



Review

Architectural proteins, transcription, and the three-dimensional organization of the genome



Caelin Cubeñas-Potts, Victor G. Corces*

Department of Biology, Emory University, 1510 Clifton Rd NE, Atlanta, GA 30322, USA

ARTICLE INFO

Article history:

Received 2 April 2015

Revised 7 May 2015

Accepted 9 May 2015

Available online 22 May 2015

Edited by Wilhelm Just

Keywords:

Chromatin

Epigenetics

Topologically associating domain

Insulator

CCCTC-binding factor

ABSTRACT

Architectural proteins mediate interactions between distant sequences in the genome. Two well-characterized functions of architectural protein interactions include the tethering of enhancers to promoters and bringing together Polycomb-containing sites to facilitate silencing. The nature of which sequences interact genome-wide appears to be determined by the orientation of the architectural protein binding sites as well as the number and identity of architectural proteins present. Ultimately, long range chromatin interactions result in the formation of loops within the chromatin fiber. In this review, we discuss data suggesting that architectural proteins mediate long range chromatin interactions that both facilitate and hinder neighboring interactions, compartmentalizing the genome into regions of highly interacting chromatin domains.

© 2015 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

1. Introduction

Multiple lines of evidence indicate that chromosome organization is a contributor to gene expression regulation [1,2]. The use of 3C-derived approaches to detect intra- and inter-chromosome interactions has led to the observation that individual chromosomes are highly organized structures. Chromatin interactions decrease with increasing linear genomic distance and occur non-randomly across the chromosome length [3–6]. Based on the frequency of these interactions, chromosomes can be divided into distinct regions of highly interacting chromatin, named topologically associating domains (TADs), which engage in few long-range interactions with loci in other TADs [7]. Architectural proteins, also known as insulator proteins, appear to play a critical role in the three-dimensional organization of the genome. Here we discuss known architectural proteins in *Drosophila* and mammals, and describe evidence suggesting that architectural proteins regulate long range chromatin contacts and ultimately, gene expression. Current results suggest that architectural proteins have two

inter-related functions, genome compartmentalization and the facilitation of interactions between regulatory elements. Finally, we end with a discussion of the molecular mechanisms regulating interactions between distant architectural protein binding sites.

2. Architectural proteins

The roles of architectural proteins in genome organization and function can be explained by their ability to facilitate the formation of long-range contacts between DNA sequences. In *Drosophila*, 11 different DNA binding architectural proteins have been identified, each recognizing a unique DNA motif: CCCTC-binding factor (CTCF), Suppressor of Hairy-wing (Su(Hw)), Boundary Element Associated Factor 32 (BEAF-32), DNA Replication Related Element binding Factor (DREF), Transcription Factor IIIC (TFIIIC), Z4 (also called Putzig), Early Boundary Activity DNA-binding Factor (Elba), Pita (also called Spotted dick), Zinc Finger Interacting with CP190 (ZIPIC), Insulator binding factor 1 (Ibf1), and Insulator binding factor 2 (Ibf2) [8–15]. ChIP-seq experiments demonstrating co-occupancy in the genome as well as coimmunoprecipitation studies have demonstrated that DNA binding architectural proteins interact with accessory proteins, which do not recognize specific DNA motifs [9,10,16]. The accessory proteins identified in *Drosophila* include Centrosomal Protein 190 (CP190), Modifier of mdg4 (Mod(mdg4)), Rad21 (a component of the cohesin complex), Cap-H2 (a component of the condensin II complex), the long

Abbreviations: ChIP-seq, chromatin immunoprecipitation-sequencing; 3C, chromosome conformation capture; ChIA-PET, chromatin interaction analysis by paired-end tag sequencing; STARR-seq, self-transcribing active regulatory region sequencing

* Corresponding author. Fax: +1 404 727 2880.

E-mail address: vcorces@emory.edu (V.G. Corces).

isoform of Female sterile homeotic on chromosome 1 (Fs(1)h-L), Lethal (3) malignant brain tumor (L3mbt), and Chromator (also called Chriz) [10]. Notably, depletion of either CTCF or CP190 reduced the chromatin interactions in the *Abd-B* locus by 3C analysis, suggesting that both DNA-binding and accessory architectural proteins can functionally contribute to chromatin looping interactions in cells [17]. Of particular interest, a recent *in vitro* analysis has provided a model for how DNA-binding and accessory architectural proteins function together to mediate long range chromatin interactions. Purified BEAF-32 protein was capable of binding its DNA motif, but only formed intermolecular interactions between two BEAF-32 motifs in the presence of the accessory proteins Chromator or CP190 [18]. However, how DNA-binding and accessory architectural proteins interact to mediate intra-chromosomal interactions within a cell remains to be determined.

The number of architectural proteins characterized in mammals is not as extensive as in *Drosophila*. Multiple lines of evidence indicate that CTCF and cohesin are mediators of chromatin interactions and thus, architectural proteins. Functional studies evaluating the function of CTCF and cohesin in chromatin looping within individual genomic loci were the first studies to indicate these proteins are regulators of chromatin interactions [19–22]. More recently, CTCF and cohesin depletion studies have shown that loss of these architectural proteins reduces genome-wide chromatin interactions [23–25]. Furthermore, ChIA-PET analysis, a technique that maps the chromatin interactions occurring between loci occupied by a specific protein, characterized a subset of the chromatin interactions that occur between cohesin and CTCF occupied sites, indicative that cohesin and CTCF may play a role mediating chromatin interactions [26,27]. In addition to CTCF and cohesin, the cohesin interacting proteins Nipbl (the protein responsible for loading cohesin rings onto DNA), and Mediator have also been implicated as mammalian architectural proteins due to their interactions with cohesin and enrichment at enhancer–promoter contact sites [28,29]. Similar to *Drosophila*, a series of proteins that co-localize or directly interact with CTCF have been identified in mammals, including Yin Yang 1 (YY1), Kaiso, Chromodomain Helicase-DNA-binding protein 8 (CHD8), Poly ADP-Ribose Polymerase 1 (PARP1), MYC-associated zing-finger protein (MAZ), jun-D proto-oncogene (JUND), ZNF143, nucleophosmin, the PR domain zinc-finger protein 5 (PRDM5), and TFII-I [30,31]. Through their association with CTCF, it is possible that CTCF interacting proteins also function as architectural proteins but additional experiments are required to address this hypothesis. Notably, there is a growing body of evidence that the CTCF-interacting protein ZNF143 is a mammalian architectural protein. The genomic occupancy of ZNF143 was shown to highly correlate with CTCF sites forming chromatin loops and ZNF143 depletion studies demonstrated a functional role for this protein in mediating long range chromatin interactions [32,33]. Beyond the cohesin and CTCF interacting proteins, a number of other potential architectural proteins have been characterized as proteins required for chromatin loop formation within specific genomic loci like CHD6 in the CFTR locus or Ldb1 in the β -globin locus [34,35]. However, the significance of CHD6 and Ldb1 as genome-wide regulators of chromatin interactions is not currently known. In summary, these data suggest that, similar to *Drosophila*, mammals express a wide array of potential architectural proteins important for regulating long range interactions that should be subject to additional characterization.

