Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Regulation of 3D chromatin organization by CTCF

Jian-Feng Xiang and Victor G Corces

Studies of nuclear architecture using chromosome conformation capture methods have provided a detailed view of how chromatin folds in the 3D nuclear space. New variants of this technology now afford unprecedented resolution and allow the identification of ever smaller folding domains that offer new insights into the mechanisms by which this organization is established and maintained. Here we review recent results in this rapidly evolving field with an emphasis on CTCF function, with the goal of gaining a mechanistic understanding of the principles by which chromatin is folded in the eukaryotic nucleus.

Address
Emory University School of Medicine, Department of Human Genetics, 615 Michael Street, Atlanta, GA 30322, USA

Corresponding author:
Corces, Victor G (vgcorces@gmail.com, vcorces@emory.edu)

Current Opinion in Genetics and Development 2021, **67**:33–40

This review comes from a themed issue on **Genome architecture and expression**

Edited by **Susan Gasser** and **Gerd Blobel**

https://doi.org/10.1016/j.gde.2020.10.005

Introduction

The eukaryotic genome is organized in the three-dimensional nuclear space in a manner that responds to and facilitates the regulation of nuclear processes. This organization therefore affects, and may be affected by, critical nuclear functions such as transcription, replication, recombination, and DNA repair. Until recently, insights into the 3D organization of the genetic material have come mostly from the use of microscopy, which has given us a broad view of nuclear architecture. Results from these studies suggest that actively transcribed regions interact to form various types of membraneless organelles such as transcription factories, regions of the genome containing H3K27me3 and Polycomb (Pc) complexes PRC1 and PRC2 congregate at Pc bodies, and regions containing H3K9me3 and HP1 such as centromeres come together to form chromocenters. Biochemical analyses of proteins involved in these contacts suggest that these structures are biomolecular condens formed by interactions among multivalent proteins [1]. The use of high throughput chromosome conformation capture methods has resulted in a more detailed view of chromatin 3D organization. Based on the original results in which the resolution of the Hi-C data ranged between 1 Mb and 50 kb, the general consensus in the field suggests that mammalian chromosomes are organized into large >1 Mb compartments, identified by PC1 after Principal Component Analysis (PCA), that can be classified into two types, A and B, which respectively correlate with active or inactive chromatin [2]. In addition to large compartments, chromosomes contain smaller Topologically Associating Domains (TADs) that can be identified using algorithms that detect changes in the directionality of interactions. A subset of TADs contains CTCF at the borders and correspond to CTCF loops whereas others are flanked by actively transcribed genes with or without CTCF [3]. Here we use the term loop to refer to a domain created by relatively stable point to point contacts, such as those mediated by CTCF/cohesin, and visible in Hi-C heatmaps as strong punctate signal at the summit of the domain. Loops are different from other contact domains formed by interactions among sequences within the domain mediated by proteins present in these sequences and corresponding to their transcriptional state. In this review we first consider the distinct contributions of transcriptional state and CTCF/cohesin to 3D chromatin organization. We then focus on mechanisms by which CTCF function is regulated to promote the formation of stable loops that can modulate enhancer–promoter communication.

3D nuclear organization, transcription, and CTCF loops

Results from high resolution Hi-C, around 1 kb for mammalian cells and 250 bp for *Drosophila*, suggest that eukaryotic genomes are organized into small contact domains containing one or several closely spaced genes in the same transcriptional state. Since these contact domains can be identified by PCA using bin sizes of 5−10 kb, they have been called compartment(al) domains to highlight their similarity to the original >1 Mb compartments [4,5]. Interactions among compartmental domains give rise to the plaid pattern observed in Hi-C heatmaps. This interaction pattern explains the formation of membraneless organelles or biomolecular condensates observed by microscopy and biochemical studies. Recent analyses of chromatin architecture using variations of the original Hi-C method have shed further light into the existence of small contact domains that correlate with transcriptional state and may represent the basic unit of chromosome organization. Micro-C employs micrococcal nuclease, rather than restriction enzymes, to digest the chromatin, thus allowing a

uniform nucleosome-size length of the digested fragments [6,7]. Micro-C allows the visualization of compartmental domains composed of single genes, including enhancer–promoter interactions, and involved in long-range interactions with other compartmental domains. When expressed, these domains are disrupted by inhibition of transcription, suggesting that they arise as a consequence of interactions among proteins involved in the transcription process [6,7], probably including initiation, splicing and termination. This may explain why simply depleting the different RNA polymerases has little effect on chromatin 3D organization [8], since it is possible that other components of the transcription machinery or proteins associated with covalent histone modifications may still remain on chromatin, although this has not been confirmed experimentally. Small contact domains that correlate with transcriptional state have also been observed using CAP-C, a derivative of Hi-C that uses multifunctional chemical crosslinkers to capture long-range chromatin interactions [9]. CAP-C achieves much higher resolution, allowing the identification of short-range interactions, which are present close to the diagonal of Hi-C heatmaps and not resolved by standard Hi-C. Using CAP-C, it has been possible to characterize loop domains that is, contact domains with strong punctate signal at the summit and identifiable by tools such as HiCCUPS. Loop domains are formed by cohesin extrusion, have convergent CTCF sites at their anchors, and are affected by CTCF depletion. Nonloop domains contain single genes, range in size between 10 kb and 40 kb, and are affected by transcription inhibition when present in the A compartment [9]. The independent contribution of CTCF and transcription to 3D chromatin organization has been recently tested in a series of elegant experiments involving either swapping the *Xist/Tsix* transcriptional units [10] or inserting a 2 kb sequence containing a CTCF site and a TSS [11]. The consequence of the presence of this 2 kb fragment in different genomic locations was analyzed by Hi-C. Depending on the genomic context, the inserted CTCF site can make loops with adjacent sites in the genome whereas the TSS creates new domains by forming directional contacts with downstream sequences. When the fragment is inserted in B compartment sequences, this TSS-dependent domain is able to form a new A compartment and establishes long-range contacts with other A-compartment domains, suggesting that domains induced by expression from the inserted TSS correspond to what we have referred to as compartmental domains. Deletion of CTCF or the TSS independently affect each type of domain. These results support previous observations in *Drosophila* suggesting that the basic units of chromosome organization are small compartmental domains that correspond to the transcriptional state of their sequences. When present in active chromatin, whether actively transcribing or paused, these domains correspond to single genes, whose sequences self-interact to form a gene loop. Adjacent, consecutively arranged active genes interact with each other to form larger compartmental domains. These domains alternate with non-transcribed sequences, which may contain H3K27me3, H3K9me3, or neither, and form self-interacting compartmental domains. Each of these types of domains interact with other domains in the same transcriptional state to give rise to the plaid pattern observed in Hi-C heatmaps away from the diagonal. These long-range interactions are heterogeneous in different cells of a population. Therefore, this organization is a consequence of the transcriptional state of sequences that is, an emergent property of the one-dimensional epigenetic information present in the chromatin fiber [12]. However, once this organization is established, it can affect gene expression by increasing the concentration of various factors in biomolecular condensates established in distinct nuclear compartments.

