CTCF and Cohesin in Genome Folding and Transcriptional Gene Regulation

Matthias Merkenschlager[1,*] and Elphège P. Nora[2,*]

[1]MRC Clinical Sciences Centre, Faculty of Medicine, Imperial College London, London W12 0NN, United Kingdom; email: matthias.merkenschlager@csc.mrc.ac.uk

[2]Gladstone Institute of Cardiovascular Disease, San Francisco, California 94158; email: elphege.nora@gladstone.ucsf.edu

Annu. Rev. Genom. Hum. Genet. 2016. 17:17–43

First published online as a Review in Advance on April 18, 2016

The *Annual Review of Genomics and Human Genetics* is online at genom.annualreviews.org

This article's doi: 10.1146/annurev-genom-083115-022339

*Contributed equally.

Keywords

CTCF, cohesin, chromosome conformation capture, genome folding, TAD, loop extrusion, gene regulation

Abstract

Genome function, replication, integrity, and propagation rely on the dynamic structural organization of chromosomes during the cell cycle. Genome folding in interphase provides regulatory segmentation for appropriate transcriptional control, facilitates ordered genome replication, and contributes to genome integrity by limiting illegitimate recombination. Here, we review recent high-resolution chromosome conformation capture and functional studies that have informed models of the spatial and regulatory compartmentalization of mammalian genomes, and discuss mechanistic models for how CTCF and cohesin control the functional architecture of mammalian chromosomes.

LEVELS OF GENOME FOLDING

3C: chromosome conformation capture

TAD: topologically associating domain

Eukaryotic genomes are partitioned into chromosomes that change their organization during the cell cycle. Chromosome conformation capture (3C) experiments have revealed that chromosomes swap their characteristic mitotic shape for an elaborate pattern of folds and loops when they decondense in the early part of G1 phase (19, 24, 101). Genome folding in interphase has several critical functions. First, folding maintains genome integrity by suppressing illegitimate recombination between repetitive sequences (60, 92). Second, units of folding correspond to units of DNA replication in S phase, indicating a role for genome folding in replication (109) (see sidebar DNA Replication Domains). Third, gene regulatory elements and their targets can be separated by large genomic distances that need to be bridged while maintaining the specificity of regulation. Genome folding into spatial domains facilitates transcriptional regulation because promoters and enhancers within the same domain preferentially interact with each other (29, 63, 137). Here, we focus on the role of genome folding in transcriptional regulation.

Topologically Associating Domains and Contact Domains as Primary Units of Genome Folding in Interphase

The primary units of interphase chromosome folding are submegabase-scale domains referred to as topologically associating domains (TADs) (**Figure 1**), which are defined by preferential interactions within them (26, 104). Quantitatively, Hi-C contact frequencies between loci in the same TAD are approximately two to three times those of genomic regions outside of the TAD.

Mammalian TADs often have nested configurations (26) in which larger TADs encompass smaller ones. Because of this hierarchical organization, the identification of TADs is highly dependent on the resolution of the Hi-C data used and the scale at which the data are analyzed (38). Operationally, TADs are defined as self-associating chromosome segments detected by 3C-based methods. The highest-resolution Hi-C maps currently available show that the smallest mammalian TADs, also called contact domains at this scale, range in size from 40 kb to 3 Mb, with a median size of 185 kb (110). High-resolution Hi-C has demonstrated that a sizable fraction of contact domains (39%) are connected at their bases (110), suggesting that many TADs are enclosed by a chromatin loop (with ~10,000 loops detected genome-wide at 1-kb resolution) (**Figure 2**). Contact domains without loops are typically flanked by loop domains and may emerge

DNA REPLICATION DOMAINS

DNA replication in S phase is organized into domains that coincide with topologically associating domains (TADs) (109). Analysis of Hi-C data from three mouse cell lines and four human cell types (26, 62, 110) indicated that replication domains that show differences in replication timing in different cell types (developmental domains) are significantly less compartmentalized than constitutive domains (24).

Whether genomic regions will be replicated early or late is determined long before S phase, in early G1 (24, 109). The point in G1 when replication timing is set (the replication timing decision point) coincides with the completion of 3-D chromatin reorganization. The nuclear lamina–associated protein Rif1 is a known regulator of replication timing in mammalian cells (13), as recently shown by circular chromosome conformation capture (4C) assays demonstrating that Rif1 controls long-range interactions between replication timing domains in mouse embryonic stem cells and fibroblasts (40). Rif1 may therefore control replication timing by limiting interdomain interactions (40).

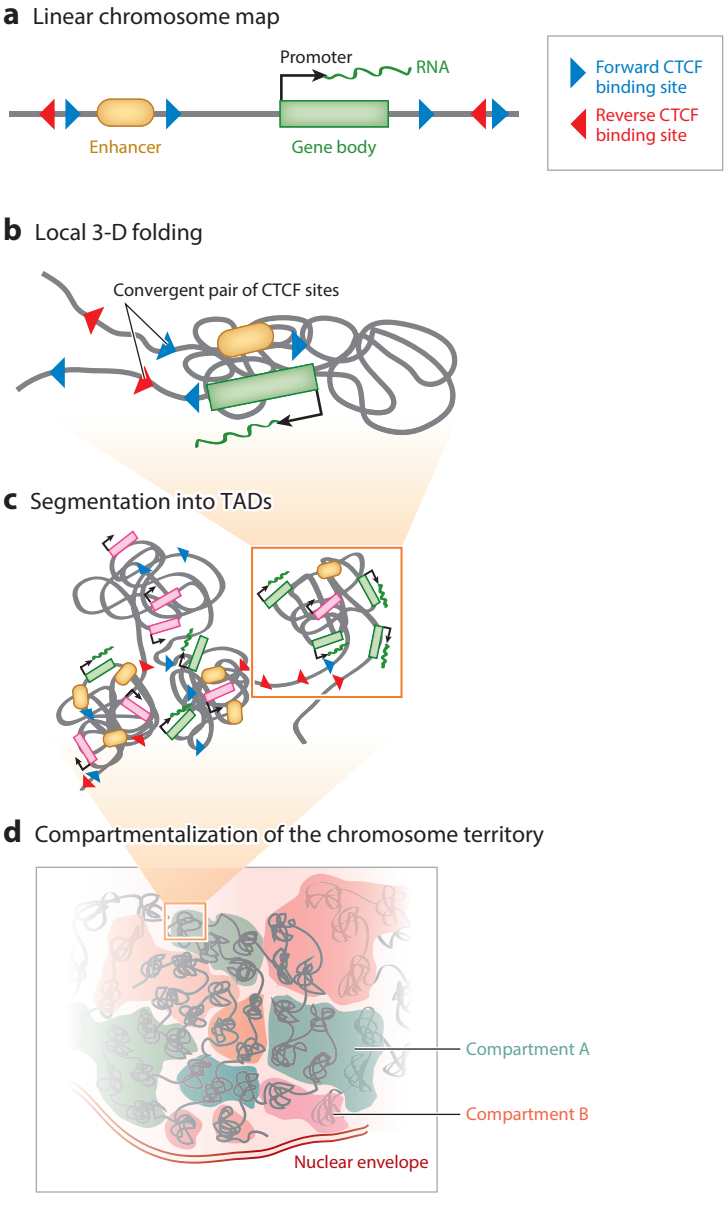


Figure 1

Levels of genome folding. (*a*) A linear map of a theoretical locus with one enhancer, one gene and its promoter, and CTCF binding sites in either a forward (*blue*) or reverse (*red*) orientation. (*b*) Local 3-D folding. This folding brings convergent pairs of CTCF sites into close spatial proximity through a loop, fostering chromosomal contacts between the enhancer and the target promoter in the intervening domain. (*c*) Segmental folding into topologically associating domains (TADs). This folding packages enhancers and promoters from the same domain together while insulating them from the regulatory elements of neighboring domains. TADs also contain inactive genes (*pink*), which are not responsive to surrounding enhancers. (*d*) Compartmentalization of the chromatin territory. The coalescence of TADs from the same or different chromosomes defines two main compartments [A (*blue-green*) and B (*light red*)], which roughly correspond to the transcriptionally active and inactive fractions of the genome. Compartments A and B can be further subdivided into a total of six subcompartments (*different shades of blue-green and red*).

