 doi:10.20944/preprints202012.0184.v1

Review

Mechanisms of enhancer-promoter interactions in higher eukaryotes

Olga Kyrchanova[1], Pavel Georgiev[1]

Department of the Control of Genetic Processes, Institute of Gene Biology Russian Academy of Sciences, 34/5 Vavilov St., Moscow 119334, Russia;

Corresponding author: Pavel Georgiev georgiev_p@mail.ru

Keywords: C2H2 proteins, CTCF, LDB1, chromatin insulator, long-distance interactions

Abstract In higher eukaryotes, enhancers determine the activation of developmental gene transcription in specific cell types and stages of embryogenesis. Enhancers transform the signals produced by various transcription factors within a given cell, activating the transcription of the targeted genes. Often, developmental genes can be associated with dozens of enhancers, some of which are located at large distances from the promoters that they regulate. Currently, the mechanisms that underly the specific distance interactions between enhancers and promoters remain unknown. This review describes the properties and activities of enhancers and discusses the mechanisms of distance interactions and potential proteins involved in this process.

1. Introduction

In higher eukaryotes, the regulation of gene expression is complicated as a consequence of cell differentiation during embryonic development [1–3]. Cell specialization is determined by differences in transcription factor (TF) repertoires, and the genes responsible for cell differentiation and organismal development are typically regulated by multiple independent enhancers, each of which stimulates a promoter in a limited population of cells during a specific time interval.

Enhancers were first described nearly 40 years ago when 72-bp tandem sequences from the SV40 virus were found to enhance gene expression when integrated at large distances from the promoter and in any orientation relative to the regulated gene [4,5]. Two years later, the first cellular enhancer was identified [6]. Currently, the human genome is predicted to encode 300,000 enhancers [7].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

Enhancers are regions of DNA, typically 100 to 1,000 bp in size, that contain TF-binding sites that stimulate the initiation and elongation of transcription from promoters [1,8–10]. In most housekeeping genes, enhancers are located in close proximity to promoters, and distinguishing the borders between these regulatory elements can be difficult. In contrast, developmental genes typically feature complex regulatory regions that can consist of dozens of enhancers located at variable distances from the regulated promoter.

During transcription activation, enhancers interact directly with gene promoters [8,11]. Currently, how chromatin loops are formed between enhancers and promoters and which architectural proteins are involved in this process remain unknown. This review provides an overview of the currently available information in mammals and *Drosophila* regarding enhancers and known architectural proteins that have been shown to support distance interactions between regulatory elements.

2. Enhancer properties and functions

Enhancers consist of combinations of short, comparative, degenerate sites, 6-12 bp in length, that are recognized by DNA-binding TFs, which determine enhancer activity [1,9]. The combination of DNA-binding TFs on a given enhancer creates a platform that attracts co-activators and co-repressors that determine the enhancer activity in each specific group of cells. In some cases, a single TF can act as both a positive and a negative regulator of the same enhancer, depending on the presence of other TFs. Thus, the ability of an enhancer to stimulate transcription depends on the combination of TF sites that positively or negatively affect enhancer activity and the relative concentrations of enhancer-binding TFs within the nuclei of a given group of cells.

Recently, super-enhancers have been identified, representing a special class of regulatory elements, characterized by large sizes, sometimes reaching tens of thousands of bp, with a high degree of TF and co-activator enrichment [12,13]. Super-enhancers are often located adjacent to genes known to be critical for cell differentiation [14]. A more detailed study of super-enhancers has shown that they often consist of separate domains that can either function together to enhance the overall activity of each domain or play independent roles during the simultaneous activation of a large number of promoters [15–17].

During the process of transcription activation (Figure 1), enhancers attract complexes that possess acetyltransferase (p300/CBP) and methyltransferase (Mll3/Mll4/COMPASS) activities

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

[18]. The proteins Mll3 and Mll4 both contain a C-terminal SET (suppressor of variegation, enhancer of zeste, trithorax) domain, which is responsible for the monomethylation of lysine 4 of histone H3 (H3K4me1) [19,20]. The complexes formed by Mll3 and Mll4 have partially overlapping and insufficiently studied functions in the regulation of enhancer activity. Mll3 and Mll4 are also known to be involved in the recruitment of the p300/CBP co-activator, which is responsible for the acetylation of histone H3 at lysine 27 (H3K27ac) [21]. H3K27ac and H3K4me1 histone marks are distinctive features of active enhancers and are used to identify enhancers in genomes [22–26]. However, although mutations that inactivate the methyltransferase activity of the SET domains in the Mll3 and Mll4 proteins significantly reduce H3K4me1 quantities, they do not have strong effects on the functional activities of enhancers [27,28].

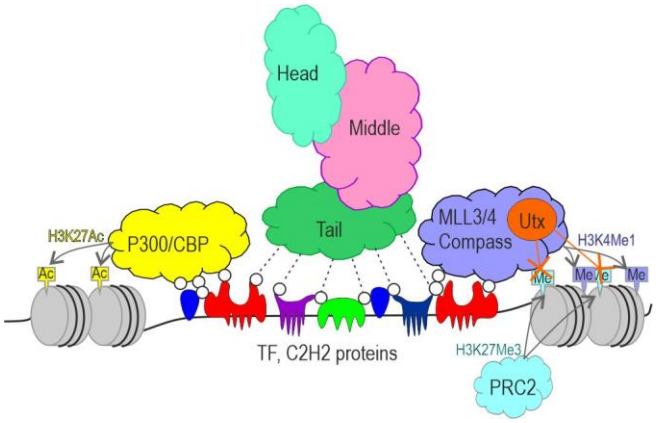


Figure 1. Schematic representation of the transcriptional complexes involved in enhancer activity. Various transcription factors (TFs) bind to enhancers sites and recruit complexes involved in transcription stimulation. p300/CBP possesses acetyltransferase activity and is responsible for H3K27 acetylation. The Mll3/4 complexes induce H3K4 monomethylation and recruit the UTX demethylase, which can remove the H3K27me3 associated with repressed chromatin. p300/CBP, Mll3/4, and UTX are thought to regulate transcription and enhancer activity through the modification of currently unknown components of primary transcriptional complexes. The subunits of the mediator complex form three main modules: head, middle, and tail. The mediator complex is recruited to the enhancer via multiple interactions between subunits of the tail module and the intrinsically disordered regions of TFs.

The H3K27ac and H3K4me1 modifications of histone H3 are thought to reduce the stability of nucleosomes [29], resulting in the formation of open chromatin and the stabilization

 doi:10.20944/preprints202012.0184.v1

of TF binding to enhancers. The Mll3/4 and p300/CBP complexes antagonize two Polycomb gene (PcG) complexes, Polycomb-repressive complexes 1 and 2 (PRC1 and PRC2, respectively), which are involved in the repression of enhancers and promoters [30]. PRC1 and PRC2 have ubiquitin transferase and methyltransferase activities, respectively, resulting in transcriptional repression [31]. The best-known activity of the PRC2 complex is the trimethylation of lysine 27 in histone H3 (H3K27me3), which is a characteristic marker of transcriptionally repressed chromatin regions. The Mll3 and Mll4 complexes are associated with UTX demethylase, which can remove the PcG-deposited H3K27me3 [18], allowing the lysine 27 in H3 to be acetylated by p300/CBP, preventing the trimethylation (H3K27me3) by PRC2 and stabilizing the active chromatin in the enhancer region. In addition to the acetylation of H3K27, p300/CBP may play an important but poorly understood role in the acetylation of transcription factors and serve as a component of the pre-initiation complex during transcription activation [32,33].

A highly conserved mediator complex, consisting of 25 subunits in yeast and 30 subunits in humans, plays a primary role in the enhancer-mediated activation of transcription [34]. The subunits found in the mediator complex form three main modules: head, middle, and tail. The head and middle modules determine the primary functions of the complex during transcription activation, whereas the tail module is responsible for interaction with the TFs bound to enhancers and gene promoters [35–39]. Together with general transcription factors assembled on the promoter, the mediator complex facilitates the assembly of a pre-initiation complex and is involved in the recruitment of RNA polymerase II to promoters [40,41]. However, many of the processes that involve the mediator complex during the activation of transcription remain insufficiently studied [42].

3. Distance interactions between enhancers and promoters

Most enhancers are located in close proximity to the promoters that they activate. However, even in the compact *Drosophila* genome, approximately 20% of enhancers are located at a distance of 50–100 kb from their associated promoters and can often be isolated by other genes [43,44]. In mammals, the distances between enhancers and promoters are typically even larger [8].

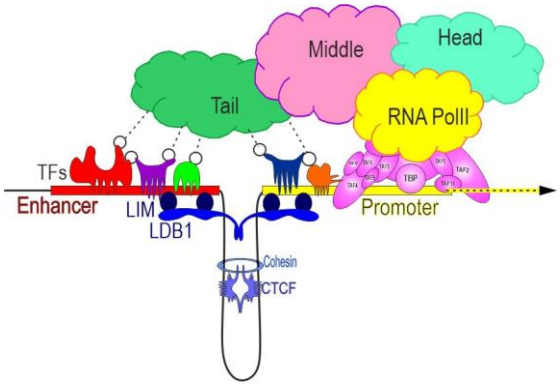
Many distance interactions between enhancers and promoters have been shown to be stable throughout *Drosophila* embryogenesis [44]. Multiple interactions have been observed between various enhancers and promoters with similar expression profiles. Similar results were obtained in mammalian cells, in which stable promoter-promoter and promoter-enhancer

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

interactions were also identified [45,46]. Interactions between enhancers and promoters can be formed both before and simultaneous with transcription activation [44,47,48]. Super-enhancers have been shown to interact with a large number of promoters during cell differentiation [49]. Recently, the mediator complex was found not to be involved in the organization of distance interactions between enhancers and promoters [50], which is consistent with the finding that the mediator interacts only transiently with an enhancer [51].

The study of *β-globin* gene activation by a strong locus control region (LCR) enhancer in real time showed that enhancers activate transcription by increasing the frequency of intermittent transcriptional bursts [52]. Improving the stability of the interaction between the LCR and the β-globin gene promoter resulted in the increased frequency of transcription bursts. A similar study was performed on reporter genes in living transgenic *Drosophila* embryos [52], which revealed that strong and weak enhancers induce transcription bursts of equal strength, but the enhancer strength determines the burst frequency. Moreover, one enhancer can simultaneously, and in a coordinated manner, activate the promoters of two reporter genes but with different activation amplitudes [53].