The regulation of the transcriptional state-dependent 3D organization of the genome takes place by well-established biochemical principles determined by the affinities of transcription factors for specific DNA sequences and for other proteins, which allow the recruitment of protein complexes to chromatin. Superimposed on this transcriptional state-dependent aspect of nuclear architecture is the phenomenon of cohesin extrusion, which takes place throughout the genome and can be stopped by CTCF. Interactions within and between compartmental domains are disrupted by CTCF and cohesin mediated loop extrusion [12]. The cohesin complex is loaded throughout the genome, or perhaps preferentially at NIPBL sites. Cohesin extrusion may be slowed down at genomic sites containing specific proteins or large protein complexes [6,7]. However, cohesin extrusion is stopped by sites bound by the CTCF protein, preferentially when the sites are arranged in a convergent orientation [13–15]. This retention of the cohesin complex at CTCF sites results in the formation of relatively stable loops that can be visualized in Hi-C heatmaps as strong puncta that represent interactions between two anchors containing CTCF. The strong punctate signal clearly visible above the surrounding background in Hi-C heatmaps suggests that these loops may be sufficiently stable to be simultaneously present in a majority of cells in a population [16]. The formation of these stable loops favors interactions between sequences inside of the loop and precludes interactions between sequences inside and outside of the loop, thus having the potential of regulating interactions among regulatory sequences in the genome. Therefore, a detailed understanding of the mechanisms by which CTCF/cohesin function can be regulated will be critical to understand this aspect of 3D chromatin organization.

Regulation of CTCF function by DNA methylation

CTCF is a ubiquitously expressed and highly conserved protein containing eleven zinc fingers (ZFs). CTCF binds to an approximately 41 bp sequence containing

four modules. A highly conserved 15 bp central core containing modules 2 and 3 is flanked by modules 1 and 4, which are much less conserved [17]. Only 6% of all CTCF sites in the genome contain all 4 modules, and 38% contain the core plus either module 1 or 4. The rest of CTCF genomic sites, around 66%, only contain the core motif [17]. This variability in the sequence of CTCF binding sites is interesting in the context of results describing the structure of the CTCF ZF domain revealed by X-ray crystallography [18,19]. Each ZFs 3–7 makes contacts with three bases in the DNA major groove of the core consensus [19], while ZF8 acts as a flexible spacer element to allow ZFs 9–11 to bind to module 1 in regions of the genome containing this sequence [18,19]. These results suggest that ZFs 1–2 and 9–11 may not be required for sequence-specific binding at most sites in the genome, opening the possibility that these ZFs may have alternative roles in CTCF function.

Approximately 30% of CTCF-bound sites are different among cell lines derived from different tissues [20], suggesting that the localization of CTCF in the genome may partially change during cell differentiation. In addition, it has been shown that CTCF occupancy can change in response to environmental stimuli, for example temperature or environmental contaminants such as chromium [21]. It is thus possible that alterations of CTCF binding play a role in regulating enhancer–promoter interactions during the establishment of specific cell fates. This may explain tissue-specific phenotypes observed when CTCF expression is altered during development. The mechanisms by which CTCF is recruited to specific sites in the genome during cell differentiation are not clear. A subset of CTCF sites contain CpGs and, therefore, CTCF binding may be susceptible to the methylation state of its target sequence. *In vitro* binding experiments indicate that cytosine methylation at position 2 of the core motif interferes with CTCF binding whereas methylation of the cytosine at position 12 has no effect [19], suggesting that CTCF would preferentially bind to sites unmethylated in the CpG at position 2 but still bind to sites containing 5mC at position 12. Site specific alteration of DNA methylation by precise genomic editing [22] or genome-wide disruption by depleting DNA methyltransferases [23] or dioxygenases [24] leads to changes in the CTCF binding pattern. In addition to directly inhibiting the interaction between CTCF and its binding site, the presence of DNA methylation also blocks CTCF binding by re-positioning nucleosomes [24]. At its binding sites, CTCF is in equilibrium with a fragile nucleosome containing H3.3 and H2A.Z [25,26], and it is flanked by approximately 10 well-positioned nucleosomes on each side [27]. It is likely that these positioned nucleosomes are a consequence, rather than a determinant, of CTCF binding. CTCF can probably bind to DNA immediately after the passage of the replication fork during S phase and before histone octamers have been deposited on nascent chromatin. Interestingly, CTCF sites are flanked by hemimethylated DNA at the DNA entry site into the nucleosome. Although this hemimethylated DNA does not appear to be required for DNA binding, its loss affects the ability of CTCF sites to form loops with other sites in the genome [23]. The presence of CTCF at unmethylated sites is necessary to maintain their unmethylated state. Downregulation of CTCF, or loss of one copy of CTCF in breast and prostate tumors [24], results in methylation of CTCF sites, which then may further interfere with binding of this protein. This effect should be limited to those CTCF sites in the genome containing CpG at position 2 in the core motif, which would limit the effect of CTCF copy number loss to the expression of a subset of genes proximal to these class of CTCF sites. This may explain the specific phenotypes caused by loss of one copy of CTCF in humans, which include autism, intellectual disability, and cleft palate [28,29].