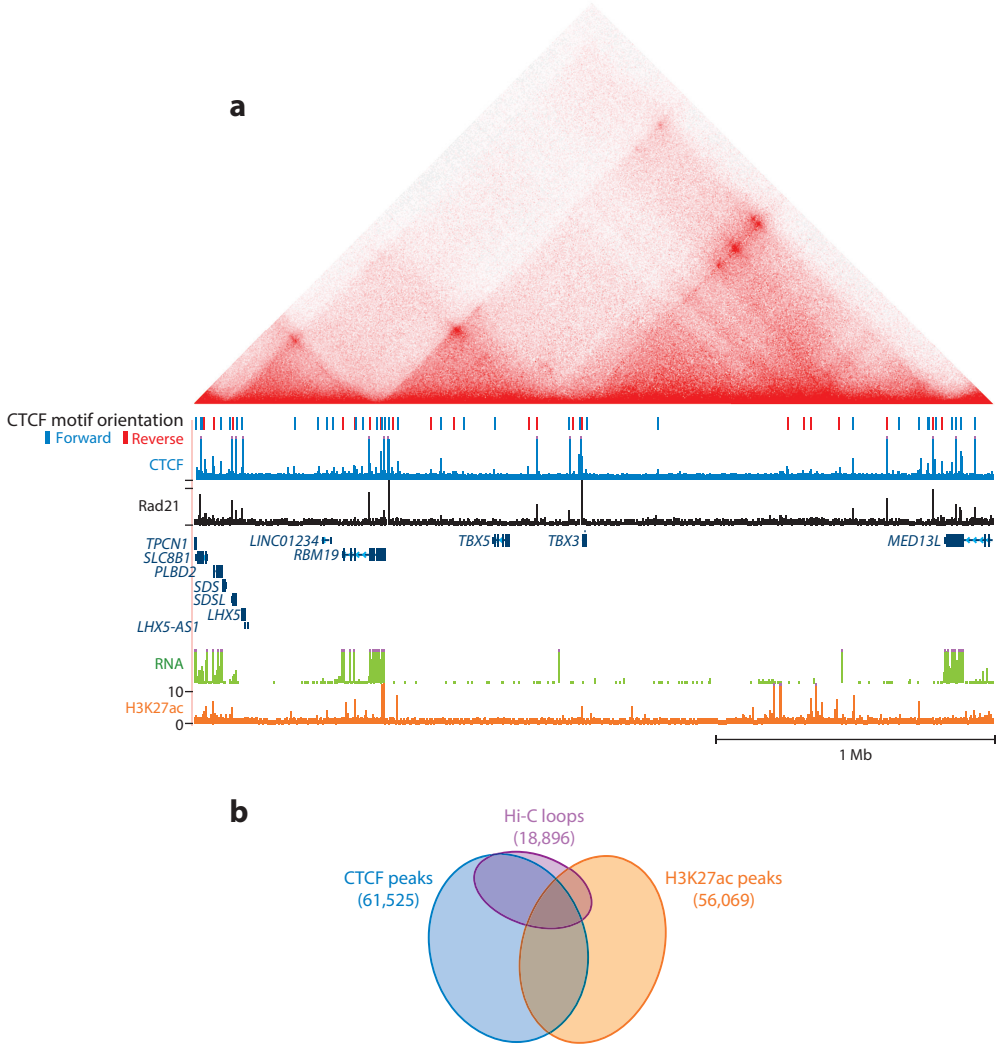


Figure 2

Genomic distribution of loops, CTCF, cohesin, and transcription. (*a*) High-resolution Hi-C data from the GM12878 human lymphoblastoid cell line at the *TBX5* locus. Topologically associating domains (TADs) appear as triangles of high signal on the Hi-C heat map; loops appear as focal points. The image shows CTCF and cohesin (Rad21 subunit) data from chromatin immunoprecipitation combined with DNA sequencing (ChIP-Seq) as well as the orientation of CTCF sites within ChIP-Seq peaks. Note that the CTCF and cohesin peaks colocalize and that CTCF sites engaged in looping show higher ChIP-Seq signals than other sites. RNA sequencing (RNA-Seq) data identify transcribed regions, and H3K27ac (acetylation of histone H3 on lysine 27) ChIP-Seq highlights active enhancers and promoters, most of which do not overlap looping elements. (*b*) Overlap of Hi-C loops, CTCF peaks, and H3K27ac peaks, highlighting that, although most loops contain a CTCF peak, the majority of CTCF peaks and active regulatory elements are not engaged in looping. Data are from Reference 110 and the Encyclopedia of DNA Elements (ENCODE) project, retrieved from the UCSC Genome Browser (**https://genome.ucsc.edu**).

from constraints exerted by neighboring loop domains (110). Although the available single-cell Hi-C data are too sparse to reconstruct TADs, they do suggest that intradomain contacts are more uniform between individual cells than contacts between domains (95).

FISH: fluorescence in situ hybridization

Higher-Order Genome Organization

TADs that share similar chromatin states tend to associate with each other in *cis* (with TADs on the same chromosome) and in *trans* (with TADs on other chromosomes) to form two genomic compartments (20, 44) (**Figure 1**). Compartment A contains most of the transcriptionally active regions, whereas compartment B contains mostly gene-poor and transcriptionally inactive regions (78). These two compartments can be subdivided into smaller subcompartments (110) (**Figure 1**). The driving forces behind compartmentalization are not fully understood, but chromatin composition and transcription may play a role. Mechanistically, the experimental recruitment of the H3K9me3 (trimethylation of histone H3 on lysine 9) methyltransferase SUV39H1 can induce a TAD in compartment A to ectopically engage compartment B (149). Interestingly, the susceptibility to nuclear repositioning varies between loci, suggesting the involvement of additional, as yet uncharacterized determinants (149). Within compartments, specific loci tend to associate in *trans*, as exemplified by clusters of Nanog/Oct4/Sox2-bound sites (17) or Polycomb-bound loci in mouse embryonic stem cells (23). The segregation of active from inactive regions is reminiscent of observations by fluorescence in situ hybridization (FISH) microscopy that chromosomes occupy characteristic nuclear territories that reflect gene density and activity, with preferential localization of euchromatin toward the center and heterochromatin toward the periphery of the nucleus and associated with the nuclear lamina (9, 14).

The familiar banding pattern of *Drosophila* polytene chromosomes reflects the organization of TADs and directly links the visual and molecular organization of chromatin (32). Polytene chromosomes are closely aligned with each other and form TADs but not compartments, suggesting that distinct mechanisms underlie TADs and compartments (32). Finally, higher-order chromatin organization reflects specific cellular functions: Segmented nuclei help granulocytes squeeze through tight spaces, and a cell type–specific central position of heterochromatin optimizes night vision in rodent photoreceptor cells (128). Sperm is an extreme case of chromatin structure following function, and radical reorganization of nucleosome-based chromatin into protamin-based chromatin allows extremely tight packing of the paternal genome. Surprisingly, however, sperm chromatin appears to retain TADs as well as compartments A and B (3).

Chromosomal Organization Is Dynamic

Single-cell Hi-C provides a genome-wide glimpse of cell-to-cell variation in chromosome organization (95). Although the position of TADs along chromosomes appears to be consistent between cells, there are tremendous differences in chromosome folding in *cis* as well as in *trans*. This is in agreement with the heterogeneity of chromosome territories observed by FISH microscopy (47). Because imaging and single-cell Hi-C do not currently have the resolution to robustly measure chromatin folding at the submegabase scale, it remains unclear how stable the loops are that demarcate domains (**Figure 3**). This limits our understanding of the molecular architecture of loops at the base of contact domains. Polymer modeling has been used extensively to help fill this gap, providing simulations of single-folded chromatin fibers as well as average configurations of chromosome ensembles that can be directly compared with population-based assays such as Hi-C (for a review, see 62).