In mammalian cells, the architectural protein cohesin is an important contributor to the regulation of nuclear size and organization. For example, depletion of cohesin in astrocytes caused an approximately 25% increase in the volume of the entire nucleus, possibly due to a loss of chromatin interactions and interphase

chromatin organization [23]. In a reciprocal analysis, deletion of the cohesin regulator Wapl caused excessive cohesin and CTCF occupancy on chromatin and a striking hypercondensation of interphase DNA was observed [36]. It is interesting to speculate that the accumulation of cohesin and CTCF on chromatin resulted in excessive chromatin interactions responsible for the condensation phenotype. However, additional characterization is required to determine if these changes in nuclear morphology were direct effects of altered chromatin interactions.

3. General mechanisms of architectural protein function

Many functions of architectural proteins can be explained by their ability to mediate interactions between distant loci and form chromatin loops. Architectural proteins directly binding enhancers and promoters increase the contact frequency between regulatory elements by forming stable protein–protein interactions between them, consistent with the classical model of enhancer–promoter interactions (Fig. 1A) [32,37–40]. In addition, polymer simulations have suggested that an architectural protein interaction can facilitate neighboring interactions by two additional mechanisms [41,42]. First, regulatory elements that are looped out by two interacting architectural protein-bound loci have higher contact frequencies, indicative that a chromatin loop highly interacts within itself (Fig. 1B) [41,42]. Secondly, the genomic elements flanking a chromatin loop are brought into closer proximity by architectural proteins, reducing the linear genomic distance between them and increasing their contact frequency (Fig. 1C) [41]. However, additional evidence is required to determine if the simulation studies are representative of mechanisms of architectural protein-mediated facilitation in cells. In addition to facilitating some interactions, the establishment of a chromatin loop by architectural proteins also precludes other interactions. The original function ascribed to architectural proteins was their ability to insulate promoters from the effect of regulatory sequences such as enhancers, consistent with a function in hindering chromatin interactions [19,43–51]. Furthermore, a reduction of contact frequency for interactions between sequences located within a chromatin loop and sequences present outside the loop was observed in polymer simulation studies, supporting the notion that an architectural protein-mediated interaction will hinder a subset of

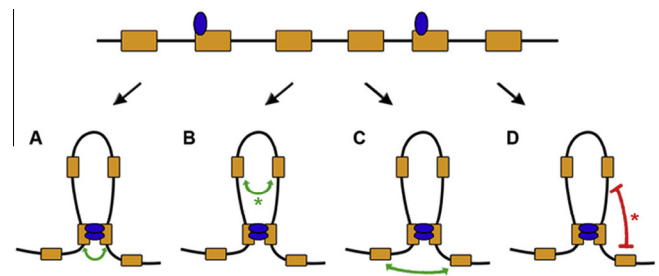


Fig. 1. The consequences of a single architectural protein interaction. Architectural proteins organize regulatory elements within the genome. The facilitating and inhibitory effects of a single architectural protein interaction between two genomic loci are shown. Regulatory elements are shown as gold boxes and architectural proteins are in blue. Facilitating interactions are depicted as green arrows, while insulating interactions are in red. (A) Architectural proteins bound to regulatory elements promote interactions between the regulatory sequences. In addition, polymer simulation studies have suggested that a single architectural protein interaction affects neighboring interactions as well (denoted by *). (B) Regulatory elements within chromatin loops formed by architectural protein interactions may interact more frequently. (C) By reducing the linear genomic distance between loci flanking a chromatin loop, architectural proteins may promote their interactions. (D) Regulatory elements within chromatin loops are insulated from interactions with elements outside the chromatin loops.

neighboring chromatin interactions (Fig. 1D) [41]. Together, this data suggests that for a given architectural protein interaction, neighboring chromatin interactions will be either facilitated or inhibited depending on their location with respect to the chromatin loop [41].

The widespread distribution of architectural proteins in the genome raises the question of what regulates which architectural protein binding sites interact to form chromatin loops (Fig. 2). Of the nearly 10000 long range chromatin interactions mapped in human cells, only 68% of the loci anchoring these contacts interact with only one other locus [32]. Thus, nearly one third of the loci involved in long range interactions interact with multiple distant loci [32]. In *Drosophila*, one factor that determines which loci interact in 3D space is the number of architectural proteins present at specific sites in the genome and the distance between those sites. More specifically, highly occupied sites containing more than 8 architectural proteins appear to mediate interactions at a longer distance and form more stable loops that interfere with interactions between sequences inside and outside of the loops [3,5,6,10,52,53] (Fig. 3A). In mammals the regulation appears to be more complex, and recent evidence suggests that the orientation of the CTCF protein in the genome could be one critical mechanism regulating chromatin interactions. Two independent reports have shown a striking correlation between the orientation of CTCF motifs in the genome and the ability of these sites to interact. The anchors of the nearly 10,000 mapped chromatin loops in humans strongly correlate with binding of architectural proteins, including CTCF, Smc1, Rad21 and ZNF143 [32]. A subset of these chromatin loops (4322) contain CTCF at both loop anchors, with 92% of loop anchors containing the CTCF motifs in a convergent head-to-head orientation [32]. The distinct orientation of CTCF motifs was also observed at TAD borders (Fig. 3A). The 5' end of TAD domains were shown to contain a CTCF motif on the minus strand, while the 3' border contains the CTCF motif on the plus strand in mice, monkeys, and dogs, supporting the hypothesis that CTCF orientation is a critical regulator of long range chromatin interactions in mammals [54]. Hadjir and colleagues speculate that the specific orientation of CTCF motifs also uniquely orients cohesin binding because Rad21 specifically interacts with the C-terminus of CTCF [54,55]. Interestingly, the distinct orientation and spacing of cohesin/CTCF peaks has been characterized previously in primary mouse liver cells [56]. Together, these data suggest that the orientation of the CTCF protein in the genome as well as the number of

architectural proteins bound are some of the mechanisms regulating which loci form long range chromatin interactions.

4. Architectural proteins mark the boundaries between TADs

Individual chromosomes are compartmentalized into TADs (880 kb and 107 kb median size in mouse and *Drosophila*, respectively) separated by regions called TAD borders or sites on contact insulation, which are depleted of interactions between adjacent sequences on either side of the border (Fig. 3A) [3,6,23]. FISH analysis demonstrated that loci spanning the same linear genomic distance are more likely to interact in 3D space if they are located in the same TAD rather than in different TADs, indicative that TADs represent regions of highly interacting chromatin [7]. TAD borders are enriched in actively transcribed genes, especially housekeeping genes, and architectural proteins (Fig. 3A) [3–7]. In general, TAD borders do not appear to separate regions of the genome with different chromatin environments, as defined by the presence of specific histone modifications (Fig. 3B). A single TAD often contains several chromatin types and TAD borders do not change after complete loss of H3K9me2 or H3K27me3 chromatin environments [6,7]. Instead, there is evidence suggesting that TADs represent a compartmentalization of regulatory elements (Fig. 3B). For example deletion of the *Xist-Tsix* or *Epha4* TAD borders resulted in novel chromatin interactions and thus, a partial merging of the two neighboring TADs and altered gene expression [2,7]. Furthermore, ectopic interactions between the enhancers and genes outside their normal TAD were detected by 4C analysis, consistent with the aberrant expression profiles [2]. Of particular interest, the ectopic interactions did not extend past the next proximal TAD border, suggesting that TAD structure partially contributes to gene regulation [2,7]. Thus, genes within a TAD appear to have a partially shared regulatory environment by interacting with similar regulatory elements. Consistent with this hypothesis, analysis of gene expression profiles during cellular differentiation demonstrated that coordinated gene expression was most correlative when analyzing genes within a single TAD [7,57].