Regulation of CTCF function by covalent modifications

Binding of CTCF to chromatin is also affected by several post-translational covalent modifications. Multiple sites in CTCF can be modified by poly(ADP-ribosyl)ation (PARylation), SUMOylation and phosphorylation [30–33]. PARylated CTCF was first shown to be present in the maternal H19 imprinted control region but also found in other parts of the genome. Inhibition of PARP-1 interferes with the ability of CTCF to restrict enhancer–promoter interactions [32]. These initial studies agree with results indicating an enrichment of PARylated CTCF in cells undergoing cell cycle arrest, which correlates with loss of CTCF at most sites in the genome [34]. Contrary to these results, PARP-1 stabilizes CTCF binding in the Epstein Barr virus genome [35]. Interestingly, PARylation of CTCF is metabolically controlled by the β-NAD + salvage pathway, pointing to possible regulatory events to control CTCF PARylation in cancer or development. In *Drosophila*, where CTCF is also PARylated, this modification is required to facilitate interactions between distant sites [36]. We suggest that the role of PARylation in the regulation of CTCF function merits further research, especially in the context of findings indicating that ATP is generated in the nucleus from poly(ADP-ribose) [37] and the requirement of ATP for cohesin extrusion [38].

SUMOylation of architectural proteins has been shown to affect their function in *Drosophila* [39] but the role of CTCF SUMOylation in controlling its interaction with chromatin or its ability to form loops has not been explored in detail in mammals. There are widespread changes in SUMOylation of CTCF sites during the heat shock response in human cells, and CTCF becomes de-SUMOylated during hypoxic stress, but the consequence

of these changes on the role of CTCF in chromatin 3D organization have not been explored [31,40].

The role of phosphorylation in CTCF function has been analyzed in more detail than that of other covalent modifications. Phosphorylated forms of CTCF are present during interphase and mitosis, suggesting that this modification may play different roles in CTCF function during the cell cycle. Early studies identified CK2 as a kinase able to phosphorylate S612 [41] but this initial work should be reanalyzed in the context of our current understanding of how CTCF affects gene expression. Phosphorylation of CTCF at T374 and S402, located in the linker space between the central ZFs, by LATS1 kinase interferes with CTCF binding to chromatin, which in turns affects expression of a specific subset of genes [42,43]. This kinase is normally cytoplasmic but translocates to the nucleus under glucose starvation and may do so under other conditions. Since LATS kinases are central components of the Hippo pathway, it will be interesting to explore whether signaling through this pathway is responsible for changes in the binding of CTCF to DNA during cell differentiation. Multiple Thr and Ser residues in CTCF are phosphorylated during mitosis [43,44]. Mitotic chromosomes have a unique structure determined by condensin loops and, therefore, the formation of this structure requires unraveling the standard interphase chromosome organization determined in large part by cohesin and CTCF [45]. Depending on the cell type, CTCF is present at low to very low levels in metaphase chromosomes [46–48], suggesting that regulation of CTCF binding by phosphorylation may represent a mechanism to ensure depletion of this protein in mitotic chromatin. This takes place at T289, T317, T346, T374, S402, S461, and T518, all residues located in the linker region of different ZFs [42,43]. Phosphorylation of CTCF at S224 by Polo-like kinase 1 during the G2/M transition affects the expression of hundreds of genes without affecting mitosis or chromatin organization [43,44]. This Ser residue is located in the amino-terminal domain of CTCF, immediately adjacent to the YDF motif that interacts with cohesin, and it would be interesting to study whether cohesin extrusion is affected by changes in phosphorylation levels of S224.

Regulation of CTCF function by interaction with other proteins

Although most studies to date have focused on the analysis of the canonical CTCF protein, human cells also produce other isoforms whose roles have not been explored in detail. In particular, a short CTCF isoform is encoded by an alternatively spliced transcript lacking exons 3 and 4 [49]. This isoforms lacks the N-terminal region responsible for interactions with cohesin and ZFs 1–2 [49]. CTCF genomic sites containing the CTCF short isoform fail to stop cohesin extrusion, leading to genome-wide decrease in cohesin occupancy [49,50]. It is unclear whether differential expression of the CTCF short isoform during development or in response to stimuli represents a strategy used by cells to regulate CTCF function. A similar role may be played by the CTCF homologue CTCFL, which is transiently expressed during germline development but at very low levels in normal somatic tissues. CTCFL binds to a very similar sequence than CTCF and, if expressed at higher levels, as it does in many cancers, could compete with CTCF for binding at a subset of sites [51,52]. Since CTCFL lacks the domain required for cohesin interaction, this protein is unable to stop cohesin extrusion and to form stable loops between distant sites in the genome [53].