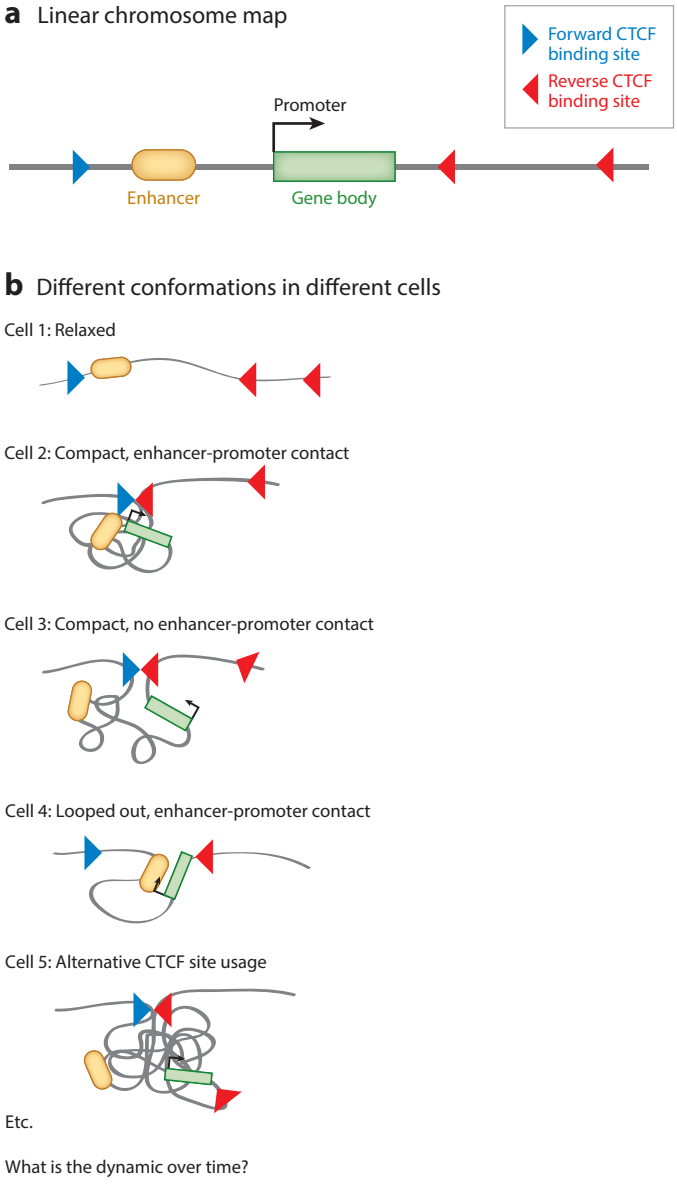


Figure 3

Cell-to-cell variability of chromatin architecture. (*a*) A linear map of a theoretical locus. (*b*) Different conformations in different cells. As shown here, 3-D folding is not identical in all cells; rather, chromatin adopts various conformations, from extended to compact, that can either enable or prevent enhancer-promoter contact. Alternative usage of looping CTCF sites can also explain the nested appearance of TADs observed in data from population-based Hi-C. Hi-C data from populations of cells do not directly capture the cell-to-cell variability in chromatin architecture or its dynamics over time.

Data-driven models use Hi-C data to derive values for biophysical parameters that, when applied to ensembles of chromatin conformation and averaged, give rise to contact frequencies similar to those shown by Hi-C. They are often used to generate spatial models (4). Other types of polymer models start with defined biophysical parameters to create conformation ensembles that are then compared with real Hi-C data (62). Both approaches generate predictions that can be explored by orthogonal approaches such as imaging. Model-derived conformation ensembles can be mined to assess the folding pattern of individual fibers in silico. By addressing how similar single fibers are to the population average, this approach can yield insights into the variability of folding patterns.

Data-driven modeling suggests that chromatin conformation within TADs is highly variable from cell to cell and that contacts between looping elements are fluctuating rather than stable (46) (**Figure 3**). High-resolution imaging indicates that TADs, and loop domains in particular, are not static structures stably connected at their base (34, 46, 104, 141). Rather, the apparent self-association of TADs and loop domains may emerge from the summation of many possible structures across the cell population. In this scenario, loops are probabilistic (42). In agreement with this view, FISH has shown that only a small subset of cells have a complete spatial overlap between loop anchors (46, 104).

The boundaries of TADs and their associated chromosome loops as detected experimentally appear to be relatively constant across cell types, and high-resolution Hi-C has shown that any two cell types share 50–75% of loops (110). Long-range folding of TADs can change during differentiation, often accompanying changes in transcriptional activity and chromatin state (9, 103). Up to 25% of the genome switches compartments during the differentiation of human embryonic stem cells (25).

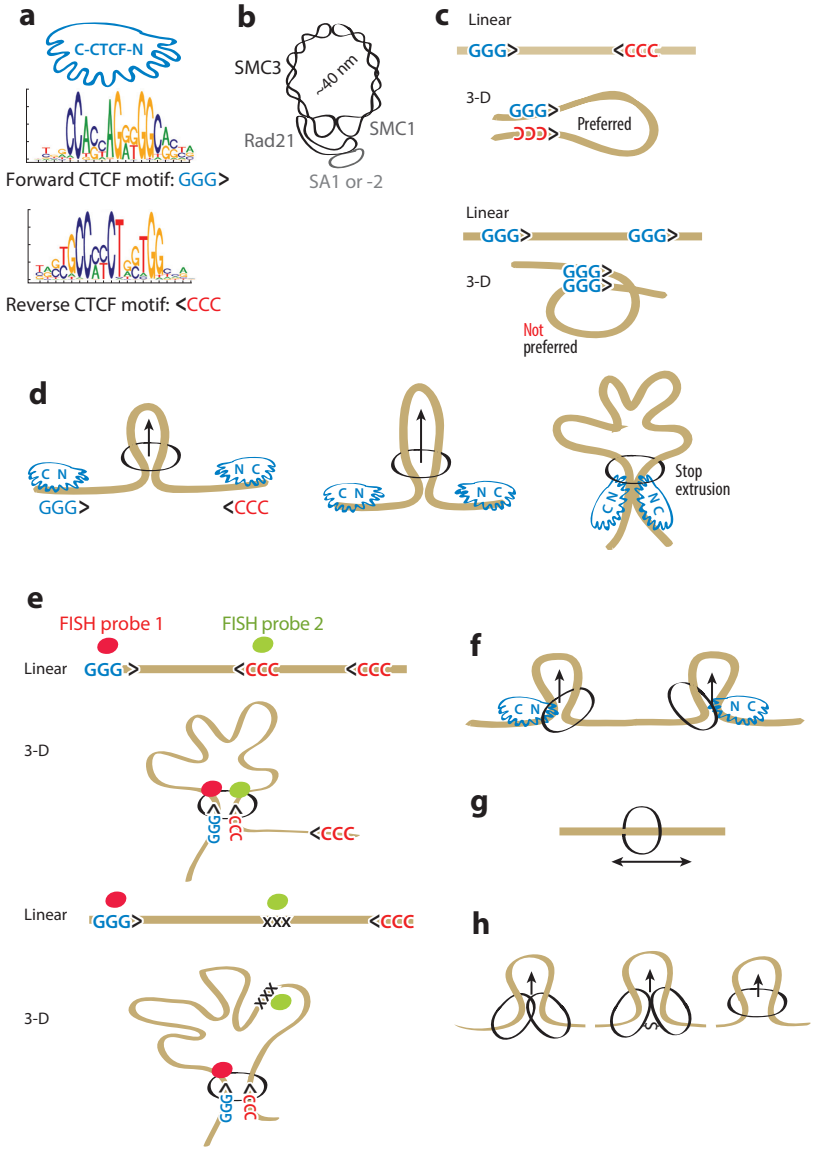
Compartments and TADs are not a constitutive feature of chromosomes. In addition to mitotic chromosomes (24, 101), other examples of loss of compartmentalization are mammalian X chromosomes subject to X inactivation, Hutchinson-Gilford progeria syndrome patient cells that enter senescence (87), and *Drosophila* cells responding to heat shock (76). Patterns in chromatin folding therefore arise from mechanisms that continuously instruct chromosome organization.

CTCF, COHESIN, AND THE MOLECULAR TIES OF LOCAL GENOME FOLDING

The boundaries of compartments, TADs, and loop domains all are enriched for the binding of CTCF, an 11-zinc-finger, sequence-specific DNA-binding protein (26, 110, 143) (**Figures 1**, **2**, and **4*a***). CTCF is conserved in most bilaterian metazoa (57, 69), broadly expressed, and essential for both cellular function (11, 36) and embryonic development (125, 129, 145). CTCF is the only known insulator protein in vertebrates (6), and it can block enhancer function in reporter plasmids (45, 108) and in its native chromatin context (8, 18, 29, 52, 53, 82, 97, 115, 122). In addition to compartment, TAD, and loop domain boundaries, CTCF is enriched at transitions between distinct chromatin states (15) and interactions with the nuclear lamina (142), which also align with TAD boundaries. CTCF enrichment at boundaries points to an important role for CTCF in boundary formation, although many CTCF sites do not appear to engage in looping or chromatin folding (**Figure 2**). At a small subset of sites (~1.5%), CTCF binding is sensitive to the methylation status of DNA (8, 37, 54, 146). In this way, CTCF binding can be epigenetically regulated to control imprinted gene expression (5, 56, 66). Similarly, disrupted DNA methylation (e.g., in cancers with metabolic deregulation) can interfere with CTCF binding (39). CTCF can also interact with RNA, and the underlying biological implications remain to be explored (70, 114).

The Functional Relationship Between CTCF and Cohesin

A functional link between CTCF (a high-affinity DNA-binding protein) and cohesin (a key component of chromatin) was provided by the finding that CTCF and cohesin co-occupy tens of thousands of binding sites across the genome (106, 113, 134, 146) (**Figure 2**). Cohesin is a ring-like multiprotein complex (**Figure 4*b***) that provides cohesion between sister chromatids from the time of DNA replication in S phase until cell division (99). This function of cohesin is essential and enables postreplicative DNA repair and proper chromosome segregation through mitosis and meiosis, ensuring the integrity of genomic information passed on from mother to daughter cells and from one generation of multicellular organisms to the next (99). Cohesin and the highly



related condensin and Smc5/6 complexes are built around heterodimers of structural maintenance of chromosomes (SMC) proteins. In contrast to cohesin, condensin complexes are not enriched at CTCF binding sites (29). In cohesin, a V-shaped Smc1-Smc3 heterodimer is complemented by Rad21/Scc1 and Scc3/SA1/SA2 subunits to form a structure large enough to topologically enclose two chromatin fibers (99) (**Figure 4*b***). In higher eukaryotes, cohesin is a major component of chromatin in cycling, noncycling, and even postmitotic cells.