These results are consistent with a model in which the stable interaction between an enhancer and one or more promoters allows the TFs bound to these regulatory elements to form a platform to which the mediator complex and other transcriptional co-activators can transiently bind (Figure 2). The unstructured acidic domains of TFs can interact with various subunits of the mediator complex, most of which are located in the tail module [54,55]. The efficiency of mediator recruitment is directly correlated with the number of TFs associated with enhancer-promoter sequences. During the transient binding with the enhancer-promoter complex, the mediator manages to induce only a short transcriptional pulse. The strength of transcription is determined by the burst frequency, which is directly dependent on the number of TFs that attract the mediator to the protein platform formed by the interaction between the enhancer and promoter.



Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

Figure 2. Model showing the remote activation of a promoter by an enhancer. The interaction between an enhancer and a promoter can be stabilized by CCCTC-binding factor (CTCF), in cooperation with the cohesin complex and the dimerization of LOB domain-containing protein 1 (LBD1), which is recruited to chromatin through interactions with LIM proteins. The TFs bind to the enhancer and a promoter to form a platform for the transient recruitment of the mediator complex. The mediator complex transfers RNA polymerase II to the promoter transcription factor IID (TFIID) complex and accelerates further transcription initiation steps to induce a short transcriptional pulse (burst). The enhancer strength is directly correlated with the efficiency of mediator recruitment to chromatin.

4. Distance interaction mechanisms between enhancers and promoters in mammalian genomes

The genome is spatially organized into compartments and topologically associated domains (TADs), which are important for proper enhancer-promoter interactions in the nucleus [56]. All eukaryotic chromosomes are organized into TADs, which are formed by boundaries, and contacts between TADs occur with relatively high frequency [57–59]. TAD boundaries are defined based on preferred interactions, and no known physical barriers restrict trans-interactions between regulatory elements located in neighboring TADs [57,59–61]. However, the TAD architecture can facilitate interactions between regulatory elements located within the TAD by reducing the physical distance between these elements.

In mammals, most TAD boundaries contain binding sites for CCCTC-binding factor (CTCF) [62]. A characteristic feature of CTCF is the presence of a cluster, consisting of 11 C2H2-type zinc finger domains [63]. Five C2H2 domains in CTCF specifically bind to an extended DNA motif that is conserved in most animals [64]. The N-termini of CTCF homologs from representative bilaterian species feature unstructured domains that are capable of homodimerization [65]. A motif that interacts with the cohesin SA2-SCC1 sub-complex was identified between the N-terminal homodimerization domain and the C2H2 cluster [66].

Together with the cohesin complex, CTCF defines the boundaries of most TADs [11,59,67]. The CTCF/cohesin complex is also involved in organizing interactions between enhancers and promoters within TADs [68,69]. CTCF inactivation resulted in the re-localization of cohesin complexes from the CTCF binding sites to the promoters of active genes, which was

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

accompanied by the partial disruption of the TADs [70]. Mutations in CTCF that disrupt the interaction with the cohesin complex also result in the loss of some distance interactions and a decrease in the efficiency of TAD formation [66,71].

Only a small fraction of CTCF sites located in the opposite convergent orientation are involved in TAD organization [72,73]. To explain the preferable formation of chromatin loops between CTCF sites located in a convergent orientation, a model was proposed in which the cohesin complex binds to the chromosome and initiates the extrusion of DNA, resulting in the formation of a chromatin loop. CTCF can block the progression of the cohesin complex when the interaction between its N-terminal domain and the SA2-SCC1 sub-complex [66] is oriented correctly relative to the moving cohesin complex. The dimerization between the N-terminal domains of the CTCF protein is thought to stabilize the formation of chromatin loops [65].

The inactivation of either cohesin or CTCF typically results in only the partial disruption of chromatin loops [70,74], and some interactions between enhancers and promoters have been found to occur independently of cohesin and CTCF [75,76]. The existence of other architectural proteins has been proposed, which may also be involved in the organization of distance interactions, together with CTCF. To date, only a few proteins have been described that can potentially participate in the organization of chromosome architecture in mammals [11]. Two of these proteins, zinc finger protein 143 (ZNF143) and Yin Yang1 (YY1), bind DNA using C2H2 domains.

ZNF143 shares similarity with CTCF and the central region of ZNF143 contains a cluster that consists of seven C2H2 domains, three of which bind to a specific CCCAGCAG motif [77]. The N-terminal domain contains three 15-aa repeats with unknown functions, and the C-terminal domain is enriched in acidic amino acids, which is typical of transcription activators. ZNF143 is essential for embryonic development in mammals [78]. ZNF143 functions in the promoter region by participating in the formation of open chromatin regions and the recruitment of complexes that activate transcription [79]. Genome-wide studies have shown that ZNF143 can participate in the formation of chromatin loops between enhancers and promoters [80–82]. In human HEK293T cells, ZNF143 functions together with CTCF to form chromatin loops at several genomic sites [82]. However, in contrast to CTCF, no experimental evidence has suggested that ZNF143 participates in the localization of the cohesin complex on chromatin. Thus, how ZNF143 supports specific distance interactions remains unknown.

The mammalian YY1 protein, which consists of only 414 amino acids, belongs to a multifunctional, evolutionarily conserved family of mammalian transcription factors and contains 4 C2H2 domains at the C-terminus [83]. *Drosophila* expresses two homologs of the YY1 protein, PHO and PHOL, which are involved in the recruitment of Polycomb proteins [84]. The N-terminal region of YY1 has been implicated in transcriptional activation, whereas the domain between 201-226 amino acids is involved in the recruitment of PcG proteins that are responsible for repression [85–87].

YY1 is predominantly associated with gene enhancers and promoters, which is consistent with a potential role in distance interactions [88,89]. The inactivation of YY1 results in a marked decrease in the number of distance interactions *in vitro* [88]. According to the proposed model, YY1 forms homodimers, which can bring the associated gene enhancers and promoters closer together. A cluster of C2H2 domains and an adjacent unstructured domain are responsible for the dimerization and oligomerization of YY1 [90,91]. Moreover, the dimerization and subsequent oligomerization of YY1 results in non-specific binding to DNA, especially with guanine quadruplexes (G4). Some experimental evidence has suggested that the dimerization of YY1 and the subsequent binding to G4 structures contribute to the YY1-mediated formation of long DNA loops [91]. YY1 may also participate in the organization of distance interactions through the regulation of proteins that are directly involved in the formation of chromatin loops. For example, YY1, together with Oct4, participates in the recruitment of the BAF remodeling complex to promoters and super-enhancers [92]. BAF can improve TF binding and stabilize the chromatin loops formed by CTCF/cohesin [89]. Therefore, YY1 appears likely to regulate enhancer activity and enhancer-promoter interactions through epigenetic mechanisms [93].

The role of the small protein LIM domain-binding factor 1 (LDB1) in the maintenance of distance interactions between enhancers and promoters has been studied in detail [94]. Unlike C2H2 proteins, LDB1 binds to enhancers and promoters through the interaction between its C-terminal domain and DNA-binding TFs that belong to the LIM family [94]. Through interactions with various LIM partners, LDB1 plays roles in several regulatory processes during embryonic development and cell differentiation, including erythropoiesis. Initially, the N-terminal domain of LDB1 was shown to be involved in the organization of interactions between a strong enhancer (LCR) and the promoters of the *beta-globin* locus [95,96]. Structural analysis showed that the N-terminal domain of LBD1 contains both alpha helices and beta sheets, which form a stable homodimer [97,98]. TFs in the LIM family predominantly bind to gene enhancers and promoters, facilitating the recruitment of LDB1 to these regulatory elements. According to the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

model, specific interaction between the N-terminal domains of LBD1 molecules associated with enhancer and promoter elements can stabilize distance interactions between these regulatory elements. Interestingly, LDB1 not only homodimerizes but can also interact with CTCF, which can promote the organization of contacts between enhancers and promoters associated with a large group of erythroid genes [99]. A small domain in LDB1, located near the N-terminal dimerization domain, interacts with an unidentified C2H2 domain in the CTCF protein. However, whether the interaction between the LBD1 domain and the C2H2 domain in CTCF is capable of forming a sufficiently stable and specific interaction between enhancers and promoters to regulate gene expression remains unknown. Recently, mutations in LDB1 that disrupt dimerization were shown to have no effect on the transcriptional activation of *beta-globin* genes [98], which suggested the existence of additional mechanisms to support the interaction between the LCR enhancer and *beta-globin* promoters. Therefore, LBD1 likely acts in cooperation with other unidentified proteins to support distance interactions.

5. Specific activation of olfactory receptor genes in mammals is supported by super-long-distance interactions between enhancers and promoters

The most interesting models for understanding the mechanisms of distance interactions are proven examples of interactions between enhancers and promoters that are separated by megabase distances. In mammals, expression mechanisms associated with a large family of genes encoding olfactory receptors have been well-studied [100]. In the mouse genome, approximately 1,100 genes encoding olfactory receptors (ORs) have been identified [101,102]. These genes are located in 40 clusters that are scattered throughout the mouse genome. Olfactory sensory neurons (OSNs) are derived from progenitor cells, in which all *OR* genes are very weakly co-expressed. Constitutive heterochromatin is formed on inactive *OR* genes, which are enriched in H3K9me3 and H4K20me3 histone modifications [103]. During OSN maturation, the transcription of one *OR* gene is activated randomly, while all others remain completely repressed [104]. A negative feedback loop likely exists, in which the strong expression of one OR receptor results in the complete inactivation of all other *OR* genes [105–107]. However, the exact mechanism associated with this process remains poorly understood.

Each cluster of *OR* genes has a nearby enhancer, which is involved in the selection of one gene that will be actively transcribed in a particular OSN [108,109]. A total of 14 specific enhancers have been identified, and the *in vivo* deletion of three of these enhancers has been

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

shown to result in the complete inability to activate any of the *OR* genes encoded in the corresponding cluster [108,110–112]. A number of studies have shown that enhancers form a single cluster in the nucleus, which regulates the activation of a single selected *OR* gene [112], and all other genes form heterochromatin. Thus, active chromatin appears to be generated by the interactions between enhancers and the promoter of a single *OR* gene that is encoded in close proximity to clusters of *OR* genes to organize heterochromatin regions.

The mechanisms that underly the physical separation and stable balance between regions containing both active and repressed chromatin remain poorly understood. *OR* gene promoters and enhancers contain binding sites for Ebf and Lhx2, which are specific TFs expressed in neurons [20,109,112]. According to the model [100], the formation of a cluster of enhancers can increase the efficiency of Ebf and Lhx2 recruitment, resulting in the significant enrichment of transcription activators associated with the active *OR* promotor compared with the promotors associated with repressed *OR* genes (Monahan et al., 2017). A high concentration of activators can also prevent the spread of heterochromatin to active promotor regions associated with enhancers.