Several different proteins have been shown to co-localize with CTCF but their involvement in the regulation of CTCF function has not been studied in detail until recently. For example, AP-1 has been shown to be recruited to CTCF sites involved in gained interactions responsible for changes in gene expression during macrophage development and in the response of hESCs to temperature stress [21,54]. A subset of Nucleoporin 153 (NUP153) sites in mESCs are present within 5 kb of CTCF sites and depletion of NUP153 leads to the loss of CTCF and cohesin binding to chromatin [55]. Given the genomic distance between the location of these proteins, it is unclear whether the effect of NUP153 is direct or indirect. As described above, CTCF is in equilibrium with a fragile nucleosome at its binding sites, and surrounded by 10 well-positioned nucleosomes on each side. It is unclear whether CTCF binding requires positioned nucleosomes flanking its binding site or, once bound, CTCF positions the flanking nucleosomes. Recent results suggest that binding of CTCF requires the ISWI complex, and depletion of the SNF2H component of ISWI results in loss of CTCF from chromatin [56]. In addition to facilitating CTCF binding, some proteins regulate CTCF function by interfering with its recruitment to DNA. An example is the ChAHP complex, which is composed of CHD4, ADNP, and HP1 [54]. ADNP is a ZF protein that recruits the ChAHP complex to specific sites in the genome to repress expression of genes involved in lineage commitment by mechanisms that do not involve H3K9me3. The ChAHP complex competes with CTCF for binding to a subset of genomic sites. Since ChAHP is unable to stop cohesin extrusion, the presence of this complex interferes with the formation of a subset of CTCF loops [57]. Interestingly, a role in facilitating loop formation has also been proposed for the ADNP protein, although in this case it is unclear whether ADNP participates with the rest of the ChAHP complex. Another protein that cooperates with CTCF in the formation of loops is TFIIIC, which is recruited to Alu elements located near cell cycle genes in response to serum starvation. TFIIIC then interacts with CTCF located at cell cycle gene promoters and these existing

loops rapidly activate gene expression upon serum exposure [58].

Regulation of CTCF function by RNA

CTCF not only interacts with DNA and other proteins, but also binds RNA and this interaction is important to regulate CTCF function [59,60]. In fact, the affinity of CTCF for RNA is an order of magnitude greater than for DNA. However, CTCF does not bind specific RNAs. Instead, the RNA interactome of CTCF is composed of thousands of transcripts from genes located in close proximity to CTCF binding sites [59,60]. Interactions between CTCF and RNA take place through ZF1, ZF10, ZF11, and 38 amino acids in the C-terminus following the ZF domain [61,62]. As described above, these ZFs are not involved in the recognition of the core CTCF motif. Deletion of the RNA-binding domains of CTCF results in decreased binding to DNA and cohesin at a subset of genomic regions and a decrease in looping interactions mediated by these anchors [61,62]. Since RNA is required to mediate CTCF–CTCF interactions, it has been suggested that the role of RNA at CTCF sites in the genome is to promote the formation of large CTCF aggregates that can help slow down cohesin extrusion to ensure orientation-dependent CTCF–cohesin interactions (see below). However, it is unclear why deletion of the RNA binding domain does not affect all genomic sites involved in loop formation. The specific nature of the RNAs involved in this process or the basis for selectivity are not known. Recently developed methods such as RD-SPRITE [63] to simultaneously measure all contacts among RNA and DNA combined with an immunoprecipitation step using antibodies to CTCF should be able to address these questions. It is tempting to speculate that RNA modifications might contribute to the specificity of CTCF–RNA interactions in view of results showing that m^6A modification of RNA can regulate chromatin accessibility [64].

Regulation of CTCF function by control of cohesin extrusion

An important role of CTCF appears to be to stop cohesin extrusion. In doing so, CTCF makes anchors of loops that have one of two functions — they either insulate genes and regulatory sequences inside the loop from those located outside or they tether regulatory sequences to their cognate promoters. Cohesin loads either at NIPBL sites or randomly throughout the sequences that eventually will form the loop. Cohesin will then extrude a loop until it encounters two CTCF sites arranged in a convergent head-to-head orientation, a phenomenon that is supported by computational modelling [13,65] and *in vitro* single molecule studies [66–68]. Loop extrusion is an evolutionarily conserved mechanism by which cells segregate interphase and mitotic chromatin, a conclusion supported by evidence in *Bacillus subtilis* [69], *yeast* [70], *Xenopus* egg [71], mouse [72] and human cells [68]. Depletion of either CTCF or the cohesin subunit Rad21 eliminates CTCF loops from the genome based on Hi-C analyses [4,72]. Recent results illustrate the structural basis for the CTCF–cohesin interaction [50]. A region in the N-terminus of CTCF containing a YDF motif interacts with a pocket formed by the SA2-SCC1 subunits of cohesin. Interestingly, a similar sequence is present in the cohesin release factor WAPL, suggesting that perhaps the role of the CTCF–cohesin interaction is to prevent unloading of cohesin by WAPL [50]. As predicted by this model, CTCF mutants lacking the N-terminus including the YDF motif cannot form loops [73,74]. The requirement for such a precise interaction between two very specific and relatively small domains have led to the suggestion that stopping cohesin extrusion by CTCF should require additional mechanisms to slow down the movement of the cohesin complex [75]. The flanking arrays of positioned nucleosomes, the formation of an unusual DNA structure as a consequence of DNA binding by CTCF, association with RNA, covalent modifications by bulky SUMO or poly-ADP-ribose, or formation of CTCF aggregates have all been suggested as candidates to pause cohesin extrusion and allow specific interactions between the CTCF and SA2-SCC1 binding interfaces [75]. In this context, it is interesting that Top2B interacts with CTCF and localizes to genomic sites for this protein such that the order at loop anchors is Top2B–CTCF–cohesin that is, Top2B is outside of the loop whereas cohesin is inside [76]. Although there is no evidence for cohesin extrusion causing DNA supercoiling, condensin is capable of introducing positive supercoils into DNA in an ATP-dependent manner [77,78]. While speculative at this time, it may be possible that supercoils resulting from transcription [79] or cohesin extrusion accumulate at the inside of loop anchors, and that this contributes to slowing cohesin extrusion to allow it to interact with CTCF.