The strength of TAD borders, defined as the ratio between intra- and inter-TAD interactions around border sequences, correlates with the number of architectural proteins bound: sites with few architectural proteins are weak borders, while sites with many architectural proteins are strong borders (Fig. 3A). Border strength also correlates with the ability of these sequences to function as classical insulators in transgene reporter experiments [10]. A potential role for architectural proteins in the formation of TAD borders is supported by a variety of experimental results. Depletion of Rad21 or CTCF results in increased inter-TAD interactions, consistent with reduced TAD border strength [23–25,53]. Similarly, the redistribution of architectural proteins from TAD borders to loci within TADs is accompanied by a concomitant decrease in TAD border strength [53]. A recent study evaluating inversions and deletions affecting TAD borders found in patients with limb malformations provided additional evidence that architectural proteins at least partially regulate chromatin interactions at TAD borders. Only deletions of TAD borders containing multiple occupied CTCF sites resulted in aberrant gene expression, indicating that the altered chromatin interactions upon deletion were not merely caused by increased proximity between regulatory elements but being regulated by architectural protein occupancy [2]. Together, these observations suggest a strong association between architectural protein occupancy and TAD borders, implicating that architectural proteins may regulate TAD structure.

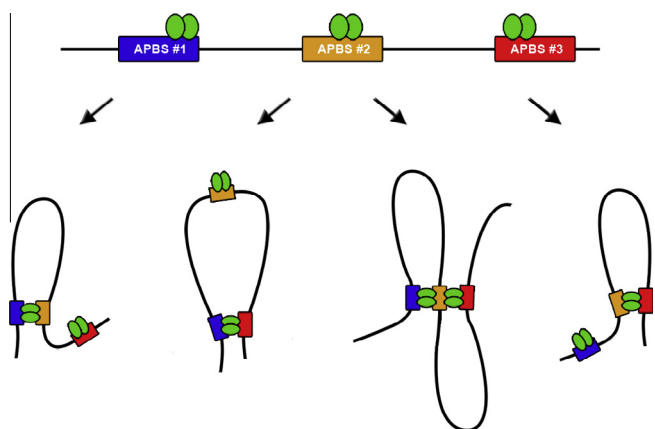


Fig. 2. Multiple architectural protein-bound loci can interact in 3D space. Three distinct architectural-protein bound loci can form multiple different types of interactions and chromatin loops. The mechanisms regulating which architectural protein-bound loci interact are still poorly understood. Architectural proteins are shown in green. (APBS: architectural protein binding site).

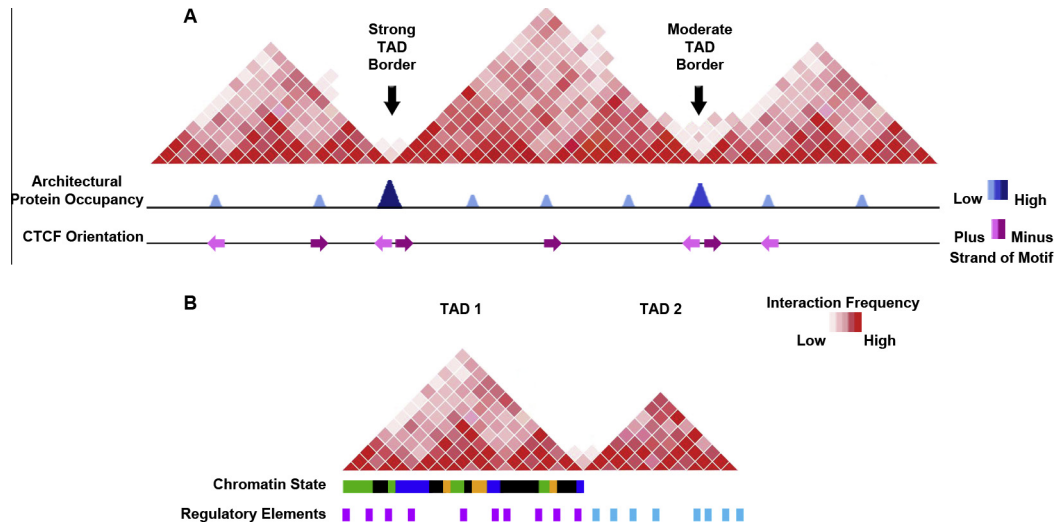


Fig. 3. Architectural proteins mark TAD borders. Cartoon schematics depicting regions of highly associating chromatin called TADs, which are separated by TAD borders. Interaction frequency is shown as a continuum from white to dark red. (A) The strength of a TAD border, defined as the ratio between intra- and inter-TAD interactions around border sequences, depends on architectural protein occupancy (occupancy shown by peak size and a continuum from light to dark blue). A specific orientation of the CTCF motif is observed at TAD borders (purple arrows). (B) Individual TADs contain many different chromatin types depicted in green, blue, black and gold. Evidence suggests that TADs represent compartments for regulatory elements. Interacting regulatory elements in the two distinct TADs are shown in purple and cyan.

5. Architectural proteins regulate chromatin interactions and gene expression within TADs

Additional compartmentalized regulatory regions have been characterized within TADs. Intra-TAD regions of enriched chromatin interactions encompassing ~100–200 kb in mammals have been named subTADs or contact domains [29,32]. The presence of subTADs is consistent with the observation that genes within the same TAD can exhibit different chromatin states and expression levels. Further evidence that intra-TAD contacts represent interactions regulating gene expression comes from cell differentiation studies. During cellular differentiation, chromatin interactions within TADs change significantly (Fig. 4A) [7,29,58]. More precisely, 893/1062 of mouse cortex TAD borders were conserved

in mouse embryonic stem cells, while only 260/425 of chromatin interactions were conserved between neural progenitor cells and embryonic stem cells in mice [3,29]. Furthermore, a subTAD at the *Evx2-HoxD* locus was strongly reduced after activation of the *HoxD* genes by retinoic acid, suggesting that dynamic subTAD structure and intra-TAD interactions are representative of unique gene expression profiles [58].

