shape 3D genome organization. In this section, we consider the converse—how does 3D genome organization affect TF activity (Figure 3)? Of course, TF genes, like any other genes, are subject to 3D-genome-mediated transcriptional regulation. For example, the activation of the Nanog locus during cellular reprogramming first requires a nearby B-to-A compartment shift that allows super-enhancer-mediated activation of the locus (Stadhouders et al., 2018). However, below, we focus on the unique effects of chromosome conformation on the activity of TF proteins.

Enhancer-Promoter Specificity

One major way in which chromosome conformation is thought to regulate TF activity is by providing specificity in pairing between enhancers and the genes they target. This idea has led to efforts to use targeted assays of 3D genome organization, such as 4C (Simonis et al., 2006; Zhao et al., 2006), chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) (Fullwood et al., 2009), Hi-C chromatin immunoprecipitation (HiChIP) (Mumbach et al., 2016), promoter capture Hi-C (Javierre et al., 2016; Mifsud et al., 2015; Schoenfelder et al., 2015), and proximity ligation-assisted ChIP-seq (PLAC-seq) (Fang et al., 2016) to associate enhancers (and their associated TF-cofactor-Pol II condensate) with candidate target genes. However, at least in some cases, physical proximity precedes enhancer-promoter activation, whether during development or in vitro differentiation (Figure 3A) (Ghavi-Helm et al., 2014; Stadhouders et al., 2018). Furthermore, enhancer-promoter pairing, both in the sense of the 3D genome and function, is not binary but quantitative. What contact frequency is sufficient for an interaction to be functional?

Until recently, there have been insufficient data on enhancer-promoter activation to evaluate the correlation between physical

Figure 3. Functional Implications of the 3D Genome

(A) The formation of an enhancer-promoter loop that activates the promoter requires enhancer activation and loop formation. These two steps can occur synchronously or asynchronously (through development or stimulation time courses), suggesting that the 3D genome contributes but is not always sufficient for activation.

(B) The specificity of enhancer-promoter activation is jointly influenced by enhancer activity, promoter responsiveness (e.g., via chromatin accessibility), and contact frequency. From the perspective of a promoter (top panel), the contribution of an enhancer is predicted by the product of enhancer activity and contact frequency. From the perspective of an enhancer (bottom panel), nearby and thus frequently contacting promoters tend to be activated, with the exception of promoters that are unresponsive.

(C) Clusters of binding sites can nucleate condensates of TFs, which can exclude specific factors and mediate binding to low-affinity TF binding sites.

(D) Promoter looping with an enhancer (and potentially its associated TF-cofactor-Pol II condensate) is correlated with bursts of transcription.
and functional pairing. Two recent studies have begun to address this by using CRISPR interference (CRISPRi) to perturb enhancer function at scale, coupled with either single-cell RNA sequencing (RNA-seq) (Gasperini et al., 2019) or Flow-FISH to identify genes with altered expression (Fulco et al., 2019). In both studies, functional enhancer-promoter pairs exhibited more frequent physical contacts than did controls. However, physical contacts alone were not sufficient to predict functional pairing; instead, a combination of 3D contacts and enhancer activity, as measured by DNase hypersensitivity and H3K27ac signal, predicts the relative contribution of an enhancer to the activation of a target gene (Figure 3B) (Fulco et al., 2019). Notably, the ability to predict experimental enhancer perturbation data was only marginally affected by excluding high-resolution Hi-C data, suggesting that genomic proximity, rather than specific TAD boundaries or focal enhancer-promoter loops, plays a major role in governing specificity. However, this model does not explain why some promoters, despite proximity to strong enhancers, are not activated. In addition to enhancer activity levels, promoter activity levels—regulated by heterochromatin and interactions with, for example, the nuclear lamina, binding of TFs and cofactors, chromatin accessibility, and histone modifications—could regulate the ability to be activated by enhancers (Figure 3B) (Haberle et al., 2019). Furthermore, there may be some logic by which TFs at enhancers pair with TFs at promoters (Gasperini et al., 2019), perhaps mediated by shared interactions with certain cofactors (Haberle et al., 2019; Stampfel et al., 2015).

The mechanistic details of physical enhancer-promoter pairing can affect enhancer function beyond simply providing specificity of pairing. For example, whether an enhancer can simultaneously pair with multiple promoters would affect whether those promoters need to compete for activation by the enhancer (Fukaya et al., 2016). Although Hi-C and related datasets are informative with respect to temporal dynamics, the timing and stability of enhancer-promoter loops may affect the variability in enhancer activation of target promoters.