ChIP: chromatin immunoprecipitation

Although strong cohesin sites usually coincide with the binding of CTCF (106, 113, 134, 146) (**Figure 2**), weaker cohesin sites often map to active promoters and enhancers (35, 65, 118, 154). Here, cohesin apparently colocalizes with its loading factor Nipbl, Mediator components, and tissue-specific transcription factors (35, 65, 118, 154).

Depletion of CTCF reduces the enrichment of cohesin at CTCF binding sites, as measured by chromatin immunoprecipitation (ChIP), but does not disrupt the presence of cohesin on chromatin (106, 146). CTCF therefore affects the distribution of cohesin on chromosomes, rather than its loading or association with chromosomes per se. Depletion of cohesin subunits does not substantially alter the binding pattern of CTCF, suggesting that CTCF is upstream of cohesin in positioning cohesin to its target sites. The SA1 and SA2 cohesin subunits can physically interact with the C terminus of CTCF in vitro (151). Upon SA1 knockout, the SMC1 and SMC3 cohesin subunits redistribute and show reduced association with some CTCF sites, suggesting that SA1 may provide a physical link between CTCF and cohesin (111). The deletion of the putative SA1/2-binding domain in CTCF reduces cohesin binding and transcriptional insulation in reporter transgenes (151), but its impact on cohesin enrichment at endogenous CTCF sites in vivo remains to be tested.

Depletion of CTCF affects looping between CTCF sites (131) and disrupts insulation between neighboring TADs (160). However, compartments are unaffected by the loss of CTCF, even when the protein is almost completely degraded using inducible strategies (E.P. Nora & J. Dekker, unpublished data). Cohesin depletion abrogates looping between CTCF sites and reduces interactions within TADs but does not affect compartmentalization (54, 120, 121, 126, 160). Chromosome folding beyond the megabase scale is disrupted only by depletion of both CTCF and cohesin, which results in general chromatin compaction (139).

Figure 4

How CTCF and cohesin contribute to genome folding. (*a*) CTCF and its motif. In subsequent panels, the forward motif is abbreviated as "GGG>" (shown in *blue*), and the reverse motif is abbreviated as "<CCC" (shown in *red*). The central zinc fingers of CTCF engage the CTCF core motif, and the C-terminal fingers are oriented toward the 5′ end of the motif (96). (*b*) The cohesin complex, showing the core subunits SMC1, SMC3, Rad21, and SA1 or -2. The internal diameter of the ring-like structure is ∼40 nm, large enough to accommodate chromatin fibers. (*c*) Loops form preferentially between convergent CTCF binding sites, suggesting that CTCF-CTCF interactions are established not by 3-D diffusion but rather by a process that tracks the chromatin fiber. (*d*) Extrusion may form chromatin loops. Extrusion stops when two occupied convergent CTCF sites are reached. (*e*) When a boundary element is lost (indicated by the "xxx"), extrusion continues to form a larger topologically associating domain (TAD). Red and green fluorescence in situ hybridization (FISH) probes mark the original boundary elements. (*f*) Simultaneous extrusion of loops from both boundaries generates two loops (see also the related sidebar Where Does Loop Extrusion Initiate?, below). (*g*) A single ring embracing a chromatin fiber moves freely without extruding a loop (see also the related sidebar The Topology of Loop Extrusion, below). (*h*) Loop extrusion could be mediated by two cohesin rings that embrace the chromatin fiber and are linked to each other either topologically (*left*) or by complex formation (*center*). Extrusion could also be mediated by a single cohesin ring sliding over the top of a preformed chromatin loop (*right*) (see also the related sidebar The Topology of Loop Extrusion, below).

How Does CTCF Mediate Chromatin Looping?

ChIA-PET: chromatin interaction analysis by paired-end tag sequencing

Reports that CTCF forms dimers or multimers have spurred the development of models in which chromosomal contacts between CTCF-bound sites are directly mediated by CTCF-DNA and CTCF-CTCF interactions (105). This model has been invoked to account for the association in *trans* of a subset of loci bound by CTCF and for the disruption of this colocalization by CTCF knockdown or deletion of the underlying CTCF binding sites (79, 153, 157–159). In contrast to *trans* interactions between loci bound by pluripotency factors in mouse embryonic stem cells, putative *trans* interactions between CTCF sites are not visible above the background in Hi-C data (78). Chromatin interaction analysis by paired-end tag sequencing (ChIA-PET) experiments have confirmed that most CTCF-mediated interactions are local and involve loci less than 1 Mb apart (55, 138). Importantly, the CTCF-multimerization models fall short of explaining why not all CTCF sites interact with each other and, in particular, why most pairs of interacting CTCF sites detected by Hi-C are in a convergent orientation (110, 143).

Long-Range Interactions Preferentially Occur Between CTCF Sites with Convergent CTCF DNA Motifs

Interestingly, the probability and strength of interactions between CTCF sites are dictated by the orientation of CTCF motifs. In mammalian cells, high-resolution Hi-C data have shown that the vast majority of CTCF-based long-range interactions occur between convergent sites (110) (**Figure 4*c***). This finding suggests that CTCF proteins that are bound to the DNA at the base of loops align with each other in a head-to-head configuration. Pairs of divergent CTCF sites are enriched at TAD boundaries (48).

Recent CTCF ChIA-PET data suggest that approximately one-third of CTCF-CTCF interactions form between CTCF motifs in a tandem orientation (138). This apparent discrepancy may result from how interactions are defined: Hi-C uses a local background (110), whereas ChIA-PET applies a uniform threshold genome-wide. As CTCF binding within TADs inherently produces a higher interaction signal, the ability of tandem CTCF sites to specifically engage in looping remains to be explored quantitatively. Circular chromosome conformation capture (4C) data confirmed that most CTCF-CTCF interactions are between convergent motifs but also found a substantial frequency of interactions between tandem motifs, which the authors attributed to difficulties in assigning interactions to specific CTCF sites in cases where these sites are in clusters of motifs with both orientations (18) (**Figure 2**).

CTCF Motif Orientation Directs Long-Range Interactions

Inverting one site of a convergent looping pair disrupts the loop (18, 53), suggesting that convergent orientation is central for interactivity between remote pairs of CTCF sites. The experimental inversion of clustered CTCF sites in the *Pcdh* protocadherin and β-globin loci disrupted local chromatin folding and allowed the inverted CTCF cluster to access previously insulated regions downstream of the CTCF site (53). The inversion of individual CTCF motifs engaged in looping abrogated loop formation but—in contrast to the inversion of clustered CTCF sites—was insufficient to disrupt local chromatin folding or to redirect long-range interactions (18). Altering clusters of CTCF binding sites may therefore have more profound effects than altering single sites. Because loops form between pairs of convergent CTCF sites, the outcome of altering one site

TRACKING

Tracking could account for the directionality of CTCF-CTCF interactions and for other orientation-dependent phenomena within TADs, such as the orientation-dependent use of recombination signal sequences in legitimate and illegitimate V(D)J recombination (60) and immunoglobulin class switch recombination (27). The site-specific bacterial T3 resolvase provides a cautionary tale, however; numerous lines of evidence suggested that this resolvase performs tracking, but careful theoretical and experimental dissection concluded otherwise (133). Site-specific recombination by the T3 resolvase in vitro requires a pair of *res* sites. Akin to loop formation between CTCF motifs (110), RAG-dependent use of recombination signal sequences (60), and immunoglobulin class switch recombination (27), recombination by the T3 resolvase prefers particular orientations of *res* sites over alternative orientations. This preference has now been explained by a two-step synapsis model in which *res* sites can initially come together regardless of orientation, but subsequent topological steps selectively retain certain *res* site orientations, while others are lost owing to topological strain (133). Pairing of CTCF in a convergent orientation may involve additional layers of molecular choreography that discriminate between appropriate and inappropriate orientations.

depends on the arrangement of surrounding CTCF sites and is therefore highly locus dependent (115).

The Case for Tracking and Extrusion Models of Genome Folding

The finding of a preferred motif orientation suggests that the mechanisms that position DNA-bound CTCF proteins relative to each other are not based on simple 3-D diffusion. Instead, the dependence on CTCF motif orientation invokes images of linear tracking, because tracking can read the directionality of DNA sequences, either directly (by DNA sequence) or indirectly (via the direction of proteins bound to the DNA sequence) (**Figure 4*c***; for a more detailed discussion, see sidebar Tracking).