Cohesin is a large ring-shaped protein complex with several alternative subunits subject to covalent modifications, both of which can serve as regulatory steps of cohesin extrusion during interphase to control the type and stability of loops. In addition to the core SMC3, SMC1A, and RAD21 subunits, the cohesin complex present in somatic cells contains one of two STAG subunits, STAG1 and STAG2. Cohesin complexes containing either of these two subunits have overlapping but also distinct binding locations on chromatin and they form different sized loops [80,81]. STAG1 can be modified by the acetyltransferase ESCO1, and this acetylation protects cohesin from release by WAPL. As a consequence, acetylated STAG1-containing loops are longer and stable for periods of several hours while those formed by SATG2 are shorter and stable for minutes [82]. This may be important during cell differentiation when stabilization of loops for long period of times may allow cells in

a population to precisely regulate the expression of specific genes in time.

Conclusions and perspectives

The 3D organization of the chromatin fiber in the eukaryotic nucleus is the result of two distinct components. One component is the interactions that take place among large protein complexes involved in the activation or repression of gene expression. This aspect of 3D organization is a consequence of the transcriptional state of specific regions of the genome and it is predictable that is, can be reproduced by computer algorithms with one-dimensional epigenetic information as the sole input. The second component is cohesin extrusion and its peculiar interaction with CTCF and, to a lesser extent, other components of the transcription machinery. This aspect of 3D nuclear organization is not predictable because we lack an understanding of the rules that allow some CTCF sites to stop cohesin extrusion whereas other sites in the same orientation are unable to do so. Dissection of the mechanisms responsible for these rules should be an area of special interest in the near future.

Funding

Funding was received for this work.

Conflict of interest statement

Nothing declared.

Acknowledgements

Work in the authors' laboratory was supported by U.S. Public Health Service AwardR01 GM035463 from the National Institutes of Health. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Sabari BR, Dall'Agnese A, Young RA: **Biomolecular condensates in the nucleus**. *Trends Biochem Sci* 2020, **45**:961-977 http://dx.doi.org/10.1016/j.tibs.2020.1006.1007.

2. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO *et al.*: **Comprehensive mapping of long-range interactions reveals folding principles of the human genome**. *Science* 2009, **326**:289-293.

3. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B: **Topological domains in mammalian genomes identified by analysis of chromatin interactions**. *Nature* 2012, **485**:376-380.

4. Rao SSP, Huang SC, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon KR, Sanborn AL, Johnstone SE, Bascom GD, Bochkov ID *et al.*: **Cohesin loss eliminates all loop domains**. *Cell* 2017, **171**:305-320 e324.

5. Rowley MJ, Nichols MH, Lyu X, Ando-Kuri M, Rivera ISM, Hermetz K, Wang P, Ruan Y, Corces VG: **Evolutionarily conserved principles predict 3D chromatin organization**. *Mol Cell* 2017, **67**:837-852 e837.

6. Hsieh TS, Fudenberg G, Goloborodko A, Rando OJ: **Micro-C XL: assaying chromosome conformation from the nucleosome to the entire genome**. *Nat Methods* 2016, **13**:1009-1011.

7. Hsieh TS, Cattoglio C, Slobodyanyuk E, Hansen AS, Rando OJ, Tjian R, Darzacq X: **Resolving the 3D landscape of transcription-linked mammalian chromatin folding**. *Mol Cell* 2020, **78**:539-553.e538.

8. Jiang Y, Huang J, Lun K, Li B, Zheng H, Li Y, Zhou R, Duan W, Wang C, Feng Y *et al.*: **Genome-wide analyses of chromatin interactions after the loss of Pol I, Pol II, and Pol III**. *Genome Biol* 2020, **21**:158.

9. You Q, Cheng AY, Gu X, Harada BT, Yu M, Wu T, Ren B, Ouyang Z, He C: **Direct DNA crosslinking with CAP-C uncovers transcription-dependent chromatin organization at high resolution**. *Nat Biotechnol* 2020 http://dx.doi.org/10.1038/s41587-41020-40643-41588.

10. van Bemmel JG, Galupa R, Gard C, Servant N, Picard C, Davies J, Szempruch AJ, Zhan Y, Zylicz JJ, Nora EP *et al.*: **The bipartite TAD organization of the X-inactivation center ensures opposing developmental regulation of Tsix and Xist**. *Nat Genet* 2019, **51**:1024-1034.

11. Zhang D, Huang P, Sharma M, Keller CA, Giardine B, Zhang H, Gilgenast TG, Phillips-Cremins JE, Hardison RC, Blobel GA: **Alteration of genome folding via contact domain boundary insertion**. *Nat Genet* 2020 http://dx.doi.org/10.1038/s41588-41020-40680-41588.

12. Rowley MJ, Corces VG: **Organizational principles of 3D genome architecture**. *Nat Rev Genet* 2018, **19**:789-800.

13. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA: **Formation of chromosomal domains by loop extrusion**. *Cell Rep* 2016, **15**:2038-2049.

14. Nichols MH, Corces VG: **A CTCF code for 3D genome architecture**. *Cell* 2015, **162**:703-705.

15. Sanborn AL, Rao SS, Huang SC, Durand NC, Huntley MH, Jewett AI, Bochkov ID, Chinnappan D, Cutkosky A, Li J *et al.*: **Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes**. *Proc Natl Acad Sci U S A* 2015, **112**:E6456-E6465.

16. Hansen AS, Pustova I, Cattoglio C, Tjian R, Darzacq X: **CTCF and cohesin regulate chromatin loop stability with distinct dynamics**. *eLife* 2017, **6**.

17. Rhee HS, Pugh BF: **Comprehensive genome-wide protein-DNA interactions detected at single-nucleotide resolution**. *Cell* 2011, **147**:1408-1419.

18. Yin M, Wang J, Wang M, Li X, Zhang M, Wu Q, Wang Y: **Molecular mechanism of directional CTCF recognition of a diverse range of genomic sites**. *Cell Res* 2017, **27**:1365-1377.