A number of depletion studies have provided evidence suggesting that architectural proteins facilitate intra-TAD interactions. Depletion of the architectural proteins Rad21 or CTCF results in a striking loss of intra-TAD contacts, implicating architectural proteins as positive regulators of chromatin interactions (Fig. 4A) [23–25,27,53]. Furthermore, these architectural protein-mediated interactions are functionally significant in gene expression and likely represent cell-type specific enhancer–promoter interactions responsible for differential gene expression. The loss of CTCF and Rad21-dependent chromatin interactions results in the misregulation of hundreds of genes, with a reduction in the expression of highly active genes and activation of poorly expressed genes [24,25,59]. Intriguingly, analysis of thymocytes after Rad21 depletion shows that nearly 50% of the misregulated genes overlap or are in close proximity to conventional enhancers or super-enhancers [59]. Misregulation of genes in close proximity to enhancers could be the result of reduced facilitating activity of architectural proteins. For example, in the absence of Rad21, enhancers may not be properly tethered to their normal gene targets, resulting in promiscuous interactions with neighboring genes. Together, these studies provide strong correlative evidence that architectural proteins are critical regulators of chromatin interactions within TADs to fine-tune gene expression and generate a robust range of transcriptional outputs (Fig. 4B).

Despite the apparent redundant function of architectural proteins in the establishment of TADs, individual architectural proteins appear to have distinct roles in regulating interactions within TADs. For example, depletion of CTCF or Rad21 exhibit many similarities but result in the misregulation of unique subsets of genes [24]. Additional evidence for distinct functions of architectural proteins comes from studies in mouse embryonic and neural progenitor cells. The unique combinations of CTCF, SMC1 and/or mediator bound at the anchors of chromatin interactions influence the genomic distance spanned by these interactions [29,60].

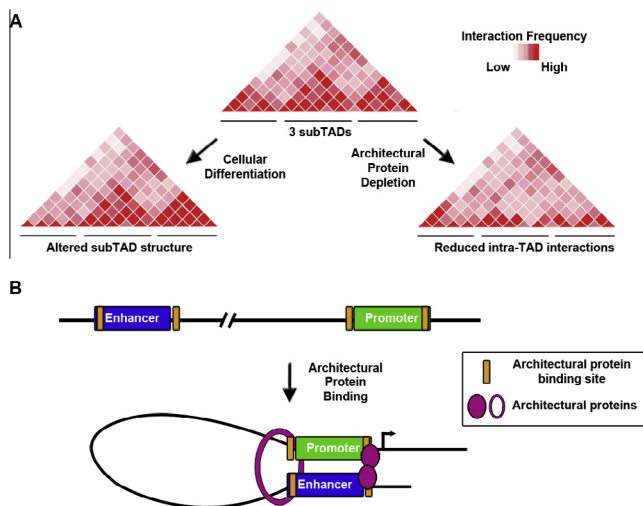


Fig. 4. Architectural proteins regulate chromatin interactions within TADs. (A) Cartoon schematic depicting the regions of enriched chromatin interactions within TADs called subTADs. Intra-TAD interactions change during cellular differentiation and are strongly reduced after architectural protein depletion. Interaction frequency is shown as a continuum from white to dark red. (B) Cartoon schematic illustrating how architectural proteins bring distant enhancers and promoters together in 3D space.

Furthermore, the architectural protein CAP-H2 has been shown to antagonize, rather than facilitate, enhancer–promoter interactions. Comparative analyses of Cap-H2 and Rad21 depletion in *Drosophila* show that CAP-H2 antagonizes while Rad21 promotes long range interactions, consistent with architectural proteins having distinct functions [53]. Overall, these results suggest that the identity of the architectural proteins bound to an intra-TAD locus may influence the type of long range interactions formed.

6. Architectural proteins regulate enhancer–promoter interactions

The molecular mechanisms governing which promoter–enhancer pairs interact in 3D space have been somewhat elusive. Only 7% of enhancer–promoter loops occur between a specific regulatory element and the closest transcriptional start site (TSS) [61]. Furthermore, each gene promoter has been estimated to contact an average of 4.75 enhancers, while 25% of enhancers were assigned to 2 or more promoters, highlighting the complexity of long range 3D interactions [60]. Recent studies have demonstrated a strong correlation between cohesin and CTCF occupancy and long range enhancer–promoter interactions [62]. ChIP-seq analysis has shown that the two cohesin subunits, Rad21 and Smc1A, as well as CTCF, exhibit higher occupancy within conventional and super-enhancers than in the neighboring genomic regions [59]. Furthermore, ChIA-PET analysis has shown that many long range chromatin interactions are anchored by enhancer–promoter pairs bound by both CTCF and cohesin, indicating a possible role for CTCF and cohesin in mediating these interactions [26,27].

Architectural proteins are associated with enhancer–promoter interactions that result in full transcription activation. Transcription factor clusters are found throughout the genome at promoter-proximal sites and distal regulatory regions and are highly correlated with cohesin occupancy [56,63]. Consistent with the different functions of architectural proteins observed in subTAD interactions, it is clear that CTCF is preferentially absent from a subset of active enhancer–promoter interactions. ChIP-seq analysis demonstrated that transcription factor hotspots contain the architectural proteins cohesin, mediator and nipbl but lack CTCF [28]. Indicative of cell-type specific gene expression profiles, only actively transcribed genes exhibit enhancer–promoter occupancy of cohesin, mediator and nipbl and long range chromatin interactions as measured by 3C [28]. In contrast, sites bound by cohesin, mediator and CTCF were more conserved between cells types, indicative of housekeeping genes or poised long range interactions [28,64]. Whether CTCF or other uncharacterized architectural proteins are truly excluded from transcription factor hotspots or are selectively released after full transcriptional activation has yet to be explored further.

Correlations between architectural protein occupancy and transcriptionally unproductive interactions between enhancers and promoters at paused genes have also been observed. For example, BEAF-32 and CP190 are enriched at sites of paused RNA polymerase II, indicative that architectural proteins can form long range interactions prior to transcription elongation [65]. It is well established that interactions between enhancers and promoters form prior to gene activation and can be maintained even after blocking transcription with pharmacological agents [35,66–70]. Architectural proteins have been implicated in regulating poised enhancer–promoter interactions during *Drosophila* embryonic development. Most interactions observed in 3–4 and 6–8 h embryos represent poised enhancer–promoter contacts because only 6% of interactions change between these two time points despite distinct changes in transcription [71]. Strikingly, nearly 50% of the poised interactions present in 6–8 h embryos in

Drosophila were bound by the architectural proteins BEAF-32, CP190, CTCF, GAF, Mod(mdg4) and Su(Hw) [71]. Notably, the 50% correlation is likely an underestimate of occupancy because many architectural proteins were not investigated, including cohesin. Overall, this data suggests that architectural proteins can facilitate enhancer–promoter interactions in the process of transcription initiation, although additional events must be required to release RNA Polymerase II into productive elongation.