The most important aim is the identification of proteins that can support the specific super-long-distance interactions between enhancers that form a single cluster. CTCF and cohesin were not identified in the regulatory regions of *OR* genes. However, Lhx2 has been shown to recruit LDB1 to *OR* enhancers, and the inactivation of LBD1 has been shown to result in the partial disruption of *OR* enhancer colocalization in the nucleus [113]. These results suggest a potential role for LBD1 in the organization of super-long-distance interactions among *OR* enhancers. However, the most likely scenario is that several unknown architectural proteins are involved in the initiation and maintenance of distance interactions and that LBD1 facilitates their functions.

6. Super-long-distance interactions in the *Drosophila* genome and the role of architectural proteins during this process

In *Drosophila*, super-long-distance interactions between regulatory elements were first described when studying the effects of an insulator, found in the *gypsy* retrotransposon, on reporter gene repression induced by a Polycomb-dependent silencer [Polycomb response element (PRE)] [114]. This study used a 660-bp PRE, found in the regulatory *bxd* region of the homeotic *Ubx* gene of the *bithorax* locus (*bxd* PRE). Typically, pairing two PRE-containing transgenes results in the increased repression of the reporter, which is likely associated with an increase in

the recruitment efficiency of PcG complexes [31]. The combination of two transgenes that both contained PREs and the *gypsy* insulator resulted in the significant repression of reporter genes [114], despite these genes being separated by several megabases or even located on different chromosomes.

The interaction between *gypsy* insulators can facilitate a functional interaction between an enhancer and a promoter that are located in association with various transgenes that are integrated into genomic sites separated by several megabases [115]. These results strongly suggested that interactions between *gypsy* insulators can support stable interactions between genome regions separated by very large distances.

A similar study was performed for transgenes containing the Miscadestral pigmentation (*Mcp*) boundary, which separates the regulatory domains that control the expression of the homeotic genes *abd-A* and *Abd-B* of the *bithorax* complex [116]. The *Mcp* boundary consists of an insulator, which is flanked on both sides by Polycomb-dependent silencers. The functional interaction between transgenes, which are inserted in different regions of the third chromosome, was investigated. The results indicated that transgenes located at a distance of several megabases, and in some cases on different arms of the same chromosome, can induce the mutual enhancement of repression, which suggested that they were able to interact physically. In a similar study, the *bxd* PRE, in combination with the 210-bp core of the *Mcp* insulator, was able to support repression between two transgenes located at super-long-distances [117]. The co-repression of reporters and their colocalization in the nucleus were observed only in the presence of the *Mcp* insulators for both tested transgenes [117–119].

Super-long-distance interactions can be also be maintained by the *Frontadominal-7* (*Fab-7*) boundary, which separates the regulatory domains responsible for the activation of the *Abd-B* gene in the *bithorax* complex [120]. At various genomic sites, a 3.6-kb DNA fragment that included the *Fab-7* boundary and an adjacent PRE functionally interacted with endogenous *Fab-7* [120]. The pairing between the 1,250-bp *Fab-7* insulators was also able to support super-long-distance interactions between transgenes in the *Drosophila* genome [117,121].

Two insulators, *Nhomie* (*Neighbor of Homie*) and *Homie* (*Homing* insulator at *eve*), were identified at the 16-kb boundaries of the regulatory region of the *even-skipped* (*eve*) gene (Fujioka et al. 1999; 2009; 2016). *Nhomie* and *Homie* interact with each other and can also maintain super-long-distance interactions between the transgene and the endogenous *eve* locus,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

which allows endogenous enhancers to activate the reporter gene promoter in the transgene [122,123].

The interaction specificity between identical regulatory elements is also manifested in the "homing" phenomenon, in which the *P*-transposon, which contains either an insulator or a promoter region, is inserted with high frequency into the genomic region where this regulatory element is located. For example, the *P*-transposon that contains the promoter region of *engrailed* or *linotte* was predominantly inserted (20%–30%) into the corresponding genes [124,125]. Similar results were obtained upon integration into the genome of the *P*-transposon containing the *Fub* boundary, which separates the regulatory regions of the *Ubx* and *abd-A* genes in the *bithorax* complex [126]. The *Fub* boundary between two TADs formed from the architecture of the 300,000-bp *bithorax* locus [127]. The "homing" effect occurs due to the interaction between architectural proteins, which is associated with two identical regions and directs the integration of the *P* transposon into the region of the corresponding gene.

In all of these examples, super-long-distance interactions were supported by a pair of identical regulatory elements that were unique to the genome. The *gypsy* insulator consists of 12 sites that bind the Su(Hw) protein [128–130]. In the genome of most *Drosophila* lines, the *gypsy* retrotransposon is only found in heterochromatin [131]. Only a small portion of the several thousand proteins with identified Su(Hw) binding regions contain 2-3 binding sites [132,133]. Thus, the *gypsy* insulator is unique due to a large number of Su(Hw) binding sites. According to ModEncode, the *homie* insulator contains binding sites for the *Drosophila* homolog of CTCF (dCTCF), Su(Hw), and GAF proteins [134]. The *Mcp* boundary contains binding sites for dCTCF and Pita [135]. Similar to dCTCF, Su(Hw) and Pita contain clusters of C2H2 domains, some of which facilitate the specific binding of these proteins with long DNA motifs [132,136,137]. The *Fub* boundary contains binding sites for CTCF, Su(Hw), and Pita [135,138,139] (Bender and Lucas 2013, Kyrchanova et al. 2017; 2020). Using *in vivo* deletion analysis, additional unknown proteins are also determined to be involved in determining the boundary function of *Mcp* and *Fub* [139].

Finally, the *Fab-7* boundary has been subdivided into three domains, which coincide with mapped DNase 1 hypersensitivity sites [140–142]. HS2 contains two Pita sites [135], whereas the central HS1 domain was found to contain six sites for the GAF protein that overlap with long, degenerate binding sites for the newly identified late boundary complex (LBC) DNA-binding complex [143], which is likely to be involved in the regulation of distance interactions between enhancers and promoters [135,144]. Thus, several proteins likely function together to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

organize active boundaries/insulators that are capable of supporting super-long-distance interactions.

In transgenic model systems, the pairing between two copies of repetitive binding sites for dCTCF, Su(Hw), or Pita was able to bring the yeast GAL4-dependent activator region and the reporter gene promoter in close proximity, resulting in the transcription of the reporter gene [145–147]. The dCTCF N-terminus contains an unstructured domain that can form tetrameric complexes [65,148], which can contribute to distance interactions. Similarly, Pita contains an N-terminal zinc-finger-associated domain (ZAD) capable of forming homodimers [147]. ZADs have also been identified in the N-termini of approximately one hundred *Drosophila* proteins containing clusters of C2H2 domains (ZAD-C2H2) [149,150]. An important feature of ZADs is their preferentially ability to homodimerize into an antiparallel dimer [147,151]. In addition to Pita, several other ZAD-C2H2 proteins, including ZIPIC (Zinc-finger protein interacting with CP190), Zw5 (Zeste-white 5), and ZAF1, have been shown to support distance interactions and form functional insulators [137,152–155]. Mutational analysis, *in vivo*, showed that the presence of ZADs determined the ability of these proteins to support distance interactions [154,155]. In transgenic lines, various combinations of reiterated binding sites from different ZAD-C2H2 proteins were unable to support the distance activation of the reporter gene by the GAL4 activator [145,147]. Thus, the homodimerization of ZAD regions is an important feature required to support specific distance interactions between ZAD-C2H2 proteins.

The C2H2 proteins, including Pita, dCTCF, and Su(Hw), can recruit chromatin to BTB (bric-a-brac, tramtrack, and broad complex)-containing proteins, such as Mod(mdg4) and CP190 [137,156–158]. The CP190 protein contains a classical N-terminal BTB domain that forms homodimers and is conserved among higher eukaryotes [159–161]. In contrast, the BTB domain of Mod(mdg4)-67.2 belongs to an insect-specific group [162,163]. The BTB domains in this group of proteins can form both homo- and heteromultimeric complexes [161]. Previously, both CP190 and Mod(mdg4)-67.2 were thought to be capable of participating in the formation of distance interactions between regulatory elements and Su(Hw)-dependent insulators [164,165]. Both proteins are required for the functional activity of the *gypsy* insulator [158,166–168]. However, the CP190 and Mod(mdg4)-67.2 can also interact with a large variety of DNA-binding proteins [137,148,156–158,169–171], which is inconsistent with their key role in the organization of specific distance interactions between *gypsy* insulators. Most likely, CP190 and Mod(mdg4)-67.2 participate in the stabilization of previously formed specific distance interactions that occur between other regulatory elements.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 December 2020 doi:10.20944/preprints202012.0184.v1

7. Conclusion

Additional study is necessary to better understand the mechanisms of distance interactions between enhancers and promoters. The homodimerization that occurs between the N-terminal domains of LDB1, which is recruited to regulatory elements through interaction with LIM proteins, has been best-described to support the maintenance of interactions between enhancers and promoters [172]. Simultaneously, evidence has suggested that LBD1 supports remote interactions between currently unknown partners. The mammalian CTCF protein contains binding sites that would allow it to interact simultaneously with enhancers and promoters associated with a number of genes, and, together with the cohesin complex, CTCF appears to be involved in facilitating distance interactions between these regulatory elements [68,69].

In *Drosophila*, identical insulators, consisting of specific combinations of binding sites for a variety of C2H2 proteins, are able to interact at large distances. According to the current model [150], many C2H2 proteins predominantly homodimerize, typically through N-terminal domains, which ensures the specificity and stability of interactions between identical insulators that are located at large distances from each other. The identification of binding sites for currently uncharacterized human and *Drosophila* C2H2 proteins will facilitate the assessment of the true contributions of this class of proteins to the organization of the chromosomal architecture. In addition, the use of gene-editing techniques, such as the (CRISPR)/Cas9 method, will allow the role of each identified C2H2 protein in the maintenance of distance interactions to be assessed using model regulatory systems.

Author Contributions: P.G. conceived the story; O.K. and P.G. wrote the article; O.K. prepared figures; P.G. provided supervision and prepared the article. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Russian Science Foundation, project no. 19-74-30026 (to P.G.)