19. Hashimoto H, Wang D, Horton JR, Zhang X, Corces VG, Cheng X: **Structural basis for the versatile and methylation-dependent binding of CTCF to DNA**. *Mol Cell* 2017, **66**:711-720 e713.

20. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, Pauli F, Lee K, Canfield T, Weaver M, Sandstrom R *et al.*: **Widespread plasticity in CTCF occupancy linked to DNA methylation**. *Genome Res* 2012, **22**:1680-1688.

21. Lyu X, Rowley MJ, Corces VG: **Architectural proteins and pluripotency factors cooperate to orchestrate the transcriptional response of hESCs to temperature stress**. *Mol Cell* 2018, **71**:940-955.e947.

22. Tarjan DR, Flavahan WA, Bernstein BE: **Epigenome editing strategies for the functional annotation of CTCF insulators**. *Nat Commun* 2019, **10**:4258.

23. Xu C, Corces VG: **Nascent DNA methylome mapping reveals inheritance of hemimethylation at CTCF/cohesin sites**. *Science* 2018, **359**:1166-1170.

24. Wiehle L, Thorn GJ, Raddatz G, Clarkson CT, Rippe K, Lyko F, Breiling A, Teif VB: **DNA (de)methylation in embryonic stem cells controls CTCF-dependent chromatin boundaries**. *Genome Res* 2019, **29**:750-761.

25. Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G: **H3.3/H2A.Z double variant-containing nucleosomes mark' nucleosome-free regions' of active promoters and other regulatory regions**. *Nat Genet* 2009, **41**:941-945.

26. Voong LN, Xi L, Sebeson AC, Xiong B, Wang JP, Wang X: **Insights into nucleosome organization in mouse embryonic stem cells through chemical mapping**. *Cell* 2016, **167**:1555-1570 e1515.

27. Owens N, Papadopoulou T, Festuccia N, Tachtsidi A, Gonzalez I, Dubois A, Vandormael-Pournin S, Nora EP, Bruneau BG, Cohen-Tannoudji M *et al.*: **CTCF confers local nucleosome resiliency after DNA replication and during mitosis**. *eLife* 2019, **8**:e47898.

28. Konrad EDH, Nardini N, Caliebe A, Nagel I, Young D, Horvath G, Santoro SL, Shuss C, Ziegler A, Bonneau D *et al.*: **CTCF variants in 39 individuals with a variable neurodevelopmental disorder broaden the mutational and clinical spectrum**. *Genet Med* 2019, **21**:2723-2733.

29. Hori I, Kawamura R, Nakabayashi K, Watanabe H, Higashimoto K, Tomikawa J, Ieda D, Ohashi K, Negishi Y, Hattori A *et al.*: **CTCF deletion syndrome: clinical features and epigenetic delineation**. *J Med Genet* 2017, **54**:836-842.

30. MacPherson MJ, Beatty LG, Zhou W, Du M, Sadowski PD: **The CTCF insulator protein is posttranslationally modified by SUMO**. *Mol Cell Biol* 2009, **29**:714-725.

31. Wang J, Wang Y, Lu L: **De-SUMOylation of CCCTC binding factor (CTCF) in hypoxic stress-induced human corneal epithelial cells**. *J Biol Chem* 2012, **287**:12469-12479.

32. Yu W, Ginjala V, Pant V, Chernukhin I, Whitehead J, Docquier F, Farrar D, Tavoosidana G, Mukhopadhyay R, Kanduri C *et al.*: **Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation**. *Nat Genet* 2004, **36**:1105-1110.

33. Kitchen NS, Schoenherr CJ: **Sumoylation modulates a domain in CTCF that activates transcription and decondenses chromatin**. *J Cell Biochem* 2010, **111**:665-675.

34. Pavlaki I, Docquier F, Chernukhin I, Kita G, Gretton S, Clarkson CT, Teif VB, Klenova E: **Poly(ADP-ribosyl)ation associated changes in CTCF-chromatin binding and gene expression in breast cells**. *Biochim Biophys Acta Gene Regul Mech* 2018, **1861**:718-730.

35. Lupey-Green LN, Caruso LB, Madzo J, Martin KA, Tan Y, Hulse M, Tempera I: **PARP1 stabilizes CTCF binding and chromatin structure to maintain Epstein-Barr virus latency type**. *J Virol* 2018, **92**:e00755-e00718.

36. Ong CT, Van Bortle K, Ramos E, Corces VG: **Poly(ADP-ribosyl)ation regulates insulator function and intrachromosomal interactions in *Drosophila***. *Cell* 2013, **155**:148-159.

37. Wright RH, Lioutas A, Le Dily F, Soronellas D, Pohl A, Bonet J, Nacht AS, Samino S, Font-Mateu J, Vicent GP *et al.*: **ADP-ribose-derived nuclear ATP synthesis by NUDIX5 is required for chromatin remodeling**. *Science* 2016, **352**:1221-1225.

38. Vian L, Pekowska A, Rao SSP, Kieffer-Kwon KR, Jung S, Baranello L, Huang SC, El Khattabi L, Dose M, Pruett N *et al.*: **The energetics and physiological impact of cohesin extrusion**. *Cell* 2018, **175**:292-294.

39. Capelson M, Corces VG: **SUMO conjugation attenuates the activity of the gypsy chromatin insulator**. *EMBO J* 2006, **25**:1906-1914.

40. Niskanen EA, Malinen M, Sutinen P, Toropainen S, Paakinaho V, Vihervaara A, Joutsen J, Kaikkonen MU, Sistonen L, Palvimo JJ: **Global SUMOylation on active chromatin is an acute heat stress response restricting transcription**. *Genome Biol* 2015, **16**:153.

41. El-Kady A, Klenova E: **Regulation of the transcription factor, CTCF, by phosphorylation with protein kinase CK2**. *FEBS Lett* 2005, **579**:1424-1434.