Of particular interest, architectural proteins are also associated with genomic elements other than enhancers and promoters. In fact, thirty-three percent of the long range interactions mediated by cohesin and CTCF were bound to genomic loci without enhancers or promoters [26]. This observation can be explained by one of three possibilities. First, the occupancy of CTCF or cohesin at these sites could be unrelated to gene expression such as sites involved in V(D)J or sister chromatid cohesion, respectively [72,73]. Secondly, the co-occupied cohesin–CTCF sites are bound to enhancers that have not been properly annotated yet. Due to the cell-type specificity of enhancer activity, the global mapping of regulatory regions is still likely far from complete. Finally, it is also possible that CTCF and cohesin form interactions between other regulatory elements such as Polycomb (Pc) response elements (PREs). Mutational analysis has demonstrated that long range *Mcp* PRE-mediated interactions do not require the PRE sequence but rather nearby architectural protein binding sites occupied by CTCF and CP190 [74]. Furthermore, depletion of the architectural proteins CTCF, CP190 or cohesin reduced or ablated long range *Mcp*-interactions, supporting a role for architectural proteins in mediating Pc interactions [75]. In addition, interacting architectural protein-bound loci have been shown to flank super-enhancers and Pc-repressed genes, forming super-enhancer and Polycomb domains [26]. Clustered regularly interspaced short palindromic repeat (CRISPR) genome-editing was utilized to delete a CTCF motif at one domain border, resulting in the misregulation of genes within and outside of the domains [26]. Thus, these CTCF sites are still important regulators of gene expression, despite not directly involving enhancers and promoters.

7. Regulation of architectural proteins

The mechanisms that regulate the genomic occupancy of architectural proteins are still poorly defined. Even though most architectural proteins are ubiquitously expressed, recent results suggest that the NIPB promoter may interact with enhancers typically found associated with housekeeping genes as well as cell type specific enhancers. Zabidi and colleagues utilized STARR-seq to identify all of the enhancers that interact with a housekeeping gene promoter compared to a developmentally regulated promoter, demonstrating a clear distinction in the nature of enhancers activating these two classes of genes [76]. Analysis of four housekeeping gene promoters and three developmentally regulated promoters suggests a clear pattern of enhancer interactions with the two distinct core promoter classes. Intriguingly, the NIPB promoter exhibited contacts with housekeeping and developmentally regulated enhancers, suggesting that Nipbl expression may be partially modulated in a cell type specific manner [76]. Thus, it is interesting to speculate that other architectural protein genes may contain a hybrid class of core promoter like NIPB and thus, exhibit a partially cell-type specific expression pattern which could contribute to cell-type specific chromatin interactions.

Another potential mechanism to regulate architectural protein function involves the modulation of DNA binding. In mammals, CTCF binding is negatively regulated by methylation of its DNA recognition sequence [77]. Thus, cell-type specific methylation can result in cell-type specific CTCF occupancy [78]. In addition,

different classes of regulatory elements may recruit distinct members of the architectural protein family based on their intrinsic sequences. For example, *Drosophila* enhancers shown to interact with housekeeping gene core promoters are enriched for the DRE motif, which can act as a binding site for the architectural proteins DREF or BEAF-32 [9,76]. Thus, DREF and BEAF-32 may be important regulators of enhancer–promoter interactions between housekeeping genes but not developmentally regulated genes.

Because functional chromatin interactions depend on DNA-binding proteins and the recruitment of accessory architectural proteins, altering the composition of architectural protein complexes can affect their function. This process may be controlled by posttranslational modifications that regulate the interaction of architectural proteins with DNA or with each other. For example, CP190, Mod(mdg4), and CTCF can be sumoylated, which regulates the formation of insulator bodies in *Drosophila* [79,80]. CTCF sumoylation has also been detected in mammals, but how the modification affects its ability to interact with DNA or other proteins is still unclear [81]. In addition, CTCF has been shown to be modified by polyADP ribosylation, which regulates DNA binding of CTCF in mammals and interactions between CTCF and CP190 in *Drosophila* [82–84]. Thus, differential covalent modifications of architectural proteins may generate unique protein subcomplexes important in regulating 3D interactions.

The presence of lncRNAs in architectural protein complexes could represent another mechanism to regulate the composition of protein complexes formed at various genomic loci. The protein–protein interactions of CTCF as well as CP190 have been shown to be dependent on the presence of ncRNAs [85,86]. In fact, interactions between *Tsix* and *Xite* RNAs are critical for targeting CTCF to the X chromosome, suggesting a role for RNA in recruiting architectural proteins to the proper genomic loci and protein complexes [87]. Expanding the identification of RNAs present in architectural protein complexes will be important to gain insight into the role of RNA in regulating long range architectural protein interactions.

8. Conclusions and perspectives

The data discussed in this review supports the hypothesis that architectural proteins play an important role in the establishment of interphase chromatin organization. These proteins mediate long range chromatin interactions to form loops that facilitate communication between some regulatory elements while reducing contacts between others. We propose a model that architectural proteins contribute to the formation of TAD borders and mediate chromatin interactions between enhancer, promoters, and PREs within TADs [53].

The mechanisms regulating the specificity of long range chromatin interactions between architectural protein-bound genomic loci are just beginning to be characterized. CTCF orientation is likely one important mechanism regulating long range chromatin interactions in mammals but is observed at less than half of the chromatin loops mapped in humans [32]. Thus, it will be interesting to determine if other architectural proteins also exhibit distinct orientations in the genome with respect to the long range chromatin interactions they mediate. Furthermore, the number of the architectural proteins bound to a genomic locus promotes specific types of long range chromatin interactions, raising the question what are high occupancy architectural protein sites interacting with to hinder chromatin interactions across TAD borders? One possibility is that TAD borders interact with each other in 3D space, looping out DNA in a rosette structure to prevent inter-TAD contacts. The 10–30 insulator bodies observed in *Drosophila* cells would then be expected to represent clusters of TAD borders from

the four chromosomes [6,9,45]. Finally, there is correlative evidence that the identity of the architectural proteins bound to a genomic locus regulate the long range chromatin interactions. Thus, future studies better characterizing distinct multimeric architectural protein complexes including the PTMs and ncRNAs present could enhance our understanding about architectural proteins promote distinct long range chromatin interactions. By refining our understanding of how architectural proteins interact in 3D space, we will gain great insight into the mechanisms that allow architectural proteins to both facilitate and hinder interactions between regulatory elements genome-wide.