**Biophysics of TF Activity**

The spatial organization of the genome, and more broadly the nucleus, fundamentally affects the biophysics of TF activity. The DNA-binding activity of TFs is a function not only of the TF affinity for specific DNA sequences and the overall concentration of the TF but also the localization of the TF within the nucleus (Figure 3C). A cluster of binding sites for a TF, whether nearby in genomic sequence or brought together by DNA loops, could produce a locally elevated concentration of the TF that in turn allows binding even at low overall TF levels and/or at low-affinity binding sites that differ considerably from the consensus motif (Mira et al., 2018; Tsai et al., 2017; Vockley et al., 2016). In some cases, such as the glucocorticoid response, the low-affinity binding sites are not sufficient for TF activity and instead amplify the activity of high-affinity sites (Vockley et al., 2016). In other contexts, including Drosophila (Tsai et al., 2017) and Ciona embryos (Farley et al., 2016), however, low-affinity binding sites are capable of and even necessary for proper tissue-specific expression.

Just as nuclear microenvironments can enrich TFs and coactivators, they can exclude factors as well (Figure 3C) (Strom et al., 2017). Furthermore, the localization of a TF (e.g., at the nuclear periphery) could sequester it away from its genomic targets (Malhas et al., 2009). Precisely how heterochromatin and other repressive environments inhibit transcription is unclear; they may physically exclude TFs and transcription machinery, perhaps through the higher density of chromatin (Ou et al., 2017), but TFs that do enter the heterochromatic regions may be transiently trapped (Bancaud et al., 2009). These spatial components of TF activity expand the potential range of mechanisms by which chromatin state can regulate TF activity (Zheng et al., 2019).

Nuclear architecture also shapes the temporal dynamics of TF activity and subsequent transcription (Figure 3D). At many genes, transcription occurs in bursts (Suter et al., 2011), with transcription levels controlled through burst frequency rather than burst size (Bartman et al., 2019). At least in cases in which the enhancer is relatively far from the promoter in genomic distance, enhancer-promoter contact frequency is correlated with transcription (Chen et al., 2018a), while the size of the Pol II focus is correlated with the number of transcripts produced in a burst (Cho et al., 2016). Thus, stable enhancer-promoter looping and perhaps even clustering among co-regulated genes could mediate robust transcription for crucial genes (Hnisz et al., 2017).

In addition to potentiating robust expression, chromosome conformation can contribute to stochastic or monoallelic expression. In both endogenous and engineered systems, infrequent chromosomal contacts can lead to gene activation, perhaps by facilitating the assembly of TFs (Apostolou and Thanos, 2008). The stochasticity of these contacts may mediate the stochastic or monoallelic activation of a promoter (Canzio et al., 2019; Guo et al., 2012; Monahan et al., 2019). However, a functional role for 3D genome organization in this context is speculative, and establishing causality remains challenging. For example, homolog pairing of the X chromosomes coincides with the random choice of one chromosome for silencing in females (Xu et al., 2008), but recent tethering experiments do not support a causative role (Pollex and Heard, 2019).

**Outlook and Future Directions**

The relation between the 3D organization of the genome and the activity of TFs is complex, with interdependencies and several modes of action. Each of the biochemical activities of TFs—binding DNA, opening chromatin, recruiting cofactors, and possibly oligomerizing—may affect 3D genome organization. Conversely, chromosome conformation affects TF function and the pairing of enhancers with promoters. Despite this complexity, we are optimistic that the ongoing development and improvement of technologies for measuring and perturbing both genome architecture and TFs offer the hope of teasing apart these interdependencies (Figure 4).

A first step is to identify which TFs play a role in 3D genome organization. In well-studied cell lines, existing ChIP-seq datasets can be analyzed to identify candidates that may shape 3D genome organization. Conversely, chromosome conformation affects TF function and the pairing of enhancers with promoters. Nonetheless, ChIP-seq is fundamentally descriptive, and it is impossible to assign causality from the descriptive
data alone. As such, it will be critical to also develop and apply high-throughput methods that test perturbations of either DNA sequences or TFs on the 3D genome (Figure 4A). For example, we recently developed MAP-C (mutation analysis in pools by chromosome conformation capture), which couples saturation mutagenesis and/or genetic screens to the 3C assay to systematically identify either loop anchor regions or TFs required for a chromosomal contact of interest (Kim et al., 2019b). Most screens to date have focused on perturbing one TF at a time; however, given the possibility of functional redundancy (e.g., due to paralogous TFs), combinatorial screens may be warranted.

However, neither perturbations to TF binding sites nor TF expression can determine how a TF affects 3D genome organization. Instead, mutations of protein domains or specific catalytic activities are necessary to reveal specific TF functions that are required (Figure 4B). For example, deletion of the CTCF RNA-binding domain (Hansen et al., 2018) supported the relevance of RNA binding to 3D genome organization.