CTCF motif orientation has revived speculative models that SMC protein complexes may extrude DNA into loops (98). An exciting proposal is that DNA loops and 3-D chromatin domains could emerge from molecular processes that operate in 1-D by linear tracking (1, 41, 102, 115). In this scenario, cohesin acts as a loop-extruding factor by binding DNA, initiating loop formation, and translocating on the chromatin fiber (**Figure 4*d***). Loop extrusion stops either when cohesin dissociates from the chromatin template or when it encounters a CTCF boundary element. Once halted at CTCF binding sites, the loop extrusion complex joins the boundaries of a chromatin domain by a loop (**Figure 4*d***). To account for the convergent orientation of CTCF sites at loop anchors, the extrusion model posits that CTCF can halt cohesin tracking only when encountered from the 3′ end of its consensus motif.

The model can potentially explain interesting features of chromatin folding, notably why domains fuse when a boundary is lost (extrusion would simply continue up to the next boundary; see **Figure 4*e***) and why interaction frequencies between specific loci are not intrinsic to these loci, but rather are dependent on the position of these loci relative to boundaries (41, 115). Loop extrusion also predicts that loss of the extrusion complex will alter interactions within domains and perhaps abrogate domain formation altogether, whereas loss of boundary elements will increase interactions between domains. Hi-C data from partially cohesin-depleted cells are broadly consistent with altered interactions within TADs (63, 120, 126, 160), whereas partial depletion of CTCF

weakens TAD boundaries (160; E.P. Nora, unpublished data). Deletion and inversion of CTCF sites at domain boundaries resulted in the loss of existing interactions and the formation of new interactions, with good agreement between experimental Hi-C data and modeling predictions (115). A loop extrusion mechanism can also account for the nested appearance of many TADs in Hi-C data from populations of cells (**Figure 2**). If a given CTCF site is not occupied all the time, loop extrusion factors can proceed across the site until they encounter the next CTCF site in a convergent orientation. These permeable boundaries would correspond to CTCF binding sites of lower occupancy [as shown by chromatin immunoprecipitation combined with DNA sequencing (ChIP-Seq)] that are not continually occupied across the cell population (**Figure 3**).

Despite their explanatory potential, current loop extrusion models remain highly speculative. In particular, it is unclear whether extrusion would initiate in the interior or at the boundaries of a domain (see sidebar Where Does Loop Extrusion Initiate?), what would drive the processivity of the loop extrusion complex (see sidebar What Drives Loop Extrusion?), what topological requirements would have to be met for cohesin to function as a loop extrusion complex (see sidebar The Topology of Loop Extrusion), and how the extrusion complex would bypass CTCF sites in one orientation but not the other (see sidebar How Is CTCF Motif Orientation Read Out?).

WHERE DOES LOOP EXTRUSION INITIATE?

An unresolved mechanistic question is whether loop extrusion initiates within the prospective loop domain or at the domain boundaries. If extrusion initiates at a single site within the prospective loop domain and progresses toward the domain boundaries, then it forms a single loop (**Figure 4*d***). Initiation at more than one site may form a series of nested domains or subdomains (26, 110). Models in which extrusion starts within the prospective loop domain explain how cohesin colocalizes with CTCF even though CTCF does not itself load cohesin onto chromosomes (**Figure 4*d***).

Alternative models envisage that loop extrusion initiates from CTCF-bound cohesin at the prospective domain boundaries (102, 152). This requires colocalization of cohesin with CTCF prior to loop extrusion. Simultaneous extrusion of loops from both boundaries would generate two loops that would then collide and be resolved to form a domain (**Figure 4*f***).

WHAT DRIVES LOOP EXTRUSION?

In prokaryotes, DNA extrusion drives genome partitioning between daughter cells (50, 150). DNA extruders, such as the FtsK and SpoIIIE proteins, form channels or pores through which they actively pump DNA (50). It is unclear how the force for cohesin-mediated loop extrusion would be generated. One suggestion is that ATP hydrolysis by the SMC ATPase domains may be channeled into mechanical force. ATP hydrolysis controls the formation and resolution of contacts within SMC complexes (99), which may drive conformational changes (94, 127) and could trigger movement of SMC complexes along the DNA (74). Alternatively, external processes such as transcription could drive extrusion. In *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*, cohesin is cleared from gene bodies by transcription and accumulates between convergent genes (51, 73). Although this provides precedent for a link between transcription and cohesin movement, it is unclear whether transcription actually translocates SMC complexes along the chromatin fiber in living cells.

THE TOPOLOGY OF LOOP EXTRUSION

Whether cohesin acts as a single ring or as a pair of rings held together has been a subject of lively debate (99, 124, 156). Notably, a single topologically bound ring would not extrude a loop; instead, it would freely translocate along the chromatin fiber (**Figure 4*g***). Tethering is therefore necessary to enable loop extrusion by a topologically bound ring. In models in which loop extrusion initiates within the prospective loop domain, tethering could be provided by the coupling of two topologically bound rings. Cohesin rings could link to each other topologically (**Figure 4*h***, left side) or by complex formation (**Figure 4*h***, center). A single cohesin ring could initiate loop extrusion by sliding over a preformed loop (**Figure 4*h***, right side). This would avoid the question of whether cohesin rings dimerize but would be inconsistent with topological embrace models of the cohesin interaction with DNA. In a hybrid model, CTCF-associated cohesin initially forms a topological embrace but then lets go and slides over the extruding loop (102, 152). Although not inconceivable, such a topological transition would need to be explained.

HOW IS CTCF MOTIF ORIENTATION READ OUT?

To explain convergent orientation of CTCF sites at loop anchors, CTCF motif orientation could be read as DNA sequence, by protein orientation, or as secondary structure. Sequence-based mechanisms such as transcription or replication operate at 0.5–3 kb/min, whereas extrusion of 1-Mb TADs or 200-kb contact domains would require considerably higher speeds to complete within ~10 min, the lifetime of chromatin-bound cohesin in G1 cells (43). Sequence-based mechanisms would allow loop formation between unoccupied CTCF sites. ChIP-Seq data map cohesin to the loop interior (35, 115, 138), but to read the sequence, the extrusion complex would have to traverse the CTCF motif, positioning it toward the exterior of the loop. An asymmetric interaction between CTCF and cohesin protein is a simpler scenario. However, the 3′ end of the CTCF motif is occupied by the N-terminal CTCF zinc fingers (96), and cohesin appears to bind the CTCF C terminus (151). Finally, CTCF may form local secondary DNA structures (83), which have been invoked to account for the directionality of loop formation between CTCF binding sites (102).

GENOME FOLDING AND REGULATORY SEGMENTATION: CTCF AND COHESIN LINK WITH GENE REGULATION THROUGH GENOME FOLDING

Why Regulatory Segmentation?

Metazoan cell types contain more than 10,000 active promoters and distal gene regulatory elements, which are not hardwired to each other. Given the opportunity, enhancers can work with many target genes, as illustrated by recurrent translocations that bring oncogenes under the control of strong enhancer elements and contribute to malignancies (112, 140). Experimentally enforced interactions between the β-globin promoter and enhancers over 40 kb away can trigger transcription (21), providing a causal link between spatial interactions and gene expression. The challenge for gene regulation is therefore to coordinate interactions between promoters and appropriate distal regulatory elements.

Spatial Segmentation Gives Rise to Regulatory Segmentation

As detailed above, spatial chromatin domains are defined by increased interaction probabilities within domains and reduced interaction probabilities between domains. Applied to gene regulatory elements, spatial segmentation effortlessly translates into functional regulatory segmentation of the genome.

Diversity Within a Shared Framework

Given the diversity of nuclear morphology and the extensive differences in genome usage between different cell types, it was surprising when genome-wide studies suggested that the organization of interphase chromatin is largely conserved between cell types (e.g., 78, 107, 117). Just as the majority of CTCF binding events are invariant between cell types, so too are the positions of TAD boundaries. Despite the conservation of boundaries, however, chromatin states, gene expression, and interactions within TADs can vary significantly between cell types (110). Because the aggregation of TADs into compartments reflects chromatin state and gene expression, the composition of compartments can change accordingly (20, 44). Although rapid or transient changes in transcriptional activity do not trigger immediate reorganization of compartments (64, 72), stable differences in transcriptional state can drive transitions between compartments in which differentially expressed loci switch between compartments A and B (25).

Contact Domains in 3-D Correspond (Loosely) to Functional Regulatory Domains

Regulatory segmentation has been directly observed in functional experiments (137). In this study, regulatory sensors (enhancer-less reporter genes packaged as transposons) introduced into the mouse genome mapped the functional range of endogenous enhancers. Enhancers interacted with reporter insertions in multiple locations, demonstrating that enhancer activity can be broadly distributed rather than gene centric. Importantly, the enhancer activity was restricted to genomic regions that broadly corresponded to TADs (137). Processes constrained by TADs include interactions between gene regulatory elements (19, 61, 89, 119, 137), association with the nuclear lamina (26, 104), the spreading of histone modifications (103), noncoding transcription (2), DNA replication (24, 109), recombination between repeat sequences in yeast (92), and recombination-activating gene (RAG)–mediated recombination in vertebrates (60) (see sidebar DNA Recombination and TAD Boundaries).