References

- [1] Pombo, A. and Dillon, N. (2015) Three-dimensional genome architecture: players and mechanisms. *Nat. Rev. Mol. Cell Biol.* 16, 245–257.
- [2] Lupiáñez, Darío G., Kraft, K., Heinrich, V., Krawitz, P., Brancati, F., Klopocki, E., Horn, D., Kayserili, H., Opitz, John M., Laxova, R., Santos-Simarro, F., Gilbert-Dussardier, B., Wittler, L., Borschiwer, M., Haas, Stefan A., Osterwalder, M., Franke, M., Timmermann, B., Hecht, J., Spielmann, M., Visel, A. and Mundlos, S. (2015) Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161, 1012–1025.
- [3] Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S. and Ren, B. (2012) Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485, 376–380.
- [4] Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragozy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S. and Dekker, J. (2009) Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293.
- [5] Sexton, T., Yaffe, E., Kenigsberg, E., Bantignies, F., Leblanc, B., Hoichman, M., Parrinello, H., Tanay, A. and Cavalli, G. (2012) Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* 148, 458–472.
- [6] Hou, C., Li, L., Qin, Z.S. and Corces, V.G. (2012) Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Mol. Cell* 48, 471–484.
- [7] Nora, E.P., Lajoie, B.R., Schulz, E.G., Giorgetti, L., Okamoto, I., Servant, N., Piolot, T., van Berkum, N.L., Meisig, J., Sedat, J., Gribnau, J., Barillot, E., Bluthgen, N., Dekker, J. and Heard, E. (2012) Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485, 381–385.
- [8] Bushey, A.M., Ramos, E. and Corces, V.G. (2009) Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes Dev.* 23, 1338–1350.
- [9] Gurudatta, B.V., Yang, J., Van Bortle, K., Donlin-Asp, P.G. and Corces, V.G. (2013) Dynamic changes in the genomic localization of DNA replication-related element binding factor during the cell cycle. *Cell Cycle* 12, 1605–1615.
- [10] Van Bortle, K., Nichols, M.H., Li, L., Ong, C.T., Takenaka, N., Qin, Z.S. and Corces, V.G. (2014) Insulator function and topological domain border strength scale with architectural protein occupancy. *Genome Biol.* 15, R82.
- [11] Gerasimova, T.I., Lei, E.P., Bushey, A.M. and Corces, V.G. (2007) Coordinated control of dCTCF and gypsy chromatin insulators in *Drosophila*. *Mol. Cell* 28, 761–772.
- [12] Yang, J., Ramos, E. and Corces, V.G. (2012) The BEAF-32 insulator coordinates genome organization and function during the evolution of *Drosophila* species. *Genome Res.* 22, 2199–2207.
- [13] Aoki, T., Sarkeshik, A., Yates, J. and Schedl, P. (2012) Elba, a novel developmentally regulated chromatin boundary factor is a hetero-tripartite DNA binding complex. *eLife* 1, e00171.
- [14] Maksimenko, O., Bartkuhn, M., Stakhov, V., Herold, M., Zolotarev, N., Jox, T., Buxa, M.K., Kirsch, R., Bonchuk, A., Fedotova, A., Kyrchanova, O., Renkawitz, R. and Georgiev, P. (2015) Two new insulator proteins, Pita and ZIPIC, target CP190 to chromatin. *Genome Res.* 25, 89–99.
- [15] Cuartero, S., Fresan, U., Reina, O., Planet, E. and Espinas, M.L. (2014) Ibf1 and Ibf2 are novel CP190-interacting proteins required for insulator function. *EMBO J.* 33, 637–647.
- [16] Kellner, W.A., Van Bortle, K., Li, L., Ramos, E., Takenaka, N. and Corces, V.G. (2013) Distinct isoforms of the *Drosophila* Brd4 homologue are present at enhancers, promoters and insulator sites. *Nucleic Acids Res.* 41, 9274–9283.
- [17] Moshkovich, N., Nisha, P., Boyle, P.J., Thompson, B.A., Dale, R.K. and Lei, E.P. (2011) RNAi-independent role for Argonaute2 in CTCF/CP190 chromatin insulator function. *Genes Dev.* 25, 1686–1701.
- [18] Vogelmann, J., Le Gall, A., Dejardin, S., Allemand, F., Gamot, A., Labesse, G., Cuvier, O., Negre, N., Cohen-Gonsaud, M., Margeat, E. and Nollmann, M. (2014) Chromatin insulator factors involved in long-range DNA interactions and their role in the folding of the *Drosophila* genome. *PLoS Genet.* 10, e1004544.
- [19] Bell, A.C., West, A.G. and Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98, 387–396.
- [20] Hadjur, S., Williams, L.M., Ryan, N.K., Cobb, B.S., Sexton, T., Fraser, P., Fisher, A.G. and Merkschlager, M. (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature* 460, 410–413.