Conflicts of Interest: The authors declare no conflict of interest.

 doi:10.20944/preprints202012.0184.v1

Abbreviations

TF Transcription Factor

CBP CREB-binding protein

Mll3/4 Mixed Lineage Leukemia 3/4

UTX demethylase

PcG Polycomb Group

PRC1/2 Polycomb repressive complexes 1/2

CTCF CTC binding factor

ZNF143 Zinc Finger Protein 143

LBD1 LIM domain – binding factor 1

CP190 Centrosomal protein 190kD

Mod(mdg4) Modifier of mdg4

CRISPR the clustered regularly interspaced short palindromic repeats

References

1. Spitz, F.; Furlong, E.E.M. Transcription factors: from enhancer binding to developmental control. *Nat. Rev. Genet.* **2012**, *13*, 613–626, doi:10.1038/nrg3207.
2. Zabidi, M.A.; Stark, A. Regulatory Enhancer–Core-Promoter Communication via Transcription Factors and Cofactors. *Trends Genet.* **2016**, *32*, 801–814, doi:10.1016/j.tig.2016.10.003.
3. Levine, M.; Cattoglio, C.; Tjian, R. Looping Back to Leap Forward: Transcription Enters a New Era. *Cell* **2014**, *157*, 13–25, doi:10.1016/j.cell.2014.02.009.

4. Moreau, P.; Hen, R.; Wasylyk, B.; Everett, R.; Gaub, M.P.; Chambon, P. The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res.* **1981**, *9*, 6047–6068, doi:10.1093/nar/9.22.6047.

5. Banerji, J.; Rusconi, S.; Schaffner, W. Expression of a β-globin gene is enhanced by remote SV40 DNA sequences. *Cell* **1981**, *27*, 299–308, doi:10.1016/0092-8674(81)90413-X.

6. Banerji, J.; Olson, L.; Schaffner, W. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell* **1983**, *33*, 729–740, doi:10.1016/0092-8674(83)90015-6.

7. Li, Y.; Shi, W.; Wasserman, W.W. Genome-wide prediction of cis-regulatory regions using supervised deep learning methods. *BMC Bioinformatics* **2018**, *19*, 202, doi:10.1186/s12859-018-2187-1.

8. Furlong, E.E.M.; Levine, M. Developmental enhancers and chromosome topology. *Science (80-.).* **2018**, *361*, 1341–1345, doi:10.1126/science.aau0320.

9. Reiter, F.; Wienerroither, S.; Stark, A. Combinatorial function of transcription factors and cofactors. *Curr. Opin. Genet. Dev.* **2017**, *43*, 73–81, doi:10.1016/j.gde.2016.12.007.

10. Rickels, R.; Shilatifard, A. Enhancer Logic and Mechanics in Development and Disease. *Trends Cell Biol.* **2018**, *28*, 608–630, doi:10.1016/j.tcb.2018.04.003.

11. Schoenfelder, S.; Fraser, P. Long-range enhancer–promoter contacts in gene expression control. *Nat. Rev. Genet.* **2019**, *20*, 437–455, doi:10.1038/s41576-019-0128-0.

12. Hnisz, D.; Abraham, B.J.; Lee, T.I.; Lau, A.; Saint-André, V.; Sigova, A.A.; Hoke, H.A.; Young, R.A. Super-Enhancers in the Control of Cell Identity and Disease. *Cell* **2013**, *155*, 934–947, doi:10.1016/j.cell.2013.09.053.

13. Whyte, W.A.; Orlando, D.A.; Hnisz, D.; Abraham, B.J.; Lin, C.Y.; Kagey, M.H.; Rahl, P.B.; Lee, T.I.; Young, R.A. Master Transcription Factors and Mediator Establish Super-Enhancers at Key Cell Identity Genes. *Cell* **2013**, *153*, 307–319, doi:10.1016/j.cell.2013.03.035.

14. Pott, S.; Lieb, J.D. What are super-enhancers? *Nat. Genet.* **2015**, *47*, 8–12, doi:10.1038/ng.3167.

15. Huang, J.; Li, K.; Cai, W.; Liu, X.; Zhang, Y.; Orkin, S.H.; Xu, J.; Yuan, G.-C. Dissecting super-enhancer hierarchy based on chromatin interactions. *Nat. Commun.* **2018**, *9*, 943,

doi:10.1038/s41467-018-03279-9.

16. Hay, D.; Hughes, J.R.; Babbs, C.; Davies, J.O.J.; Graham, B.J.; Hanssen, L.L.P.; Kassouf, M.T.; Oudelaar, A.M.; Sharpe, J.A.; Suciu, M.C.; et al. Genetic dissection of the α-globin super-enhancer in vivo. *Nat. Genet.* **2016**, *48*, 895–903, doi:10.1038/ng.3605.

17. Shin, H.Y.; Willi, M.; Yoo, K.H.; Zeng, X.; Wang, C.; Metser, G.; Hennighausen, L. Hierarchy within the mammary STAT5-driven Wap super-enhancer. *Nat. Genet.* **2016**, *48*, 904–911, doi:10.1038/ng.3606.

18. Cenik, B.K.; Shilatifard, A. COMPASS and SWI/SNF complexes in development and disease. *Nat. Rev. Genet.* **2020**, doi:10.1038/s41576-020-0278-0.

19. Li, Y.; Han, J.; Zhang, Y.; Cao, F.; Liu, Z.; Li, S.; Wu, J.; Hu, C.; Wang, Y.; Shuai, J.; et al. Structural basis for activity regulation of MLL family methyltransferases. *Nature* **2016**, *530*, 447–452, doi:10.1038/nature16952.

20. Zhang, Y.; Mittal, A.; Reid, J.; Reich, S.; Gamblin, S.J.; Wilson, J.R. Evolving Catalytic Properties of the MLL Family SET Domain. *Structure* **2015**, *23*, 1921–1933, doi:10.1016/j.str.2015.07.018.

21. Wang, C.; Lee, J.E.; Lai, B.; Macfarlan, T.S.; Xu, S.; Zhuang, L.; Liu, C.; Peng, W.; Ge, K. Enhancer priming by H3K4 methyltransferase MLL4 controls cell fate transition. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, *113*, 11871–11876, doi:10.1073/pnas.1606857113.

22. Bonn, S.; Zinzen, R.P.; Girardot, C.; Gustafson, E.H.; Perez-Gonzalez, A.; Delhomme, N.; Ghavi-Helm, Y.; Wilczyåski, B.; Riddell, A.; Furlong, E.E.M. Tissue-specific analysis of chromatin state identifies temporal signatures of enhancer activity during embryonic development. *Nat. Genet.* **2012**, *44*, 148–156, doi:10.1038/ng.1064.

23. Heintzman, N.D.; Hon, G.C.; Hawkins, R.D.; Kheradpour, P.; Stark, A.; Harp, L.F.; Ye, Z.; Lee, L.K.; Stuart, R.K.; Ching, C.W.; et al. Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* **2009**, *459*, 108–112, doi:10.1038/nature07829.

24. Zentner, G.E.; Tesar, P.J.; Scacheri, P.C. Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. *Genome Res.* **2011**, *21*, 1273–1283, doi:10.1101/gr.122382.111.

25. Buecker, C.; Wysocka, J. Enhancers as information integration hubs in development: lessons from genomics. *Trends Genet.* **2012**, *28*, 276–284, doi:10.1016/j.tig.2012.02.008.

26. Shlyueva, D.; Stampfel, G.; Stark, A. Transcriptional enhancers: from properties to

genome-wide predictions. *Nat. Rev. Genet.* **2014**, *15*, 272–286, doi:10.1038/nrg3682.

27. Dorighi, K.M.; Swigut, T.; Henriques, T.; Bhanu, N. V.; Scruggs, B.S.; Nady, N.; Still, C.D.; Garcia, B.A.; Adelman, K.; Wysocka, J. Mll3 and Mll4 Facilitate Enhancer RNA Synthesis and Transcription from Promoters Independently of H3K4 Monomethylation. *Mol. Cell* **2017**, *66*, 568-576.e4, doi:10.1016/j.molcel.2017.04.018.

28. Rickels, R.; Herz, H.-M.; Sze, C.C.; Cao, K.; Morgan, M.A.; Collings, C.K.; Gause, M.; Takahashi, Y.; Wang, L.; Rendleman, E.J.; et al. Histone H3K4 monomethylation catalyzed by Trr and mammalian COMPASS-like proteins at enhancers is dispensable for development and viability. *Nat. Genet.* **2017**, *49*, 1647–1653, doi:10.1038/ng.3965.

29. Tessarz, P.; Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat. Rev. Mol. Cell Biol.* **2014**, *15*, 703–708, doi:10.1038/nrm3890.

30. Douillet, D.; Sze, C.C.; Ryan, C.; Piunti, A.; Shah, A.P.; Ugarenko, M.; Marshall, S.A.; Rendleman, E.J.; Zha, D.; Helmin, K.A.; et al. Uncoupling histone H3K4 trimethylation from developmental gene expression via an equilibrium of COMPASS, Polycomb and DNA methylation. *Nat. Genet.* **2020**, *52*, 615–625, doi:10.1038/s41588-020-0618-1.

31. Kuroda, M.I.; Kang, H.; De, S.; Kassis, J.A. Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming. *Annu. Rev. Biochem.* **2020**, *89*, 235–253, doi:10.1146/annurev-biochem-120219-103641.

32. Schröder, S.; Herker, E.; Itzen, F.; He, D.; Thomas, S.; Gilchrist, D.A.; Kaehlcke, K.; Cho, S.; Pollard, K.S.; Capra, J.A.; et al. Acetylation of RNA Polymerase II Regulates Growth-Factor-Induced Gene Transcription in Mammalian Cells. *Mol. Cell* **2013**, *52*, 314–324, doi:10.1016/j.molcel.2013.10.009.

33. Boija, A.; Mahat, D.B.; Zare, A.; Holmqvist, P.-H.; Philip, P.; Meyers, D.J.; Cole, P.A.; Lis, J.T.; Stenberg, P.; Mannervik, M. CBP Regulates Recruitment and Release of Promoter-Proximal RNA Polymerase II. *Mol. Cell* **2017**, *68*, 491-503.e5, doi:10.1016/j.molcel.2017.09.031.

34. Soutourina, J. Transcription regulation by the Mediator complex. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 262–274, doi:10.1038/nrm.2017.115.

35. Cevher, M.A.; Shi, Y.; Li, D.; Chait, B.T.; Malik, S.; Roeder, R.G. Reconstitution of active human core Mediator complex reveals a critical role of the MED14 subunit. *Nat. Struct. Mol. Biol.* **2014**, *21*, 1028–1034, doi:10.1038/nsmb.2914.

36. Nozawa, K.; Schneider, T.R.; Cramer, P. Core Mediator structure at 3.4 Å extends model

of transcription initiation complex. *Nature* **2017**, *545*, 248–251, doi:10.1038/nature22328.

37. Plaschka, C.; Larivière, L.; Wenzeck, L.; Seizl, M.; Hemann, M.; Tegunov, D.; Petrotchenko, E. V.; Borchers, C.H.; Baumeister, W.; Herzog, F.; et al. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **2015**, *518*, 376–380, doi:10.1038/nature14229.