42. Luo H, Yu Q, Liu Y, Tang M, Liang M, Zhang D, Xiao TS, Wu L, Tan M, Ruan Y *et al.*: **LATS kinase-mediated CTCF phosphorylation and selective loss of genomic binding**. *Sci Adv* 2020, **6**:eaaw4651.

43. Sekiya T, Murano K, Kato K, Kawaguchi A, Nagata K: **Mitotic phosphorylation of CCCTC-binding factor (CTCF) reduces its DNA binding activity**. *FEBS Open Bio* 2017, **7**:397-404.

44. Del Rosario BC, Kriz AJ, Del Rosario AM, Anselmo A, Fry CJ, White FM, Sadreyev RI, Lee JT: **Exploration of CTCF post-translation modifications uncovers Serine-224 phosphorylation by PLK1 at pericentric regions during the G2/M transition**. *eLife* 2019, **8**:e42341.

45. Gibcus JH, Samejima K, Goloborodko A, Samejima I, Naumova N, Nuebler J, Kanemaki MT, Xie L, Paulson JR, Earnshaw WC *et al.*: **A pathway for mitotic chromosome formation**. *Science* 2018, **359**:eaao6135.

46. Oomen ME, Hansen AS, Liu Y, Darzacq X, Dekker J: **CTCF sites display cell cycle-dependent dynamics in factor binding and nucleosome positioning**. *Genome Res* 2019, **29**:236-249.

47. Owens N, Papadopoulou T, Festuccia N, Tachtsidi A, Gonzalez I, Dubois A, Vandormael-Pournin S, Nora EP, Bruneau BG, Cohen-Tannoudji M *et al.*: **CTCF confers local nucleosome resiliency after DNA replication and during mitosis**. *eLife* 2019, **8**.

48. Zhang H, Emerson DJ, Gilgenast TG, Titus KR, Lan Y, Huang P, Zhang D, Wang H, Keller CA, Giardine B *et al.*: **Chromatin structure dynamics during the mitosis-to-G1 phase transition**. *Nature* 2019, **576**:158-162.

49. Li J, Huang K, Hu G, Babarinde IA, Li Y, Dong X, Chen YS, Shang L, Guo W, Wang J *et al.*: **An alternative CTCF isoform antagonizes canonical CTCF occupancy and changes chromatin architecture to promote apoptosis**. *Nat Commun* 2019, **10**:1535.

50. Li Y, Haarhuis JHI, Sedeno Cacciatore A, Oldenkamp R, van Ruiten MS, Willems L, Teunissen H, Muir KW, de Wit E, Rowland BD *et al.*: **The structural basis for cohesin-CTCF-anchored loops**. *Nature* 2020, **578**:472-476.

51. Nishana M, Ha C, Rodriguez-Hernaez J, Ranjbaran A, Chio E, Nora EP, Badri SB, Kloetgen A, Bruneau BG, Tsirigos A *et al.*: **Defining the relative and combined contribution of CTCF and CTCFL to genomic regulation**. *Genome Biol* 2020, **21**:108.

52. Debaugny RE, Skok JA: **CTCF and CTCFL in cancer**. *Curr Opin Genet Dev* 2020, **61**:44-52.

53. Vara C, Paytuvi-Gallart A, Cuartero Y, Le Dily F, Garcia F, Salva-Castro J, Gomez HL, Julia E, Moutinho C, Aiese Cigliano R *et al.*: **Three-dimensional genomic structure and cohesin occupancy correlate with transcriptional activity during spermatogenesis**. *Cell Rep* 2019, **28**:352-367 e359.

54. Phanstiel DH, Van Bortle K, Spacek D, Hess GT, Shamim MS, Machol I, Love MI, Aiden EL, Bassik MC, Snyder MP: **Static and dynamic DNA loops form AP-1-bound activation hubs during macrophage development**. *Mol Cell* 2017, **67**:1037-1048 e1036.

55. Kadota S, Ou J, Shi Y, Lee JT, Sun J, Yildirim E: **Nucleoporin 153 links nuclear pore complex to chromatin architecture by mediating CTCF and cohesin binding**. *Nat Commun* 2020, **11**:2606.

56. Barisic D, Stadler MB, Iurlaro M, Schubeler D: **Mammalian ISWI and SWI/SNF selectively mediate binding of distinct transcription factors**. *Nature* 2019, **569**:136-140.

57. Kaaij LJT, Mohn F, van der Weide RH, de Wit E, Buhler M: **The ChAHP complex counteracts chromatin looping at CTCF sites that emerged from SINE expansions in mouse**. *Cell* 2019, **178**:1437-1451 e1414.

58. Ferrari R, de Llobet Cucalon LI, Di Vona C, Le Dilly F, Vidal E, Lioutas A, Oliete JQ, Jochem L, Cutts E, Dieci G *et al.*: **TFIIIC binding to Alu elements controls gene expression via chromatin looping and histone acetylation**. *Mol Cell* 2020, **77**:475-487 e411.

59. Kung JT, Kesner B, An JY, Ahn JY, Cifuentes-Rojas C, Colognori D, Jeon Y, Szanto A, del Rosario BC, Pinter SF *et al.*: **Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF**. *Mol Cell* 2015, **57**:361-375.

60. Saldana-Meyer R, Gonzalez-Buendia E, Guerrero G, Narendra V, Bonasio R, Recillas-Targa F, Reinberg D: **CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53**. *Genes Dev* 2014, **28**:723-734.

61. Hansen AS, Hsieh TS, Cattoglio C, Pustova I, Saldaña-Meyer R, Reinberg D, Darzacq X, Tjian R: **Distinct classes of chromatin**

loops revealed by deletion of an RNA-binding region in CTCF. *Mol Cell* 2019, **76**:395-411.e313.