- [21] Mishiro, T., Ishihara, K., Hino, S., Tsutsumi, S., Aburatani, H., Shirahige, K., Kinoshita, Y. and Nakao, M. (2009) Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J.* 28, 1234–1245.
- [22] Nativio, R., Wendt, K.S., Ito, Y., Huddleston, J.E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J.M. and Murrell, A. (2009) Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. *PLoS Genet.* 5, e1000739.
- [23] Sofueva, S., Yaffe, E., Chan, W.C., Georgopoulou, D., Vietri Rudan, M., Mira-Bontenbal, H., Pollard, S.M., Schroth, G.P., Tanay, A. and Hadjur, S. (2013) Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* 32, 3119–3129.
- [24] Zuin, J., Dixon, J.R., van der Reijden, M.I., Ye, Z., Kolovos, P., Brouwer, R.W., van de Corput, M.P., van de Werken, H.J., Knoch, T.A., van, I.W.F., Grosveld, F.G., Ren, B. and Wendt, K.S. (2014) Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 111, 996–1001.
- [25] Seitan, V.C., Faure, A.J., Zhan, Y., McCord, R.P., Lajoie, B.R., Ing-Simmons, E., Lenhard, B., Giorgetti, L., Heard, E., Fisher, A.G., Flicek, P., Dekker, J. and Merckenschlager, M. (2013) Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res.* 23, 2066–2077.
- [26] Downen, J.M., Fan, Z.P., Hnisz, D., Ren, G., Abraham, B.J., Zhang, L.N., Weintraub, A.S., Schuijers, J., Lee, T.I., Zhao, K. and Young, R.A. (2014) Control of cell identity genes occurs in insulated neighborhoods in Mammalian chromosomes. *Cell* 159, 374–387.
- [27] Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F., Wong, E., Sheng, J., Zhang, Y., Poh, T., Chan, C.S., Kunarso, G., Shahab, A., Bourque, G., Cacheux-Rataboul, V., Sung, W.K., Ruan, Y. and Wei, C.L. (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* 43, 630–638.
- [28] Kagey, M.H., Newman, J.J., Bilodeau, S., Zhan, Y., Orlando, D.A., van Berkum, N.L., Ebmeier, C.C., Goossens, J., Rahl, P.B., Levine, S.S., Taatjes, D.J., Dekker, J. and Young, R.A. (2010) Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467, 430–435.
- [29] Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y., Bland, M.J., Wagstaff, W., Dalton, S., McDewitt, T.C., Sen, R., Dekker, J., Taylor, J. and Corces, V.G. (2013) Architectural protein subclasses shape 3D organization of genomes during lineage commitment. *Cell* 153, 1281–1295.
- [30] Pena-Hernandez, R., Marques, M., Hilmi, K., Zhao, T., Saad, A., Alaoui-Jamali, M.A., Del Rincon, S.V., Ashworth, T., Roy, A.L., Emerson, B.M. and Witcher, M. (2015) Genome-wide targeting of the epigenetic regulatory protein CTCF to gene promoters by the transcription factor TFII-I. *Proc. Natl. Acad. Sci. U.S.A.* 112, E677–E686.
- [31] Zlatanova, J. and Caiafa, P. (2009) CTCF and its protein partners: divide and rule? *J. Cell Sci.* 122, 1275–1284.
- [32] Rao, S.S., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S. and Aiden, E.L. (2014) A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159, 1665–1680.
- [33] Bailey, S.D., Zhang, X., Desai, K., Aid, M., Corradin, O., Cowper-Sal Lari, R., Akhtar-Zaidi, B., Scacheri, P.C., Haibe-Kains, B. and Lupien, M. (2015) ZNF143 provides sequence specificity to secure chromatin interactions at gene promoters. *Nat. Commun.* 2, 6186.
- [34] Sancho, A., Li, S., Paul, T., Zhang, F., Aguilo, F., Vashisht, A., Balasubramanian, N., Leleiko, N.S., Suchy, F.J., Wohlschlegel, J.A., Zhang, W. and Walsh, M.J. (2015) CHD6 regulates the topological arrangement of the CFTR locus. *Hum. Mol. Genet.*
- [35] Deng, W., Lee, J., Wang, H., Miller, J., Reik, A., Gregory, P.D., Dean, A. and Blobel, G.A. (2012) Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149, 1233–1244.
- [36] Tedeschi, A., Wutz, G., Huet, S., Jaritz, M., Wuensche, A., Schirghuber, E., Davidson, I.F., Tang, W., Cisneros, D.A., Bhaskara, V., Nishiyama, T., Vaziri, A., Wutz, A., Ellenberg, J. and Peters, J.M. (2013) Wapl is an essential regulator of chromatin structure and chromosome segregation. *Nature* 501, 564–568.
- [37] Bulger, M. and Groudine, M. (2011) Functional and mechanistic diversity of distal transcription enhancers. *Cell* 144, 327–339.
- [38] Levine, M. (2010) Transcriptional enhancers in animal development and evolution. *Curr. Biol.* 20, R754–R763.
- [39] Pennacchio, L.A., Bickmore, W., Dean, A., Nobrega, M.A. and Bejerano, G. (2013) Enhancers: five essential questions. *Nat. Rev. Genet.* 14, 288–295.
- [40] Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F. and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell* 10, 1453–1465.
- [41] Doyle, B., Fudenberg, G., Imakaev, M. and Mirny, L.A. (2014) Chromatin loops as allosteric modulators of enhancer–promoter interactions. *PLoS Comput. Biol.* 10, e1003867.
- [42] Mukhopadhyay, S., Schedl, P., Studitsky, V.M. and Sengupta, A.M. (2011) Theoretical analysis of the role of chromatin interactions in long-range action of enhancers and insulators. *Proc. Natl. Acad. Sci. U.S.A.* 108, 19919–19924.
- [43] Ramos, E., Ghosh, D., Baxter, E. and Corces, V.G. (2006) Genomic organization of gypsy chromatin insulators in *Drosophila melanogaster*. *Genetics* 172, 2337–2349.
- [44] Parnell, T.J., Vierung, M.M., Skjesol, A., Helou, C., Kuhn, E.J. and Geyer, P.K. (2003) An endogenous suppressor of hairy-wing insulator separates regulatory domains in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 13436–13441.
- [45] Spana, C., Harrison, D.A. and Corces, V.G. (1988) The *Drosophila melanogaster* suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. *Genes Dev.* 2, 1414–1423.
- [46] Geyer, P.K. and Corces, V.G. (1992) DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes Dev.* 6, 1865–1873.
- [47] Smith, P.A. and Corces, V.G. (1992) The suppressor of Hairy-wing binding region is required for gypsy mutagenesis. *Mol. Gen. Genet.* 233, 65–70.
- [48] Pai, C.Y., Lei, E.P., Ghosh, D. and Corces, V.G. (2004) The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol. Cell* 16, 737–748.
- [49] Gerasimova, T.I., Gdula, D.A., Gerasimov, D.V., Simonova, O. and Corces, V.G. (1995) A *Drosophila* protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell* 82, 587–597.
- [50] Gaszner, M., Vazquez, J. and Schedl, P. (1999) The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer–promoter interaction. *Genes Dev.* 13, 2098–2107.
- [51] Gilbert, M.K., Tan, Y.Y. and Hart, C.M. (2006) The *Drosophila* boundary element-associated factors BEAF-32A and BEAF-32B affect chromatin structure. *Genetics* 173, 1365–1375.
- [52] Phillips-Cremins, J.E. and Corces, V.G. (2013) Chromatin insulators: linking genome organization to cellular function. *Mol. Cell* 50, 461–474.
- [53] Li, L., Lyu, X., Hou, C., Takenaka, N., Nguyen, Huy Q., Ong, C.-T., Cubeñas-Potts, C., Hu, M., Lei, Elissa P., Bosco, G., Qin, Zhaohui S., and Corces, Victor G. (2015) Widespread rearrangement of 3D chromatin organization underlies polycomb-mediated stress-induced silencing. *Mol. Cell.*
- [54] Vietri Rudan, M., Barrington, C., Henderson, S., Ernst, C., Odom, D.T., Tanay, A. and Hadjur, S. (2015) Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* 10, 1297–1309.
- [55] Xiao, T., Wallace, J. and Felsenfeld, G. (2011) Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol. Cell Biol.* 31, 2174–2183.
- [56] Faure, A.J., Schmidt, D., Watt, S., Schwalie, P.C., Wilson, M.D., Xu, H., Ramsay, R.G., Odom, D.T. and Flicek, P. (2012) Cohesin regulates tissue-specific expression by stabilizing highly occupied cis-regulatory modules. *Genome Res.* 22, 2163–2175.
- [57] Arner, E., Daub, C.O., Vitting-Seerup, K., Andersson, R., Lilje, B., Drablos, F., Lernerstsson, A., Ronnerblad, M., Hrydziuszko, O., Vitezic, M., Freeman, T.C., Alhendi, A.M., Arner, P., Axton, R., Baillie, J.K., Beckhouse, A., Bodega, B., Briggs, J., Brombacher, F., Davis, M., Detmar, M., Ehrlund, A., Endoh, M., Eslami, A., Fagiolini, M., Fairbairn, L., Faulkner, G.J., Ferrai, C., Fisher, M.E., Forrester, L., Goldowitz, D., Guler, R., Ha, T., Hara, M., Herlyn, M., Ikawa, T., Kai, C., Kawamoto, H., Khachigian, L.M., Klinken, S.P., Kojima, S., Koseki, H., Klein, S., Mejhert, N., Miyaguchi, K., Mizuno, Y., Morimoto, M., Morris, K.J., Mummery, C., Nakachi, Y., Ogishima, S., Okada-Hatakeyama, M., Okazaki, Y., Orlando, V., Ovchinnikov, D., Passier, R., Patrikakis, M., Pombo, A., Qin, X.Y., Roy, S., Sato, H., Savvi, S., Saxena, A., Schwegmann, A., Sugiyama, D., Swoboda, R., Tanaka, H., Tomoiu, A., Winteringham, L.N., Wolvetang, E., Yanagi-Mizuochi, C., Yoneda, M., Zabierowski, S., Zhang, P., Abugessaisa, I., Bertin, N., Diehl, A.D., Fukuda, S., Furuno, M., Harshbarger, J., Hasegawa, A., Hori, F., Ishikawa-Kato, S., Ishizu, Y., Itoh, M., Kawashima, T., Kojima, M., Kondo, N., Lizio, M., Meehan, T.F., Mungall, C.J., Murata, M., Nishiyori-Sueki, H., Sahin, S., Nagao-Sato, S., Severin, J., de Hoon, M.J., Kawai, J., Kasukawa, T., Lassmann, T., et al. (2015) Gene regulation. Transcribed enhancers lead waves of coordinated transcription in transitioning mammalian cells. *Science* 347, 1010–1014.
- [58] Williamson, I., Berlivet, S., Eskeland, R., Boyle, S., Illingworth, R.S., Paquette, D., Dostie, J. and Bickmore, W.A. (2014) Spatial genome organization: contrasting views from chromosome conformation capture and fluorescence in situ hybridization. *Genes Dev.* 28, 2778–2791.
- [59] Ing-Simmons, E., Seitan, V., Faure, A., Flicek, P., Carroll, T., Dekker, J., Fisher, A., Lenhard, B. and Merckenschlager, M. (2015) Spatial enhancer clustering and regulation of enhancer–proximal genes by cohesin. *Genome Res.*
- [60] Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A. and Ren, B. (2013) A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503, 290–294.
- [61] Sanyal, A., Lajoie, B.R., Jain, G. and Dekker, J. (2012) The long-range interaction landscape of gene promoters. *Nature* 489, 109–113.
- [62] Narendra, V., Rocha, P.P., An, D., Raviram, R., Skok, J.A., Mazzoni, E.O. and Reinberg, D. (2015) Transcription. CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. *Science* 347, 1017–1021.
- [63] Yan, J., Enge, M., Whittington, T., Dave, K., Liu, J., Sur, I., Schmierer, B., Jolma, A., Kivioja, T., Taipale, M. and Taipale, J. (2013) Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* 154, 801–813.
- [64] He, B., Chen, C., Teng, L. and Tan, K. (2014) Global view of enhancer–promoter interactome in human cells. *Proc. Natl. Acad. Sci. U.S.A.* 111, E2191–E2199.
- [65] Liang, J., Lacroix, L., Gamot, A., Cuddapah, S., Queille, S., Lhoumaud, P., Lepetit, P., Martin, P.G., Vogelmann, J., Court, F., Hennion, M., Micas, G., Urbach, S., Bouchez, O., Nollmann, M., Zhao, K., Emberly, E. and Cuvier, O. (2014) Chromatin immunoprecipitation indirect peaks highlight long-range interactions of insulator proteins and Pol II pausing. *Mol. Cell* 53, 672–681.
- [66] Vernimmen, D., De Gobbi, M., Sloane-Stanley, J.A., Wood, W.G. and Higgs, D.R. (2007) Long-range chromosomal interactions regulate the timing of the transition between poised and active gene expression. *EMBO J.* 26, 2041–2051.