38. Robinson, P.J.; Trnka, M.J.; Bushnell, D.A.; Davis, R.E.; Mattei, P.-J.; Burlingame, A.L.; Kornberg, R.D. Structure of a Complete Mediator-RNA Polymerase II Pre-Initiation Complex. *Cell* **2016**, *166*, 1411-1422.e16, doi:10.1016/j.cell.2016.08.050.

39. Tsai, K.-L.; Yu, X.; Gopalan, S.; Chao, T.-C.; Zhang, Y.; Florens, L.; Washburn, M.P.; Murakami, K.; Conaway, R.C.; Conaway, J.W.; et al. Mediator structure and rearrangements required for holoenzyme formation. *Nature* **2017**, *544*, 196–201, doi:10.1038/nature21393.

40. Schier, A.C.; Taatjes, D.J. Structure and mechanism of the RNA polymerase II transcription machinery. *Genes Dev.* **2020**, *34*, 465–488, doi:10.1101/gad.335679.119.

41. Sierecki, E. The Mediator complex and the role of protein-protein interactions in the gene regulation machinery. *Semin. Cell Dev. Biol.* **2020**, *99*, 20–30, doi:10.1016/j.semcdb.2018.08.006.

42. André, K.M.; Sipos, E.H.; Soutourina, J. Mediator Roles Going Beyond Transcription. *Trends Genet.* **2020**, doi:10.1016/j.tig.2020.08.015.

43. Kvon, E.Z.; Kazmar, T.; Stampfel, G.; Yáñez-Cuna, J.O.; Pagani, M.; Schernhuber, K.; Dickson, B.J.; Stark, A. Genome-scale functional characterization of Drosophila developmental enhancers in vivo. *Nature* **2014**, *512*, 91–95, doi:10.1038/nature13395.

44. Ghavi-Helm, Y.; Klein, F.A.; Pakozdi, T.; Ciglar, L.; Noordermeer, D.; Huber, W.; Furlong, E.E. Enhancer loops appear stable during development and are associated with paused polymerase. *Nature* **2014**, *512*, 96–100, doi:10.1038/nature13417.

45. Jin, F.; Li, Y.; Dixon, J.R.; Selvaraj, S.; Ye, Z.; Lee, A.Y.; Yen, C.-A.; Schmitt, A.D.; Espinoza, C.A.; Ren, B. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* **2013**, *503*, 290–294, doi:10.1038/nature12644.

46. Li, G.; Ruan, X.; Auerbach, R.K.; Sandhu, K.S.; Zheng, M.; Wang, P.; Poh, H.M.; Goh, Y.; Lim, J.; Zhang, J.; et al. Extensive Promoter-Centered Chromatin Interactions Provide a Topological Basis for Transcription Regulation. *Cell* **2012**, *148*, 84–98, doi:10.1016/j.cell.2011.12.014.

47. Rubin, A.J.; Barajas, B.C.; Furlan-Magaril, M.; Lopez-Pajares, V.; Mumbach, M.R.; Howard, I.; Kim, D.S.; Boxer, L.D.; Cairns, J.; Spivakov, M.; et al. Lineage-specific dynamic and pre-established enhancer–promoter contacts cooperate in terminal differentiation. *Nat. Genet.* **2017**, *49*, 1522–1528, doi:10.1038/ng.3935.

48. Chen, H.; Levo, M.; Barinov, L.; Fujioka, M.; Jaynes, J.B.; Gregor, T. Dynamic interplay between enhancer–promoter topology and gene activity. *Nat. Genet.* **2018**, *50*, 1296–1303, doi:10.1038/s41588-018-0175-z.

49. Novo, C.L.; Javierre, B.-M.; Cairns, J.; Segonds-Pichon, A.; Wingett, S.W.; Freire-Pritchett, P.; Furlan-Magaril, M.; Schoenfelder, S.; Fraser, P.; Rugg-Gunn, P.J. Long-Range Enhancer Interactions Are Prevalent in Mouse Embryonic Stem Cells and Are Reorganized upon Pluripotent State Transition. *Cell Rep.* **2018**, *22*, 2615–2627, doi:10.1016/j.celrep.2018.02.040.

50. El Khattabi, L.; Zhao, H.; Kalchschmidt, J.; Young, N.; Jung, S.; Van Blerkom, P.; Kieffer-Kwon, P.; Kieffer-Kwon, K.-R.; Park, S.; Wang, X.; et al. A Pliable Mediator Acts as a Functional Rather Than an Architectural Bridge between Promoters and Enhancers. *Cell* **2019**, *178*, 1145-1158.e20, doi:10.1016/j.cell.2019.07.011.

51. Jeronimo, C.; Robert, F. The Mediator Complex: At the Nexus of RNA Polymerase II Transcription. *Trends Cell Biol.* **2017**, *27*, 765–783, doi:10.1016/j.tcb.2017.07.001.

52. Bartkuhn, M.; Straub, T.; Herold, M.; Herrmann, M.; Rathke, C.; Saumweber, H.; Gilfillan, G.D.; Becker, P.B.; Renkawitz, R. Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J.* **2009**, *28*, 877–88, doi:10.1038/emboj.2009.34.

53. Fukaya, T.; Lim, B.; Levine, M. Enhancer Control of Transcriptional Bursting. *Cell* **2016**, *166*, 358–368, doi:10.1016/j.cell.2016.05.025.

54. Petrenko, N.; Jin, Y.; Wong, K.H.; Struhl, K. Mediator Undergoes a Compositional Change during Transcriptional Activation. *Mol. Cell* **2016**, *64*, 443–454, doi:10.1016/j.molcel.2016.09.015.

55. Jeronimo, C.; Langelier, M.-F.; Bataille, A.R.; Pascal, J.M.; Pugh, B.F.; Robert, F. Tail and Kinase Modules Differently Regulate Core Mediator Recruitment and Function In Vivo. *Mol. Cell* **2016**, *64*, 455–466, doi:10.1016/j.molcel.2016.09.002.

56. van Steensel, B.; Furlong, E.E.M. The role of transcription in shaping the spatial organization of the genome. *Nat. Rev. Mol. Cell Biol.* 2019, *20*, 327–337.

57. Sikorska, N.; Sexton, T. Defining Functionally Relevant Spatial Chromatin Domains: It is

a TAD Complicated. *J. Mol. Biol.* **2020**, *432*, 653–664, doi:10.1016/j.jmb.2019.12.006.

58. Hansen, A.S.; Cattoglio, C.; Darzacq, X.; Tjian, R. Recent evidence that TADs and chromatin loops are dynamic structures. *Nucleus* **2018**, *9*, 20–32, doi:10.1080/19491034.2017.1389365.

59. Szabo, Q.; Bantignies, F.; Cavalli, G. Principles of genome folding into topologically associating domains. *Sci. Adv.* **2019**, *5*, eaaw1668, doi:10.1126/sciadv.aaw1668.

60. Chang, L.-H.; Ghosh, S.; Noordermeer, D. TADs and Their Borders: Free Movement or Building a Wall? *J. Mol. Biol.* **2020**, *432*, 643–652, doi:10.1016/j.jmb.2019.11.025.

61. Luppino, J.M.; Joyce, E.F. Single cell analysis pushes the boundaries of TAD formation and function. *Curr. Opin. Genet. Dev.* **2020**, *61*, 25–31, doi:10.1016/j.gde.2020.03.005.

62. Bozza, T.; Vassalli, A.; Fuss, S.; Zhang, J.-J.; Weiland, B.; Pacifico, R.; Feinstein, P.; Mombaerts, P. Mapping of Class I and Class II Odorant Receptors to Glomerular Domains by Two Distinct Types of Olfactory Sensory Neurons in the Mouse. *Neuron* **2009**, *61*, 220–233, doi:10.1016/j.neuron.2008.11.010.

63. Heger, P.; Marin, B.; Bartkuhn, M.; Schierenberg, E.; Wiehe, T. The chromatin insulator CTCF and the emergence of metazoan diversity. *Proc Natl Acad Sci U S A* **2012**, *109*, 17507–17512, doi:10.1073/pnas.1111941109.

64. Hashimoto, H.; Wang, D.; Horton, J.R.; Zhang, X.; Corces, V.G.; Cheng, X. Structural Basis for the Versatile and Methylation-Dependent Binding of CTCF to DNA. *Mol. Cell* **2017**, *66*, 711-720.e3, doi:10.1016/j.molcel.2017.05.004.

65. Bonchuk, A.; Kamalyan, S.; Mariasina, S.; Boyko, K.; Popov, V.; Maksimenko, O.; Georgiev, P. N-terminal domain of the architectural protein CTCF has similar structural organization and ability to self-association in bilaterian organisms. *Sci. Rep.* **2020**, *10*, 2677, doi:10.1038/s41598-020-59459-5.

66. Li, Y.; Haarhuis, J.H.I.; Sedeño Cacciatore, Á.; Oldenkamp, R.; van Ruiten, M.S.; Willems, L.; Teunissen, H.; Muir, K.W.; de Wit, E.; Rowland, B.D.; et al. The structural basis for cohesin–CTCF-anchored loops. *Nature* **2020**, doi:10.1038/s41586-019-1910-z.

67. Fudenberg, G.; Imakaev, M.; Lu, C.; Goloborodko, A.; Abdennur, N.; Mirny, L.A. Formation of Chromosomal Domains by Loop Extrusion. *Cell Rep.* **2016**, *15*, 2038–2049, doi:10.1016/j.celrep.2016.04.085.

68. Ren, G.; Jin, W.; Cui, K.; Rodrigez, J.; Hu, G.; Zhang, Z.; Larson, D.R.; Zhao, K. CTCF-Mediated Enhancer-Promoter Interaction Is a Critical Regulator of Cell-to-Cell Variation

of Gene Expression. *Mol. Cell* **2017**, *67*, 1049-1058.e6, doi:10.1016/j.molcel.2017.08.026.

69. Hyle, J.; Zhang, Y.; Wright, S.; Xu, B.; Shao, Y.; Easton, J.; Tian, L.; Feng, R.; Xu, P.; Li, C. Acute depletion of CTCF directly affects MYC regulation through loss of enhancer–promoter looping. *Nucleic Acids Res.* **2019**, *47*, 6699–6713, doi:10.1093/nar/gkz462.

70. Nora, E.P.; Goloborodko, A.; Valton, A.-L.; Gibcus, J.H.; Uebersohn, A.; Abdennur, N.; Dekker, J.; Mirny, L.A.; Bruneau, B.G. Targeted Degradation of CTCF Decouples Local Insulation of Chromosome Domains from Genomic Compartmentalization. *Cell* **2017**, *169*, 930-944.e22, doi:10.1016/j.cell.2017.05.004.