Contact Domains as Scaffolds for Transcriptional Regulation

5C: chromosome conformation capture carbon copy

Studies using 3C, chromosome conformation capture carbon copy (5C), capture Hi-C, and high-resolution Hi-C all show specific (or private) enhancer-promoter interactions (61, 64, 89, 110, 117). In GM12878 cells, for example, 30% of observed loops (2,854 out of 9,448) linked annotated promoters with known enhancers, compared with 7% expected by chance (110). For comparison, the fraction of loops involving domain boundaries was ~38% (110). Most promoter-genome interactions detected by capture Hi-C occur within individual TADs. These data are consistent with the concept of TADs as discrete regulatory domains but also show that a significant fraction (up to 12%) of promoters, especially active promoters, interact with sites outside their own TADs (119).

DNA RECOMBINATION AND TAD BOUNDARIES

The RAG recombinase initiates V(D)J recombination in developing lymphocytes by generating DNA double-strand breaks. Within immunoglobulin and T cell receptor loci, these breaks are directed on-target by matched pairs of recombination signal sequences. The choice of recombination signal sequences is determined by their accessibility within the chromatin landscape (132) and histone modifications (84), which reflect cell lineage and developmental stage. In addition to these on-target breaks, RAG can initiate off-target breaks, often at DNA sequences that resemble recombination signal sequences. Off-target breaks can result in chromosomal rearrangements and neoplastic transformation (112, 140).

Interestingly, RAG recombinase activity has recently been linked to TADs. In a model system in which RAG was experimentally directed to chromosomal domains, off-target breaks were confined by TAD boundaries, and deletion of convergent CTCF binding sites extended the range of RAG activity (60). Mechanistically, confinement of RAG recombinase to loop domains requires the DNA repair protein ataxia telangiectasia mutated (ATM), pointing to the involvement of cofactors that restrict RAG activity to TADs. In contrast to RAG recombinase activity, the activity of Cas9 was not confined to TADs, demonstrating that domain boundaries are not physical or mechanical; rather, they are actively read by some chromatin proteins and ignored by others.

Enhancer Sharing and Coregulation of Genes in the Same Interaction Domain

Enhancer-promoter interactions are rarely exclusive (one to one), and instead tend to be promiscuous. Studies using 5C have indicated that an average of approximately four enhancers contact an active gene (109). Similarly, RNA polymerase II ChIA-PET (111–113), Hi-C (105), and capture Hi-C all show that most promoters contact more than one enhancer (108). Extended super-enhancers in murine embryonic stem cells interacted not only with a subset of genes (142 out of 210 genes) they had been tentatively assigned to (148), but also with 361 more distal genes (119).

Genes within the same TAD showed more similar regulation than genes in different TADs during embryonic stem cell differentiation (104). Similarly, in response to progesterone treatment of T47D breast cancer cells, a subset of TADs behaved as discrete regulatory units within which chromatin states and gene expression were regulated in a coordinated fashion (104). The TAD-wide correlation for coregulated gene expression was significant but nevertheless modest ($R < 0.5$), which may change based on the more precise definition of TADs afforded by high-resolution Hi-C data.

CTCF-Based Interactions as a Scaffold for Cell Type–Specific Factors

As recently reviewed in depth (16), enhancer-promoter contacts can be preformed (predate gene expression) or inducible (reflect gene expression), or can include both preformed and inducible components. The *T cell receptor alpha* (*Tcra*) locus provides an example of a CTCF-based local loop that promotes interactions between T cell–specific promoter and enhancer elements (121). The molecular separation between architectural players such as CTCF and the transcription machinery itself may allow for additional layers of gene expression control.

Domain-Wide Interactions Between Gene Regulatory Elements

A surprising result from large-scale 5C analysis (117), capture Hi-C (119), and high-resolution Hi-C (64, 110) was that—despite the evidence for private enhancer-promoter interactions discussed above—most enhancers and promoters do not appear to engage in specific loops. Most domain-wide contacts that promoters make are in fact not with other regulatory regions, such as enhancers, promoters, or CTCF sites. Similarly, most enhancers do not specifically contact promoters, but rather participate in domain-wide contacts. This suggests that many gene regulatory elements may not preferentially engage with each other and instead scan their local interaction domain, in line with the idea that the functional activity of enhancers is constrained by genome folding. In support of this notion, disruption of a TAD boundary causes ectopic chromosomal contacts and long-range transcriptional deregulation (104).

Such domain-wide—or public—interactions may account for observations that many (reported as ~50% in Reference 119 and as ~46% in Reference 64) or most (reported as >50% in Reference 110) promoters do not detectably engage in specific contacts with enhancers. Some of these are promoters of housekeeping genes, which may not always rely on distal regulatory elements. Other enhancer interactions may not have been detected because these promoters do not interact with enhancers more often than they do with other constituent elements of their contact domain (110). Domain-wide or public promoter-enhancer contacts are likely transient and invisible above the local interaction background. Transient promoter-enhancer interactions may support the expression of genes with a burst-like pattern of transcription (135, 136) and in this way form a probabilistic mechanism for the regulation of gene expression.

X Chromosome Inactivation as a Paradigm for the Connection Between Chromosome Structure and Transcriptional Activity

The inactivation of one of the two X chromosomes in female cells highlights that the relationship between genome folding and genome function is reciprocal and context dependent. X chromosome inactivation is initiated following the upregulation of *X inactive specific transcript* (*Xist*) from one of the two X chromosomes in female cells. *Xist* physically coats the future inactive X chromosome (Xi) by exploiting the 3-D configuration of the chromosome. Here, chromosome structure instructs transcriptional dynamics: TADs from compartment A that are positioned close to the *Xist* locus are readily exposed to *Xist* RNA and silenced first (33). As silencing progresses, chromosome structure is reconfigured: TADs are effaced, and the entire Xi switches to the same compartment (104, 131). Loci on the Xi lack preferential long-range partners, and the entire Xi has an overall random organization, with the exception of a handful of peculiar regions that exhibit Xi-specific CTCF binding and participate in setting up multimegabase loops that generate two superdomains (22, 90, 110).

These observations could suggest that the loss of loops, TADs, and compartments is a consequence of transcriptional silencing, and that structure simply follows activity. However, this is not the case: Deletion of *Xist* after X inactivation is established restores Xi loops, TADs, and compartments but not transcriptional activity (90, 130). This highlights that *Xist* RNA directly or indirectly disrupts chromosome folding through mechanisms that can be decoupled from transcriptional repression.

How does *Xist* RNA wipe out TADs and compartments? One hypothesis is that *Xist* interferes with the control of chromosome architecture by CTCF and cohesin. Although the binding pattern of CTCF is similar between the active X chromosome and the Xi (12, 155), cohesin enrichment at CTCF sites is reduced on the Xi (90). Deletion of *Xist* restores cohesin enrichment at CTCF

sites, which suggests that *Xist* RNA may disconnect cohesin from CTCF on the Xi. An intriguing hypothesis is that *Xist* interferes with the CTCF-cohesin interaction through its ability to interact physically with these proteins (90). Importantly, additional key players, such as the SMC domain–containing SMCHD1 protein, are recruited to the Xi and contribute to its unusual structure and transcriptional silencing (10).

WHEN SPATIAL SEGMENTATION BREAKS DOWN

CTCF Binding Site Mutations and Disruption of Boundary Elements Compromise Regulatory Segmentation

In addition to the CTCF site deletion and inversion experiments discussed above (18, 53, 115), ablation of individual CTCF binding sites identified by cohesin ChIA-PET disrupted the expression of selected nearby genes (30), and the deletion of CTCF sites altered spatial interactions and transcriptional regulation in the mouse *HoxA* gene cluster (97). CTCF binding site mutations can disrupt the binding of CTCF and cohesin and are common in colorectal cancer (67), acute T cell leukemia (59), and other types of cancer. When boundary elements are lost, either experimentally or as a result of natural mutations, new interdomain contacts can be formed, and genes located in the affected domains can become deregulated. Human limb malformations are caused by deletions, inversions, or duplications that alter the structure of the TAD-spanning *WNT6/IHH/EPHA4/PAX3* locus, and analogous mutations have been engineered in mice (82). Loss of boundary elements that included CTCF sites and other sequence motifs was linked to ectopic interactions between gene regulatory elements and deregulated gene expression (82). Similarly, inversions on chromosome 3 in acute myeloid leukemia disrupt TAD boundaries. The resulting TAD fusions allow a distal GATA2 enhancer to ectopically activate the *EVI1* oncogene and at the same time reduce endogenous GATA2 expression. A causal link was established by genetic excision of the ectopic enhancer, which restored *EVI1* silencing and led to growth inhibition and differentiation of acute myeloid leukemia cells (49).