- [67] Spilianakis, C.G. and Flavell, R.A. (2004) Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* 5, 1017–1027.
- [68] Palstra, R.J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B. and de Laat, W. (2008) Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. *PLoS ONE* 3, e1661.
- [69] Mitchell, J.A. and Fraser, P. (2008) Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev.* 22, 20–25.
- [70] Kim, Y.W., Lee, S., Yun, J. and Kim, A. (2015) Chromatin looping and eRNA transcription precede the transcriptional activation of gene in the beta-globin locus. *Biosci. Rep.*
- [71] Ghavi-Helm, Y., Klein, F.A., Pakozdi, T., Ciglar, L., Noordermeer, D., Huber, W. and Furlong, E.E. (2014) Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* 512, 96–100.
- [72] Haarhuis, J.H., Elbatsh, A.M. and Rowland, B.D. (2014) Cohesin and its regulation: on the logic of X-shaped chromosomes. *Dev. Cell* 31, 7–18.
- [73] Ong, C.T. and Corces, V.G. (2014) CTCF: an architectural protein bridging genome topology and function. *Nat. Rev. Genet.* 15, 234–246.
- [74] Li, H.B., Muller, M., Bahechar, I.A., Kyrchanova, O., Ohno, K., Georgiev, P. and Pirrotta, V. (2011) Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in *Drosophila melanogaster*. *Mol. Cell. Biol.* 31, 616–625.
- [75] Li, H.B., Ohno, K., Gui, H. and Pirrotta, V. (2013) Insulators target active genes to transcription factories and polycomb-repressed genes to polycomb bodies. *PLoS Genet.* 9, e1003436.
- [76] Zabidi, M.A., Arnold, C.D., Schernhuber, K., Pagani, M., Rath, M., Frank, O. and Stark, A. (2015) Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation. *Nature* 518, 556–559.
- [77] Engel, N., West, A.G., Felsenfeld, G. and Bartolomei, M.S. (2004) Antagonism between DNA hypermethylation and enhancer-blocking activity at the H19 DMD is uncovered by CpG mutations. *Nat. Genet.* 36, 883–888.
- [78] Wang, H., Maurano, M.T., Qu, H., Varley, K.E., Gertz, J., Pauli, F., Lee, K., Canfield, T., Weaver, M., Sandstrom, R., Thurman, R.E., Kaul, R., Myers, R.M. and Stamatoyannopoulos, J.A. (2012) Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* 22, 1680–1688.
- [79] Capelson, M. and Corces, V.G. (2006) SUMO conjugation attenuates the activity of the gypsy chromatin insulator. *EMBO J.* 25, 1906–1914.
- [80] Golovnin, A., Volkov, I. and Georgiev, P. (2012) SUMO conjugation is required for the assembly of *Drosophila* Su(Hw) and Mod(mdg4) into insulator bodies that facilitate insulator complex formation. *J. Cell Sci.* 125, 2064–2074.
- [81] MacPherson, M.J., Beatty, L.G., Zhou, W., Du, M. and Sadowski, P.D. (2009) The CTCF insulator protein is posttranslationally modified by SUMO. *Mol. Cell. Biol.* 29, 714–725.
- [82] Witcher, M. and Emerson, B.M. (2009) Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol. Cell* 34, 271–284.
- [83] Ong, C.T., Van Bortle, K., Ramos, E. and Corces, V.G. (2013) Poly(ADP-ribosyl)ation regulates insulator function and intrachromosomal interactions in *Drosophila*. *Cell* 155, 148–159.
- [84] Yu, W., Ginja, V., Pant, V., Chernukhin, I., Whitehead, J., Docquier, F., Farrar, D., Tavoosidana, G., Mukhopadhyay, R., Kanduri, C., Oshimura, M., Feinberg, A.P., Lobanenko, V., Klenova, E. and Ohlsson, R. (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat. Genet.* 36, 1105–1110.
- [85] Lei, E.P. and Corces, V.G. (2006) RNA interference machinery influences the nuclear organization of a chromatin insulator. *Nat. Genet.* 38, 936–941.
- [86] Yao, H., Brick, K., Evrard, Y., Xiao, T., Camerini-Otero, R.D. and Felsenfeld, G. (2010) Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev.* 24, 2543–2555.
- [87] Kung, J.T., Kesner, B., An, J.Y., Ahn, J.Y., Cifuentes-Rojas, C., Colognori, D., Jeon, Y., Szanto, A., del Rosario, B.C., Pinter, S.F., Erwin, J.A. and Lee, J.T. (2015) Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell* 57, 361–375.