71. Pugacheva, E.M.; Kubo, N.; Loukinov, D.; Tajmul, M.; Kang, S.; Kovalchuk, A.L.; Strunnikov, A. V.; Zentner, G.E.; Ren, B.; Lobanenkov, V. V. CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention. *Proc. Natl. Acad. Sci.* **2020**, *117*, 2020–2031, doi:10.1073/pnas.1911708117.

72. 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.; et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **2014**, *159*, 1665–1680, doi:10.1016/j.cell.2014.11.021.

73. Vietri Rudan, M.; Barrington, C.; Henderson, S.; Ernst, C.; Odom, D.T.; Tanay, A.; Hadjur, S. Comparative Hi-C Reveals that CTCF Underlies Evolution of Chromosomal Domain Architecture. *Cell Rep.* **2015**, *10*, 1297–1309, doi:10.1016/j.celrep.2015.02.004.

74. Rao, S.S.P.; Huang, S.-C.; Glenn St Hilaire, B.; Engreitz, J.M.; Perez, E.M.; Kieffer-Kwon, K.-R.; Sanborn, A.L.; Johnstone, S.E.; Bascom, G.D.; Bochkov, I.D.; et al. Cohesin Loss Eliminates All Loop Domains. *Cell* **2017**, *171*, 305-320.e24, doi:10.1016/j.cell.2017.09.026.

75. Thiecke, M.J.; Wutz, G.; Muhar, M.; Tang, W.; Bevan, S.; Malysheva, V.; Stocsits, R.; Neumann, T.; Zuber, J.; Fraser, P.; et al. Cohesin-Dependent and -Independent Mechanisms Mediate Chromosomal Contacts between Promoters and Enhancers. *Cell Rep.* **2020**, *32*, 107929, doi:10.1016/j.celrep.2020.107929.

76. Krivega, I.; Dean, A. LDB1-mediated enhancer looping can be established independent of mediator and cohesin. *Nucleic Acids Res.* **2017**, *45*, 8255–8268, doi:10.1093/nar/gkx433.

77. Schuster, C.; Myslinski, E.; Krol, A.; Carbon, P. Staf, a novel zinc finger protein that activates the RNA polymerase III promoter of the selenocysteine tRNA gene. *EMBO J.* **1995**, *14*, 3777–3787, doi:10.1002/j.1460-2075.1995.tb00047.x.

78. Myslinski, E.; Krol, A.; Carbon, P. ZNF76 and ZNF143 Are Two Human Homologs of the Transcriptional Activator Staf. *J. Biol. Chem.* **1998**, *273*, 21998–22006, doi:10.1074/jbc.273.34.21998.

79. Sathyan, K.M.; McKenna, B.D.; Anderson, W.D.; Duarte, F.M.; Core, L.; Guertin, M.J. An improved auxin-inducible degron system preserves native protein levels and enables rapid and specific protein depletion. *Genes Dev.* **2019**, *33*, 1441–1455, doi:10.1101/gad.328237.119.

80. Bailey, S.D.; Zhang, X.; Desai, K.; Aid, M.; Corradin, O.; Cowper-Sal·lari, R.; Akhtar-Zaidi, B.; Scacheri, P.C.; Haibe-Kains, B.; Lupien, M. ZNF143 provides sequence specificity to secure chromatin interactions at gene promoters. *Nat. Commun.* **2015**, *6*, 6186, doi:10.1038/ncomms7186.

81. Heidari, N.; Phanstiel, D.H.; He, C.; Grubert, F.; Jahanbani, F.; Kasowski, M.; Zhang, M.Q.; Snyder, M.P. Genome-wide map of regulatory interactions in the human genome. *Genome Res* **2014**, *24*, 1905–1917, doi:10.1101/gr.176586.114.

82. Wen, Z.; Huang, Z.-T.; Zhang, R.; Peng, C. ZNF143 is a regulator of chromatin loop. *Cell Biol. Toxicol.* **2018**, *34*, 471–478, doi:10.1007/s10565-018-9443-z.

83. Deng, Z.; Cao, P.; Wan, M.M.; Sui, G. Yin Yang 1. *Transcription* **2010**, *1*, 81–84, doi:10.4161/trns.1.2.12375.

84. Kassis, J.A.; Kennison, J.A.; Tamkun, J.W. Polycomb and trithorax group genes in drosophila. *Genetics* **2017**, *206*, 1699–1725, doi:10.1534/genetics.115.185116.

85. Wilkinson, F.H.; Park, K.; Atchison, M.L. Polycomb recruitment to DNA in vivo by the YY1 REPO domain. *Proc. Natl. Acad. Sci.* **2006**, *103*, 19296–19301, doi:10.1073/pnas.0603564103.

86. Yang, W.-M.; Inouye, C.; Zeng, Y.; Bearss, D.; Seto, E. Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. *Proc. Natl. Acad. Sci.* **1996**, *93*, 12845–12850, doi:10.1073/pnas.93.23.12845.

87. Lewis, B.A.; Tullis, G.; Seto, E.; Horikoshi, N.; Weinmann, R.; Shenk, T. Adenovirus E1A proteins interact with the cellular YY1 transcription factor. *J. Virol.* **1995**, *69*, 1628–1636, doi:10.1128/JVI.69.3.1628-1636.1995.

88. Weintraub, A.S.; Li, C.H.; Zamudio, A. V.; Sigova, A.A.; Hannett, N.M.; Day, D.S.; Abraham, B.J.; Cohen, M.A.; Nabet, B.; Buckley, D.L.; et al. YY1 Is a Structural Regulator of Enhancer-Promoter Loops. *Cell* **2017**, *171*, 1573-1588.e28,

doi:10.1016/j.cell.2017.11.008.

89. Beagan, J.A.; Duong, M.T.; Titus, K.R.; Zhou, L.; Cao, Z.; Ma, J.; Lachanski, C. V.; Gillis, D.R.; Phillips-Cremins, J.E. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Res.* **2017**, *27*, 1139–1152, doi:10.1101/gr.215160.116.

90. López-Perrote, A.; Alatwi, H.E.; Torreira, E.; Ismail, A.; Ayora, S.; Downs, J.A.; Llorca, O. Structure of Yin Yang 1 Oligomers That Cooperate with RuvBL1-RuvBL2 ATPases. *J. Biol. Chem.* **2014**, *289*, 22614–22629, doi:10.1074/jbc.M114.567040.

91. Li, L.; Williams, P.; Ren, W.; Wang, M.Y.; Gao, Z.; Miao, W.; Huang, M.; Song, J.; Wang, Y. YY1 interacts with guanine quadruplexes to regulate DNA looping and gene expression. *Nat. Chem. Biol.* **2020**, doi:10.1038/s41589-020-00695-1.

92. Wang, X.; Wang, S.; Troisi, E.C.; Howard, T.P.; Haswell, J.R.; Wolf, B.K.; Hawk, W.H.; Ramos, P.; Oberlick, E.M.; Tzvetkov, E.P.; et al. BRD9 defines a SWI/SNF sub-complex and constitutes a specific vulnerability in malignant rhabdoid tumors. *Nat. Commun.* **2019**, *10*, 1881, doi:10.1038/s41467-019-09891-7.

93. Zhou, X.; Xian, W.; Zhang, J.; Zhu, Y.; Shao, X.; Han, Y.; Qi, Y.; Ding, X.; Wang, X. YY1 binds to the E3′ enhancer and inhibits the expression of the immunoglobulin κ gene via epigenetic modifications. *Immunology* **2018**, *155*, 491–498, doi:10.1111/imm.12990.

94. Krivega, I.; Dean, A. Chromatin looping as a target for altering erythroid gene expression. *Ann. N. Y. Acad. Sci.* **2016**, *1368*, 31–39, doi:10.1111/nyas.13012.

95. Krivega, I.; Dale, R.K.; Dean, A. Role of LDB1 in the transition from chromatin looping to transcription activation. *Genes Dev.* **2014**, *28*, 1278–1290, doi:10.1101/gad.239749.114.

96. Deng, W.; Lee, J.; Wang, H.; Miller, J.; Reik, A.; Gregory, P.D.; Dean, A.; Blobel, G.A. Controlling Long-Range Genomic Interactions at a Native Locus by Targeted Tethering of a Looping Factor. *Cell* **2012**, *149*, 1233–1244, doi:10.1016/j.cell.2012.03.051.

97. Renko, M.; Fiedler, M.; Rutherford, T.J.; Schaefer, J. V.; Plückthun, A.; Bienz, M. Rotational symmetry of the structured Chip/LDB-SSDP core module of the Wnt enhanceosome. *Proc. Natl. Acad. Sci.* **2019**, *116*, 20977–20983, doi:10.1073/pnas.1912705116.

98. Wang, H.; Kim, J.; Wang, Z.; Yan, X.-X.; Dean, A.; Xu, W. Crystal structure of human LDB1 in complex with SSBP2. *Proc. Natl. Acad. Sci.* **2020**, *117*, 1042–1048,

doi:10.1073/pnas.1914181117.

99. Lee, J.; Krivega, I.; Dale, R.K.; Dean, A. The LDB1 Complex Co-opts CTCF for Erythroid Lineage-Specific Long-Range Enhancer Interactions. *Cell Rep.* **2017**, *19*, 2490–2502, doi:10.1016/j.celrep.2017.05.072.

100. Bashkirova, E.; Lomvardas, S. Olfactory receptor genes make the case for inter-chromosomal interactions. *Curr. Opin. Genet. Dev.* **2019**, *55*, 106–113, doi:10.1016/j.gde.2019.07.004.

101. Buck, L.; Axel, R. A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell* **1991**, *65*, 175–187, doi:10.1016/0092-8674(91)90418-X.

102. Degl'Innocenti, A.; D'Errico, A. Regulatory Features for Odorant Receptor Genes in the Mouse Genome. *Front. Genet.* **2017**, *8*, doi:10.3389/fgene.2017.00019.

103. Magklara, A.; Yen, A.; Colquitt, B.M.; Clowney, E.J.; Allen, W.; Markenscoff-Papadimitriou, E.; Evans, Z.A.; Kheradpour, P.; Mountoufaris, G.; Carey, C.; et al. An Epigenetic Signature for Monoallelic Olfactory Receptor Expression. *Cell* **2011**, *145*, 555–570, doi:10.1016/j.cell.2011.03.040.

104. Hanchate, N.K.; Kondoh, K.; Lu, Z.; Kuang, D.; Ye, X.; Qiu, X.; Pachter, L.; Trapnell, C.; Buck, L.B. Single-cell transcriptomics reveals receptor transformations during olfactory neurogenesis. *Science (80-.).* **2015**, *350*, 1251–1255, doi:10.1126/science.aad2456.