62. Saldana-Meyer R, Rodriguez-Hernaez J, Escobar T, Nishana M, Jacome-Lopez K, Nora EP, Bruneau BG, Tsirigos A, Furlan-Magaril M, Skok J *et al.*: **RNA interactions are essential for CTCF-mediated genome organization**. *Mol Cell* 2019, **76**:412-422.e415.

63. Quinodoz SA, Bhat P, Ollikainen N, Jachowicz JW, Banerjee AK, Chovanec P, Blanco MR, Chow A, Markaki Y, Plath K *et al.*: **RNA promotes the formation of spatial compartments in the nucleus**. *bioRxiv* 2020 http://dx.doi.org/10.1101/2020.1108.1125.267435.

64. Liu J, Dou X, Chen C, Chen C, Liu C, Xu MM, Zhao S, Shen B, Gao Y, Han D *et al.*: **N (6)-methyladenosine of chromosome-associated regulatory RNA regulates chromatin state and transcription**. *Science* 2020, **367**:580-586.

65. Nuebler J, Fudenberg G, Imakaev M, Abdennur N, Mirny LA: **Chromatin organization by an interplay of loop extrusion and compartmental segregation**. *Proc Natl Acad Sci U S A* 2018, **115**:E6697-E6706.

66. Davidson IF, Bauer B, Goetz D, Tang W, Wutz G, Peters JM: **DNA loop extrusion by human cohesin**. *Science* 2019, **366**:1338-1345.

67. Ganji M, Shaltiel IA, Bisht S, Kim E, Kalichava A, Haering CH, Dekker C: **Real-time imaging of DNA loop extrusion by condensin**. *Science* 2018, **360**:102-105.

68. Kim Y, Shi Z, Zhang H, Finkelstein IJ, Yu H: **Human cohesin compacts DNA by loop extrusion**. *Science* 2019, **366**:1345-1349.

69. Wang X, Hughes AC, Brandão HB, Walker B, Lierz C, Cochran JC, Oakley MG, Kruse AC, Rudner DZ: **In vivo evidence for ATPase-dependent DNA translocation by the *Bacillus subtilis* SMC condensin complex**. *Mol Cell* 2018, **71**:841-847.e845.

70. Dauban L, Montagne R, Thierry A, Lazar-Stefanita L, Bastie N, Gadal O, Cournac A, Koszul R, Beckouet F: **Regulation of cohesin-mediated chromosome folding by Eco1 and other partners**. *Mol Cell* 2020, **77**:1279-1293 e1274.

71. Golfier S, Quail T, Kimura H, Brugues J: **Cohesin and condensin extrude DNA loops in a cell cycle-dependent manner**. *eLife* 2020, **9**:e53885.

72. Nora EP, Goloborodko A, Valton AL, Gibcus JH, Uebersohn A, Abdennur N, Dekker J, Mirny LA, Bruneau BG: **Targeted degradation of CTCF decouples local insulation of chromosome domains from genomic compartmentalization**. *Cell* 2017, **169**:930-944 e922.

73. Nora EP, Caccianini L, Fudenberg G, Kameswaran V, Nagle A, Uebersohn A, So K, Hajj B, Le Saux A, Coulon A *et al.*: **Molecular basis of CTCF binding polarity in genome folding**. *bioRxiv* 2019 http://dx.doi.org/10.1101/2019.1112.1113.876177.

74. Pugacheva EM, Kubo N, Loukinov D, Tajmul M, Kang S, Kovalchuk AL, Strunnikov AV, Zentner GE, Ren B, Lobanenkov VV: **CTCF mediates chromatin looping via N-terminal domain-dependent cohesin retention**. *Proc Natl Acad Sci U S A* 2020, **117**:2020-2031.

75. Hansen AS: **CTCF as a boundary factor for cohesin-mediated loop extrusion: evidence for a multi-step mechanism**. *Nucleus* 2020, **11**:132-148.

76. Uuskula-Reimand L, Hou H, Samavarchi-Tehrani P, Rudan MV, Liang M, Medina-Rivera A, Mohammed H, Schmidt D, Schwalie P, Young EJ *et al.*: **Topoisomerase II beta interacts with cohesin and CTCF at topological domain borders**. *Genome Biol* 2016, **17**:182.

77. Bazett-Jones DP, Kimura K, Hirano T: **Efficient supercoiling of DNA by a single condensin complex as revealed by electron spectroscopic imaging**. *Mol Cell* 2002, **9**:1183-1190.

78. Kimura K, Cuvier O, Hirano T: **Chromosome condensation by a human condensin complex in *Xenopus* egg extracts**. *J Biol Chem* 2001, **276**:5417-5420.

79. Bjorkegren C, Baranello L: **DNA supercoiling, topoisomerases, and cohesin: partners in regulating chromatin architecture?** *Int J Mol Sci* 2018, **19**:884.

80. Casa V, Moronta Gines M, Gade Gusmao E, Slotman JA, Zirkel A, Josipovic N, Oole E, van IWFJ, Houtsmuller AB, Papantonis A *et al.*: **Redundant and specific roles of cohesin STAG subunits in chromatin looping and transcriptional control**. *Genome Res* 2020, **30**:515-527.

81. Kojic A, Cuadrado A, De Koninck M, Gimenez-Llorente D, Rodriguez-Corsino M, Gomez-Lopez G, Le Dily F, Marti-Renom MA, Losada A: **Distinct roles of cohesin-SA1 and cohesin-SA2 in 3D chromosome organization**. *Nat Struct Mol Biol* 2018, **25**:496-504.

82. Wutz G, Ladurner R, St Hilaire BG, Stocsits RR, Nagasaka K, Pignard B, Sanborn A, Tang W, Varnai C, Ivanov MP *et al.*: **ESCO1 and CTCF enable formation of long chromatin loops by protecting cohesin(STAG1) from WAPL**. *eLife* 2020, **9**:e52091.