Aberrant DNA Methylation Can Interfere with Regulatory Segmentation

CTCF binding is sensitive to DNA methylation at a small subset (1.5%) of genomic sites (37, 85, 146). In this way, CTCF binding can be epigenetically regulated to control imprinted gene expression (5, 56, 66). *IDH* mutations are common in tumors and can result in the production of the onco-metabolite 2-hydroxyglutarate, which interferes with iron-dependent hydroxylases, including the TET family of 5′-methylcytosine hydroxylases. TET enzymes catalyze the active removal of DNA methylation, and *IDH* mutant gliomas display increased methylation of CpG islands and, crucially, a subset of CTCF binding sites. A recent study described both loss and gain of CTCF binding in *IDH* mutant gliomas (39). Interestingly, altered CTCF binding was linked to an increase in the correlation of gene expression across TAD boundaries in *IDH* mutant gliomas. In one critical example, loss of CTCF binding at a domain boundary allowed a constitutive enhancer to interact aberrantly with the receptor tyrosine kinase gene *PDGFRA*, a prominent glioma oncogene (39).

CTCF Mutations

Mice with heterozygous *Ctcf* deletion have been employed in an attempt to rationalize the frequent occurrence of human *CTCF* mutations in cancer (68). Interestingly, *Ctcf* hemizygous mice showed increased variability in DNA methylation and were susceptible to a range of cancers (68).

Germline and Somatically Acquired Cohesin Mutations

Germline mutations in genes encoding cohesin subunits or the cohesin-loading protein NIPBL are rare (1/10,000–1/20,000) and cause Cornelia de Lange syndrome. Somatic cohesin mutations are much more common. Recurrent mutations of cohesin subunits are present in many types of cancer. Initially, somatic cohesin mutations were suspected to lead to genomic instability, but it is now recognized that cohesin mutations often occur early in the development of cancer, before genomic instability (81). Cohesin mutations have been identified in preleukemic hematopoietic stem cells and during the early development of acute myeloid leukemia and other myeloid malignancies. Reduced cohesin expression has been modeled by heterozygosity in SMC3 (144) and inducible RNA interference–mediated cohesin knockdown (93) in mice, and results in altered hematopoietic stem cell homeostasis with increased proliferation of myeloid progenitor cells and increased susceptibility to myeloproliferative disease. Enforced expression of mutant cohesin proteins can block the differentiation of human hematopoietic stem and progenitor cells in vitro and in mouse xenograft models (86). This finding has been attributed to increased chromatin accessibility, promoting the binding of the transcription factors GATA2 and RUNX1. Knockdown of these transcription factors rescued the differentiation block, suggesting that mutant cohesin proteins impair hematopoietic progenitor cell differentiation by controlling chromatin accessibility and transcription factor activity (86).

The Impact of CTCF and Cohesin Deletion

Both CTCF and cohesin are essential factors in cycling cells, and experimental deletion can give misleading results in which phenotypes related to stress and DNA damage can obscure direct effects of CTCF and cohesin on gene expression or genome folding. Controlled short-term deletion of cohesin and deletion of cohesin from noncycling cells (63, 71, 120, 121, 126) combined with Hi-C and gene expression profiling have been useful in disentangling direct from indirect effects of cohesin deletion. Complex loci (88, 121) and enhancer-proximal genes (63) were particularly susceptible to the loss of cohesin.

CTCF and Cohesin Can Separate or Connect Genomic Elements

The punctuation of chromatin into domains by CTCF and cohesin is often represented by schematic drawings in which the image of neatly laid out loops raises the question of why placement in different chromatin loops would prevent spatial and functional interactions between genomic loci. Parallel loops could provide ample opportunities for side-by-side interactions, but insulation may occur if loops crumple, contract, or collapse into self-interacting domains with a sequestered interior. Chromatin loops may shorten the spatial distance between genomic features or align sequences for preferential interactions (31, 63). In this sense, the punctuation of the genome by CTCF- and cohesin-based interactions may serve both to bring genomic sites together and to keep them isolated from each other. Such interactions are illustrated by complex genomic loci, including lymphocyte receptor loci, clustered protocadherin loci, the β-globin locus, and many other developmentally regulated genes where one or more transcription units are subject to multiple regulatory inputs (88).

At clustered protocadherin loci, for example, CTCF sites are located upstream of the start site of each transcription unit. The motif orientation of these CTCF sites is in the sense of transcription and points to a cluster of CTCF sites with reverse orientation at the protocadherin enhancer.

Inversion of this cluster disrupts interactions across the locus and interferes with protocadherin promoter choice and expression (53).

A similar arrangement of CTCF sites is found at the *immunoglobulin heavy chain* (*IgH*) locus, where multiple transcription units are associated with CTCF sites oriented toward a cluster of CTCF sites with inverse orientation at the *IgH* constant region. This arrangement is thought to anchor long-range interactions with multiple *IgH* gene segments, equalizing their opportunity to contribute to the somatic recombination process (7). Major promoter and enhancer elements at the 3′ end of the *Tcra* locus are marked by prominent CTCF sites. Analysis of high-resolution Hi-C data from CH12 pre-B cells indicated that the 3′ part of *Tcra* forms a local loop braced by CTCF (110). 3C data show that this interaction is at least partly cohesin dependent and is further strengthened when *Tcra* becomes transcriptionally active in developing thymocytes (121). *Tcra* transcription and the associated histone modifications recruit the RAG recombinases and thereby promote somatic recombination. Long-range interactions between CTCF-decorated enhancer and promoter elements are reduced in cohesin-deficient thymocytes, which impairs *Tcra* transcription and rearrangement (121). Despite this, the 3′ end of *Tcra* is not insular, but rather interacts with variable gene segments throughout the locus to promote the rearrangement of functional T cell receptors. These and other examples suggest that sequences marked by appropriately oriented CTCF sites will not necessarily be spatially and functionally insulated from each other.

CTCF and Cohesin Functions Beyond Spatial Partitioning

Early evidence that CTCF (6, 147) and cohesin (106, 147) can prevent productive interactions between promoters and enhancers comes from reporter plasmids that are presumably too small to fold into spatially isolated domains. This phenomenon is hard to explain using the framework of interaction domains and may point to the mechanisms that actively interfere with enhancer-promoter interactions. In addition to the role of CTCF and cohesin in genome folding and regulatory segmentation, the literature contains numerous examples of these proteins affecting gene expression by modulating chromatin accessibility (80), transcription factor binding (35, 154), transcriptional elongation (28), alternative splicing (123), interactions with noncoding RNAs (77), and a range of other functions (reviewed in 28, 88).

FUTURE ISSUES

1. What is the structure of CTCF and cohesin bound to DNA?
2. To what extent are CTCF and cohesin required for chromatin architecture, and what levels of genome folding do they control?
3. What are the molecular connections between CTCF and cohesin in vivo, and how does CTCF binding lead to cohesin enrichment at CTCF sites?
4. What are the topological details of cohesin-mediated chromosomal interactions, and what are their dynamics over time?
5. How does CTCF motif orientation determine loop formation, is it the sole determinant, and are there additional layers of regulation?
6. Does CTCF-based, cohesin-mediated loop extrusion account for local genome folding?

7. Where is cohesin loaded, where does loop formation initiate, and how stable are loops over time?
8. What is the force driving extrusion?
9. Is genome folding set once after mitosis and then lasts for the remaining interphase, or is it continuously shaped during interphase?
10. What is the link between the dynamics of chromosome folding and transcription?
11. What is the functional impact of polymorphisms in the human genome that affect CTCF occupancy?
12. To what extent has the evolution of chromosomes been constrained by how CTCF and cohesin control their regulatory architecture?
13. Could the disruption of specific CTCF binding sites have allowed the sudden emergence of mammalian traits by rewiring enhancer-promoter communication?

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

We thank Drs. Ya Guo, Kim Nasmyth, David Sherratt, and Benoit Bruneau for stimulating discussions. We sincerely apologize to our colleagues if we could not discuss their seminal work owing to space limitations or thematic constraints. This work was supported by the Wellcome Trust (WT099276/Z/12/Z) and the Medical Research Council (M.M.) and by a long-term fellowship (LT000302) from the Human Frontier Science Program (E.P.N.).