105. Sokpor, G.; Abbas, E.; Rosenbusch, J.; Staiger, J.F.; Tuoc, T. Transcriptional and Epigenetic Control of Mammalian Olfactory Epithelium Development. *Mol. Neurobiol.* **2018**, *55*, 8306–8327, doi:10.1007/s12035-018-0987-y.

106. Sharma, R.; Ishimaru, Y.; Davison, I.; Ikegami, K.; Chien, M.-S.; You, H.; Chi, Q.; Kubota, M.; Yohda, M.; Ehlers, M.; et al. Olfactory receptor accessory proteins play crucial roles in receptor function and gene choice. *Elife* **2017**, *6*, doi:10.7554/eLife.21895.

107. Ferreira, T.; Wilson, S.R.; Choi, Y.G.; Risso, D.; Dudoit, S.; Speed, T.P.; Ngai, J. Silencing of odorant receptor genes by G Protein βγ signaling ensures the expression of one odorant receptor per olfactory sensory neuron. *Neuron* **2014**, doi:10.1016/j.neuron.2014.01.001.

108. Khan, M.; Vaes, E.; Mombaerts, P. Regulation of the Probability of Mouse Odorant Receptor Gene Choice. *Cell* **2011**, *147*, 907–921, doi:10.1016/j.cell.2011.09.049.

109. Vassalli, A.; Feinstein, P.; Mombaerts, P. Homeodomain binding motifs modulate the probability of odorant receptor gene choice in transgenic mice. *Mol. Cell. Neurosci.* **2011**,

46, 381–396, doi:10.1016/j.mcn.2010.11.001.

110. Nishizumi, H.; Kumasaka, K.; Inoue, N.; Nakashima, A.; Sakano, H. Deletion of the core-H region in mice abolishes the expression of three proximal odorant receptor genes in cis. *Proc. Natl. Acad. Sci.* **2007**, *104*, 20067–20072, doi:10.1073/pnas.0706544105.

111. Bozza, T.; Vassalli, A.; Fuss, S.; Zhang, J.-J.; Weiland, B.; Pacifico, R.; Feinstein, P.; Mombaerts, P. Mapping of Class I and Class II Odorant Receptors to Glomerular Domains by Two Distinct Types of Olfactory Sensory Neurons in the Mouse. *Neuron* **2009**, *61*, 220–233, doi:10.1016/j.neuron.2008.11.010.

112. Markenscoff-Papadimitriou, E.; Allen, W.E.; Colquitt, B.M.; Goh, T.; Murphy, K.K.; Monahan, K.; Mosley, C.P.; Ahituv, N.; Lomvardas, S. Enhancer Interaction Networks as a Means for Singular Olfactory Receptor Expression. *Cell* **2014**, *159*, 543–557, doi:10.1016/j.cell.2014.09.033.

113. Monahan, K.; Horta, A.; Lomvardas, S. LHX2- and LDB1-mediated trans interactions regulate olfactory receptor choice. *Nature* **2019**, *565*, 448–453, doi:10.1038/s41586-018-0845-0.

114. Sigrist, C.J.; Pirrotta, V. Chromatin insulator elements block the silencing of a target gene by the Drosophila polycomb response element (PRE) but allow trans interactions between PREs on different chromosomes. *Genetics* **1997**, *147*, 209–221.

115. Kravchenko, E.; Savitskaya, E.; Kravchuk, O.; Parshikov, A.; Georgiev, P.; Savitsky, M. Pairing between gypsy insulators facilitates the enhancer action in trans throughout the Drosophila genome. *Mol Cell Biol* **2005**, *25*, 9283–9291, doi:10.1128/MCB.25.21.9283-9291.2005.

116. Muller, M.; Hagstrom, K.; Gyurkovics, H.; Pirrotta, V.; Schedl, P. The mcp element from the Drosophila melanogaster bithorax complex mediates long-distance regulatory interactions. *Genetics* **1999**, *153*, 1333–1356.

117. Li, H.B.; Muller, M.; Bahechar, I.A.; Kyrchanova, O.; Ohno, K.; Georgiev, P.; Pirrotta, V. Insulators, not Polycomb response elements, are required for long-range interactions between Polycomb targets in Drosophila melanogaster. *Mol Cell Biol* **2011**, *31*, 616–625, doi:10.1128/MCB.00849-10.

118. Vazquez, J.; Muller, M.; Pirrotta, V.; Sedat, J.W. The Mcp element mediates stable long-range chromosome-chromosome interactions in Drosophila. *Mol Biol Cell* **2006**, *17*, 2158–2165, doi:10.1091/mbc.E06-01-0049.

119. Kyrchanova, O. V; Georgiev, P.G. Functional interaction between an insulator and a Pc-dependent silencer with the example of the Mcp boundary of the Drosophila melanogaster Abd-B gene. *Genetika* **2010**, *46*, 593–603, doi:10.1134/S1022795410050029.

120. Bantignies, F.; Grimaud, C.; Lavrov, S.; Gabut, M.; Cavalli, G. Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. *Genes Dev* **2003**, *17*, 2406–2420, doi:10.1101/gad.269503.

121. Kyrchanova, O.; Wolle, D.; Sabirov, M.; Kurbidaeva, A.; Aoki, T.; Maksimenko, O.; Kyrchanova, M.; Georgiev, P.; Schedl, P. Distinct Elements Confer the Blocking and Bypass Functions of the Bithorax Fab-8 Boundary. *Genetics* **2019**, *213*, 865–876, doi:10.1534/genetics.119.302694.

122. Fujioka, M.; Mistry, H.; Schedl, P.; Jaynes, J.B. Determinants of Chromosome Architecture: Insulator Pairing in cis and in trans. *PLoS Genet.* **2016**, *2*, e1005889, doi:10.1371/journal.pgen.1005889.

123. Fujioka, M.; Wu, X.; Jaynes, J.B. A chromatin insulator mediates transgene homing and very long-range enhancer-promoter communication. *Development* **2009**, *136*, 3077–3087, doi:10.1242/dev.036467.

124. Hama, C.; Ali, Z.; Kornberg, T.B. Region-specific recombination and expression are directed by portions of the Drosophila engrailed promoter. *Genes Dev.* **1990**, *4*, 1079–1093, doi:10.1101/gad.4.7.1079.

125. Taillebourg, E.; Dura, J.-M. A novel mechanism for P element homing in Drosophila. *Proc. Natl. Acad. Sci.* **1999**, *96*, 6856–6861, doi:10.1073/pnas.96.12.6856.

126. Bender, W.; Hudson, A.; Barges, S.; Mihaly, J.; Galloni, M.; Hagstrom, K.; Muller, M.; Shanower, G.; Schedl, P.; Gyurkovics, H.; et al. P element homing to the Drosophila bithorax complex. *Development* **2000**, doi:10.1126/science.221.4605.23.

127. Mateo, L.J.; Murphy, S.E.; Hafner, A.; Cinquini, I.S.; Walker, C.A.; Boettiger, A.N. Visualizing DNA folding and RNA in embryos at single-cell resolution. *Nature* **2019**, *568*, 49–54, doi:10.1038/s41586-019-1035-4.

128. Scott, K.C.; Taubman, A.D.; Geyer, P.K. Enhancer blocking by the Drosophila gypsy insulator depends upon insulator anatomy and enhancer strength. *Genetics* **1999**, *153*, 787–798.

129. Holdridge, C.; Dorsett, D. Repression of hsp70 heat shock gene transcription by the suppressor of hairy-wing protein of Drosophila melanogaster. *Mol Cell Biol* **1991**, *11*,

1894–1900.

130. Geyer, P.K.; Corces, V.G. DNA position-specific repression of transcription by a Drosophila zinc finger protein. *Genes Dev* **1992**, *6*, 1865–1873.

131. Modolell, J.; Bender, W.; Meselson, M. Drosophila melanogaster mutations suppressible by the suppressor of Hairy-wing are insertions of a 7.3-kilobase mobile element. *Proc. Natl. Acad. Sci.* **1983**, *80*, 1678–1682, doi:10.1073/pnas.80.6.1678.

132. Schwartz, Y.B.; Linder-Basso, D.; Kharchenko, P. V; Tolstorukov, M.Y.; Kim, M.; Li, H.B.; Gorchakov, A.A.; Minoda, A.; Shanower, G.; Alekseyenko, A.A.; et al. Nature and function of insulator protein binding sites in the Drosophila genome. *Genome Res* **2012**, *22*, 2188–2198, doi:10.1101/gr.138156.112.

133. Soshnev, A.A.; He, B.; Baxley, R.M.; Jiang, N.; Hart, C.M.; Tan, K.; Geyer, P.K. Genome-wide studies of the multi-zinc finger Drosophila Suppressor of Hairy-wing protein in the ovary. *Nucleic Acids Res* **2012**, *40*, 5415–5431, doi:10.1093/nar/gks225.

134. Negre, N.; Brown, C.D.; Ma, L.; Bristow, C.A.; Miller, S.W.; Wagner, U.; Kheradpour, P.; Eaton, M.L.; Loriaux, P.; Sealfon, R.; et al. A cis-regulatory map of the Drosophila genome. *Nature* **2011**, *471*, 527–531, doi:10.1038/nature09990.

135. Kyrchanova, O.; Zolotarev, N.; Mogila, V.; Maksimenko, O.; Schedl, P.; Georgiev, P. Architectural protein Pita cooperates with dCTCF in organization of functional boundaries in Bithorax complex. *Development* **2017**, *144*, 2663–2672, doi:10.1242/dev.149815.

136. Baxley, R.M.; Bullard, J.D.; Klein, M.W.; Fell, A.G.; Morales-Rosado, J.A.; Duan, T.; Geyer, P.K. Deciphering the DNA code for the function of the Drosophila polydactyl zinc finger protein Suppressor of Hairy-wing. *Nucleic Acids Res.* **2017**, *45*, 4463–4478, doi:10.1093/nar/gkx040.

137. Maksimenko, O.; Bartkuhn, M.; Stakhov, V.; Herold, M.; Zolotarev, N.; Jox, T.; Buxa, M.K.; Kirsch, R.; Bonchuk, A.; Fedotova, A.; et al. Two new insulator proteins, Pita and ZIPIC, target CP190 to chromatin. *Genome Res* **2015**, *25*, 89–99, doi:10.1101/gr.174169.114.

138. Bender, W.; Lucas, M. The border between the ultrabithorax and abdominal-A regulatory domains in the Drosophila bithorax complex. *Genetics* **2013**, *193*, 1135–1147, doi:10.1534/genetics.112.146340.

139. Kyrchanova, O.; Maksimenko, O.; Ibragimov, A.; Sokolov, V.; Postika, N.; Lukyanova, M.; Schedl, P.; Georgiev, P. The insulator functions of the Drosophila polydactyl C2H2

zinc finger protein CTCF: Necessity versus sufficiency. *Sci. Adv.* **2020**, *6*, eaaz3152, doi:10.1126/sciadv.aaz3152.