To determine which specific functions of a given TF (e.g., cofactor recruitment) are sufficient for structuring the genome, TF recruitment and bypass experiments can be used (Figure 4C). The canonical example is a study in which a zinc finger protein fused to the Ldb1 self-association domain was
sufficient to induce looping and activation of globin genes (Deng et al., 2012, 2014). Numerous other studies have used similar TF bypass experiments but did not examine 3D genome organization. Stark and colleagues have used fusions between the Gal4 DNA-binding domain and TFs or cofactors to study TF and cofactor function within enhancers (Stampfel et al., 2015), while others have used dCas9 fusion proteins to recruit histone acetyltransferase (Hilton et al., 2015), DNA methyltransferase (Pflueger et al., 2018), and other chromatin remodelers. Catalytic or domain mutations can be combined with TF bypass experiments to further narrow down specific cofactor functions, such as SUV39H1 binding to H3K9me3, which is necessary for restructuring the genome (Wijchers et al., 2016). In addition, dCas9 and other DNA-binding domains have been used to perturb specific aspects of 3D genome organization (Kim et al., 2019a; Morgan et al., 2017; Pollex and Heard, 2019; Reddy et al., 2008; Wang et al., 2018; Wijchers et al., 2016) and could be used to study their effects on TF function.

An exciting alternative approach for dissecting the mechanisms shaping 3D genome organization is dCas9 proteomics (Figure 4D). Studies of the proteins binding to DNA generally require a priori knowledge of proteins of interest, as they must be labeled via genomic addition of epitope tags or with specific antibodies. However, recent studies have enabled the converse, by coupling mass spectrometry with enrichment for the proteins in the proximity of a specific genomic locus, for example, targetting dCas9 either fused with a biotinylated tag (Liu et al., 2017) or APEX2 (Gao et al., 2018; Myers et al., 2018), which can create radicals that covalently tag nearby proteins. These methods extend to individual genomic loci previous approaches that identified interaction partners of chromatin complexes (Alekseyenko et al., 2014; Mohammed et al., 2016) and may aid the systematic search for proteins involved in specific chromosome conformation contacts.

An added challenge to resolving the interplay between TFs and 3D genome organization is the inherent cell-to-cell variability in both TF abundance and chromosome conformation, in addition to the mixture of cell types in most tissue samples. Single-cell versions of Hi-C can resolve distinct cell types and states mixed in a sample (Nagano et al., 2017; Ramani et al., 2017) and with imputation are capable of determining low-resolution structures for the entire diploid genome (Tan et al., 2018). With further improvements in molecular efficiency and other simultaneous measurements of chromatin state, these approaches may help clarify the sources and roles of cellular heterogeneity in gene regulation.

In addition to perturbations, temporal resolution is another means of clarifying the mechanism. Although we have focused here on 3C-based technologies for studying 3D genome organization, microscopy methods are dramatically improving in resolution, throughput, and scalability (Bintu et al., 2018; Mateo et al., 2019). The molecular efficiency of FISH allows a high-resolution single-cell view of genome conformation that is not yet achievable by single-cell Hi-C methods (Nagano et al., 2017; Ramani et al., 2017; Tan et al., 2018), with the added possibility of measuring RNA in the same cells (Mateo et al., 2019). However, both FISH and Hi-C require fixed cells, which limits temporal resolution to bulk time courses (Vian et al., 2018) or reconstruction of trajectories such as the cell cycle from distinct single cells (Nagano et al., 2017). In contrast, live imaging methods allow the direct tracking of individual genomic loci, or even molecules, over the course of minutes or hours (Figure 4E) (Chen et al., 2018a; Gu et al., 2018; Ma et al., 2015; Maass et al., 2018; Mir et al., 2018). These approaches are revealing the nature of TF and genomic motion and its relation with transcription. For example, enhancer-promoter looping is temporally correlated with active transcription (Chen et al., 2018a), and the size of a Pol II cluster is correlated with the number of transcripts subse- quently produced (Cho et al., 2016). Single-molecule imaging can reveal the timescales of TF binding to DNA, which temporarily slows TF motion (Hansen et al., 2017). Recent advances have enabled the single-molecule resolution tracking of individual gene loci and nearby Pol II, TFs, or cofactors (Li et al., 2019). However, live imaging remains limited in throughput—for example, in terms of number of genomic loci and the number of features that can be labeled simultaneously. As the capacity to simultaneously label features grows, we may eventually be able to see how different TFs work together and with other factors in the context of the 3D genome.

The 3D organization of the nucleus, including the genome, has shed new light on our understanding of how TFs function. This new perspective explains puzzles ranging from how enhancers can act at a distance upon multiple genes to how low-affinity binding sites can function. Conversely, a TF-centric view of the 3D genome reveals the physical basis of sequence-specific DNA conformation. Looking forward, the mechanistic dissection of the role of TFs in shaping the 3D genome promises to yield insights into both how TFs work and how the genome is organized in 3D.

ACKNOWLEDGMENTS

We thank K. Xue, M. Spielmann, W. Noble, and C. Disteche for comments and discussion. S.K. was supported by a National Science Foundation (NSF) Graduate Research Fellowship (DGE-1256082). J.S. is an Investigator of the Howard Hughes Medical Institute. J.S. is a co-director of the University of Washington Center for Nuclear Organization and Function (U54 DK107979).

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES


Molecular Cell

Review