140. Galloni, M.; Gyurkovics, H.; Schedl, P.; Karch, F. The bluetail transposon: evidence for independent cis-regulatory domains and domain boundaries in the bithorax complex. *EMBO J* **1993**, *12*, 1087–1097.

141. Hagstrom, K.; Muller, M.; Schedl, P. Fab-7 functions as a chromatin domain boundary to ensure proper segment specification by the Drosophila bithorax complex. *Genes Dev* **1996**, *10*, 3202–3215.

142. Zhou, J.; Barolo, S.; Szymanski, P.; Levine, M. The Fab-7 element of the bithorax complex attenuates enhancer-promoter interactions in the Drosophila embryo. *Genes Dev* **1996**, *10*, 3195–3201.

143. Wolle, D.; Cleard, F.; Aoki, T.; Deshpande, G.; Schedl, P.; Karch, F. Functional Requirements for Fab-7 Boundary Activity in the Bithorax. *Mol. Cell. Biol.* **2015**, doi:10.1128/MCB.00456-15.Address.

144. Kyrchanova, O.; Sabirov, M.; Mogila, V.; Kurbidaeva, A.; Postika, N.; Maksimenko, O.; Schedl, P.; Georgiev, P. Complete reconstitution of bypass and blocking functions in a minimal artificial Fab-7 insulator from Drosophila bithorax complex. *Proc. Natl. Acad. Sci.* **2019**, *116*, 13462–13467, doi:10.1073/pnas.1907190116.

145. Kyrchanova, O.; Chetverina, D.; Maksimenko, O.; Kullyev, A.; Georgiev, P. Orientation-dependent interaction between Drosophila insulators is a property of this class of regulatory elements. *Nucleic Acids Res* **2008**, *36*, 7019–7028, doi:10.1093/nar/gkn781.

146. Kyrchanova, O.; Ivlieva, T.; Toshchakov, S.; Parshikov, A.; Maksimenko, O.; Georgiev, P. Selective interactions of boundaries with upstream region of Abd-B promoter in Drosophila bithorax complex and role of dCTCF in this process. *Nucleic Acids Res.* **2011**, *39*, 3042–3052, doi:10.1093/nar/gkq1248.

147. Zolotarev, N.; Fedotova, A.; Kyrchanova, O.; Bonchuk, A.; Penin, A.A.; Lando, A.S.; Eliseeva, I.A.; Kulakovskiy, I. V.; Maksimenko, O.; Georgiev, P. Architectural proteins Pita, Zw5,and ZIPIC contain homodimerization domain and support specific long-range interactions in Drosophila. *Nucleic Acids Res.* **2016**, *44*, 7228–7241, doi:10.1093/nar/gkw371.

148. Bonchuk, A.; Maksimenko, O.; Kyrchanova, O.; Ivlieva, T.; Mogila, V.; Deshpande, G.; Wolle, D.; Schedl, P.; Georgiev, P. Functional role of dimerization and CP190 interacting domains of CTCF protein in Drosophila melanogaster. *BMC Biol* **2015**, *13*, 63,

doi:10.1186/s12915-015-0168-7.

149. Chung, H.R.; Schafer, U.; Jackle, H.; Bohm, S. Genomic expansion and clustering of ZAD-containing C2H2 zinc-finger genes in Drosophila. *EMBO Rep* **2002**, *3*, 1158–1162, doi:10.1093/embo-reports/kvf243.

150. Fedotova, A.A.; Bonchuk, A.N.; Mogila, V.A.; Georgiev, P.G. C2H2 zinc finger proteins: The largest but poorly explored family of higher eukaryotic transcription factors. *Acta Naturae* **2017**, *9*, 47–58.

151. Jauch, R.; Bourenkov, G.P.; Chung, H.R.; Urlaub, H.; Reidt, U.; Jackle, H.; Wahl, M.C. The zinc finger-associated domain of the Drosophila transcription factor grauzone is a novel zinc-coordinating protein-protein interaction module. *Structure* **2003**, *11*, 1393–1402.

152. Gaszner, M.; Vazquez, J.; Schedl, P. The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. *Genes Dev* **1999**, *13*, 2098–2107.

153. Blanton, J.; Gaszner, M.; Schedl, P. Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev* **2003**, *17*, 664–675, doi:10.1101/gad.1052003.

154. Zolotarev, N.A.; Maksimenko, O.G.; Georgiev, P.G.; Bonchuk, A.N. ZAD-Domain Is Essential for Nuclear Localization of Insulator Proteins in Drosophila melanogaster. *Acta Naturae* **2016**, *8*, 97–102.

155. Maksimenko, O.; Kyrchanova, O.; Klimenko, N.; Zolotarev, N.; Elizarova, A.; Bonchuk, A.; Georgiev, P. Small Drosophila zinc finger C2H2 protein with an N-terminal zinc finger-associated domain demonstrates the architecture functions. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2020**, *1863*, 194446, doi:10.1016/j.bbagrm.2019.194446.

156. Mohan, M.; Bartkuhn, M.; Herold, M.; Philippen, A.; Heinl, N.; Bardenhagen, I.; Leers, J.; White, R.A.; Renkawitz-Pohl, R.; Saumweber, H.; et al. The Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J* **2007**, *26*, 4203–4214, doi:10.1038/sj.emboj.7601851.

157. Melnikova, L.; Kostyuchenko, M.; Molodina, V.; Parshikov, A.; Georgiev, P.; Golovnin, A. Interactions between BTB domain of CP190 and two adjacent regions in Su(Hw) are required for the insulator complex formation. *Chromosoma* **2018**, *127*, 59–71, doi:10.1007/s00412-017-0645-6.

158. Melnikova, L.; Kostyuchenko, M.; Molodina, V.; Parshikov, A.; Georgiev, P.; Golovnin,

A. Multiple interactions are involved in a highly specific association of the Mod(mdg4)-67.2 isoform with the Su(Hw) sites in Drosophila. *Open Biol.* **2017**, *7*, 170150, doi:10.1098/rsob.170150.

159. Oliver, D.; Sheehan, B.; South, H.; Akbari, O.; Pai, C.Y. The chromosomal association/dissociation of the chromatin insulator protein Cp190 of Drosophila melanogaster is mediated by the BTB/POZ domain and two acidic regions. *BMC Cell Biol* **2010**, *11*, 101, doi:10.1186/1471-2121-11-101.

160. Plevock, K.M.; Galletta, B.J.; Slep, K.C.; Rusan, N.M. Newly Characterized Region of CP190 Associates with Microtubules and Mediates Proper Spindle Morphology in Drosophila Stem Cells. *PLoS One* **2015**, *10*, e0144174, doi:10.1371/journal.pone.0144174.

161. Bonchuk, A.; Denisov, S.; Georgiev, P.; Maksimenko, O. Drosophila BTB/POZ domains of “ttk group” can form multimers and selectively interact with each other. *J Mol Biol* **2011**, *412*, 423–436, doi:10.1016/j.jmb.2011.07.052.

162. Zollman, S.; Godt, D.; Privé, G.G.; Couderc, J.L.; Laski, F.A. The BTB domain, found primarily in zinc finger proteins, defines an evolutionarily conserved family that includes several developmentally regulated genes in Drosophila. *Proc. Natl. Acad. Sci. U. S. A.* **1994**.

163. Büchner, K.; Roth, P.; Schotta, G.; Krauss, V.; Saumweber, H.; Reuter, G.; Dorn, R. Genetic and molecular complexity of the position effect variegation modifier mod(mdg4) in Drosophila. *Genetics* **2000**, *155*, 141–57.

164. Vogelmann, J.; Le Gall, A.; Dejardin, S.; Allemand, F.; Gamot, A.; Labesse, G.; Cuvier, O.; Negre, N.; Cohen-Gonsaud, M.; Margeat, E.; et al. Chromatin insulator factors involved in long-range DNA interactions and their role in the folding of the Drosophila genome. *PLoS Genet* **2014**, *10*, e1004544, doi:10.1371/journal.pgen.1004544.

165. Ghosh, D. Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. *EMBO J.* **2001**, *20*, 2518–2527, doi:10.1093/emboj/20.10.2518.

166. Pai, C.Y.; Lei, E.P.; Ghosh, D.; Corces, V.G. The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol Cell* **2004**, *16*, 737–748, doi:10.1016/j.molcel.2004.11.004.

167. Melnikova, L.; Kostyuchenko, M.; Parshikov, A.; Georgiev, P.; Golovnin, A. Role of Su(Hw) zinc finger 10 and interaction with CP190 and Mod(mdg4) proteins in recruiting the Su(Hw) complex to chromatin sites in Drosophila. *PLoS One* **2018**, *13*, e0193497,

doi:10.1371/journal.pone.0193497.

168. Golovnin, A.; Mazur, A.; Kopantseva, M.; Kurshakova, M.; Gulak, P. V.; Gilmore, B.; Whitfield, W.G.F.; Geyer, P.; Pirrotta, V.; Georgiev, P. Integrity of the Mod(mdg4)-67.2 BTB Domain Is Critical to Insulator Function in Drosophila melanogaster. *Mol. Cell. Biol.* **2007**, *27*, 963–974, doi:10.1128/MCB.00795-06.

169. Cuartero, S.; Fresan, U.; Reina, O.; Planet, E.; Espinas, M.L. Ibf1 and Ibf2 are novel CP190-interacting proteins required for insulator function. *EMBO J.* **2014**, *33*, 637–647, doi:10.1002/embj.201386001.

170. Fedotova, A.; Clendinen, C.; Bonchuk, A.; Mogila, V.; Aoki, T.; Georgiev, P.; Schedl, P. Functional dissection of the developmentally restricted BEN domain chromatin boundary factor Insensitive. *Epigenetics Chromatin* **2019**, *12*, 2, doi:10.1186/s13072-018-0249-2.

171. Zolotarev, N.; Maksimenko, O.; Kyrchanova, O.; Sokolinskaya, E.; Osadchiy, I.; Girardot, C.; Bonchuk, A.; Ciglar, L.; Furlong, E.E.M.; Georgiev, P. Opbp is a new architectural/insulator protein required for ribosomal gene expression. *Nucleic Acids Res.* **2017**, *45*, 12285–12300, doi:10.1093/nar/gkx840.

172. Liu, G.; Dean, A. Enhancer long-range contacts: The multi-adaptor protein LDB1 is the tie that binds. *Biochim. Biophys. Acta - Gene Regul. Mech.* **2019**, *1862*, 625–633, doi:10.1016/j.bbagrm.2019.04.003.