LITERATURE CITED

1. Alipour E, Marko JF. 2012. Self-organization of domain structures by DNA-loop-extruding enzymes. *Nucleic Acids Res.* 40:11202–12
2. Austenaa LM, Barozzi I, Simonatto M, Masella S, Della Chiara G, et al. 2015. Transcription of mammalian *cis*-regulatory elements is restrained by actively enforced early termination. *Mol. Cell* 60:460–74
3. Battulin N, Fishman VS, Mazur AM, Pomaznoy M, Khabarova AA, et al. 2015. Comparison of the three-dimensional organization of sperm and fibroblast genomes using the Hi-C approach. *Genome Biol.* 16:77
4. Baù D, Sanyal A, Lajoie BR, Capriotti E, Byron M, et al. 2011. The three-dimensional folding of the α-globin gene domain reveals formation of chromatin globules. *Nat. Struct. Mol. Biol.* 18:107–14
5. Bell AC, Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405:482–85
6. Bell AC, West AG, Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98:387–96
7. Benner C, Isoda T, Murre C. 2015. New roles for DNA cytosine modification, eRNA, anchors, and superanchors in developing B cell progenitors. *PNAS* 112:12776–81
8. Beygo J, Citro V, Sparago A, De Crescenzo A, Cerrato F, et al. 2013. The molecular function and clinical phenotype of partial deletions of the *IGF2/H19* imprinting control region depends on the spatial arrangement of the remaining CTCF-binding sites. *Hum. Mol. Genet.* 22:544–57

9. Bickmore WA, van Steensel B. 2013. Genome architecture: domain organization of interphase chromosomes. *Cell* 152:1270–84
10. Blewitt ME, Gendrel AV, Pang Z, Sparrow DB, Whitelaw N, et al. 2008. SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nat. Genet.* 40:663–69
11. Burcin M, Arnold R, Lutz M, Kaiser B, Runge D, et al. 1997. Negative protein 1, which is required for function of the chicken lysozyme gene silencer in conjunction with hormone receptors, is identical to the multivalent zinc finger repressor CTCF. *Mol. Cell. Biol.* 17:1281–88
12. Calabrese JM, Sun W, Song L, Mugford JW, Williams L, et al. 2012. Site-specific silencing of regulatory elements as a mechanism of X inactivation. *Cell* 151:951–63
13. Cornacchia D, Dileep V, Quivy JP, Foti R, Tili F, et al. 2012. Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J.* 31:3678–90
14. Cremer T, Cremer M, Dietzel S, Muller S, Solovei I, Fakan S. 2006. Chromosome territories—a functional nuclear landscape. *Curr. Opin. Cell Biol.* 18:307–16
15. Cuddapah S, Jothi R, Schones DE, Roh TY, Cui K, Zhao K. 2009. Global analysis of the insulator binding protein CTCF in chromatin barrier regions reveals demarcation of active and repressive domains. *Genome Res.* 19:24–32
16. de Laat W, Duboule D. 2013. Topology of mammalian developmental enhancers and their regulatory landscapes. *Nature* 502:499–506
17. de Wit E, Bouwman BA, Zhu Y, Klous P, Splinter E, et al. 2013. The pluripotent genome in three dimensions is shaped around pluripotency factors. *Nature* 501:227–31
18. de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, et al. 2015. CTCF binding polarity determines chromatin looping. *Mol. Cell* 60:676–84
19. Dekker J. 2014. Two ways to fold the genome during the cell cycle: insights obtained with chromosome conformation capture. *Epigenetics Chromatin* 7:25
20. Dekker J, Heard E. 2015. Structural and functional diversity of Topologically Associating Domains. *FEBS Lett.* 589:2877–84
21. Deng W, Lee J, Wang H, Miller J, Reik A, et al. 2012. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell* 149:1233–44
22. Deng X, Ma W, Ramani V, Hill A, Yang F, et al. 2015. Bipartite structure of the inactive mouse X chromosome. *Genome Biol.* 16:152
23. Denholtz M, Bonora G, Chronis C, Splinter E, de Laat W, et al. 2013. Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell* 13:602–16
24. Dileep V, Ay F, Sima J, Vera DL, Noble WS, Gilbert DM. 2015. Topologically associating domains and their long-range contacts are established during early G1 coincident with the establishment of the replication-timing program. *Genome Res.* 25:1104–13
25. Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* 518:331–36
26. Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, et al. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* 485:376–80
27. Dong J, Panchakshari RA, Zhang T, Zhang Y, Hu J, et al. 2015. Orientation-specific joining of AID-initiated DNA breaks promotes antibody class switching. *Nature* 525:134–39
28. Dorsett D, Merkenschlager M. 2013. Cohesin at active genes: a unifying theme for cohesin and gene expression from model organisms to humans. *Curr. Opin. Cell Biol.* 25:327–33
29. Dowen JM, Bilodeau S, Orlando DA, Hübner MR, Abraham BJ, et al. 2013. Multiple structural maintenance of chromosome complexes at transcriptional regulatory elements. *Stem Cell Rep.* 1:371–78
30. Dowen JM, Fan ZP, Hnisz D, Ren G, Abraham BJ, et al. 2014. Control of cell identity genes occurs in insulated neighborhoods in mammalian chromosomes. *Cell* 159:374–87
31. Doyle B, Fudenberg G, Imakaev M, Mirny LA. 2014. Chromatin loops as allosteric modulators of enhancer-promoter interactions. *PLOS Comput. Biol.* 10:e1003867
32. Eagen KP, Hartl TA, Kornberg RD. 2015. Stable chromosome condensation revealed by chromosome conformation capture. *Cell* 163:934–46

33. Engreitz JM, Pandya-Jones A, McDonel P, Shishkin A, Sirokman K, et al. 2013. The Xist lncRNA exploits three-dimensional genome architecture to spread across the X chromosome. *Science* 341:1237973
34. Fabre PJ, Benke A, Joye E, Nguyen Huynh TH, Manley S, Duboule D. 2015. Nanoscale spatial organization of the *HoxD* gene cluster in distinct transcriptional states. *PNAS* 112:13964–69
35. Faure AJ, Schmidt D, Watt S, Schwalie PC, Wilson MD, et al. 2012. Cohesin regulates tissue-specific expression by stabilising highly occupied *cis*-regulatory modules. *Genome Res.* 22:2163–75
36. Fedoriw AM, Stein P, Svoboda P, Schultz RM, Bartolomei MS. 2004. Transgenic RNAi reveals essential function for CTCF in *H19* gene imprinting. *Science* 303:238–40
37. Feldmann A, Ivanek R, Murr R, Gaidatzis D, Burger L, Schübeler D. 2013. Transcription factor occupancy can mediate active turnover of DNA methylation at regulatory regions. *PLOS Genet.* 9:e1003994
38. Filippova D, Patro R, Duggal G, Kingsford C. 2014. Identification of alternative topological domains in chromatin. *Algorithms Mol. Biol.* 9:14
39. Flavahan WA, Drier Y, Liau BB, Gillespie SM, Venteicher AS, et al. 2016. Insulator dysfunction and oncogene activation in IDH mutant gliomas. *Nature* 529:110–14
40. Foti R, Gnan S, Cornacchia D, Dileep V, Bulut-Karslioglu A, et al. 2016. Nuclear architecture organized by Rif1 underpins the replication-timing program. *Mol. Cell* 61:260–73
41. Fudenberg G, Imakaev M, Lu C, Goloborodko A, Abdennur N, Mirny LA. 2015. Formation of chromosomal domains by loop extrusion. bioRxiv. doi: 10.1101/024620
42. Fudenberg G, Mirny LA. 2012. Higher-order chromatin structure: bridging physics and biology. *Curr. Opin. Genet. Dev.* 22:115–24
43. Gerlich D, Koch B, Dupeux F, Peters JM, Ellenberg J. 2006. Live-cell imaging reveals a stable cohesin-chromatin interaction after but not before DNA replication. *Curr. Biol.* 16:1571–78
44. Gibcus JH, Dekker J. 2013. The hierarchy of the 3D genome. *Mol. Cell* 49:773–82
45. Giles KE, Gowher H, Ghirlando R, Jin C, Felsenfeld G. 2010. Chromatin boundaries, insulators, and long-range interactions in the nucleus. *Cold Spring Harb. Symp. Quant. Biol.* 75:79–85
46. Giorgetti L, Galupa R, Nora EP, Piolot T, Lam F, et al. 2014. Predictive polymer modeling reveals coupled fluctuations in chromosome conformation and transcription. *Cell* 157:950–63
47. Goetze S, Mateos-Langerak J, Gierman HJ, de Leeuw W, Giromus O, et al. 2007. The three-dimensional structure of human interphase chromosomes is related to the transcriptome map. *Mol. Cell. Biol.* 2:4475–87
48. Gómez-Marín C, Tena JJ, Acemal RD, Lopez-Mayorga M, Naranjo S, et al. 2015. Evolutionary comparison reveals that diverging CTCF sites are signatures of ancestral topological associating domains borders. *PNAS* 112:7542–47
49. Gröschel S, Sanders MA, Hoogenboezem R, de Wit E, Bouwman BA, et al. 2014. A single oncogenic enhancer rearrangement causes concomitant *EVI1* and *GATA2* deregulation in leukemia. *Cell* 157:369–81
50. Gruber S. 2014. Multilayer chromosome organization through DNA bending, bridging and extrusion. *Curr. Opin. Microbiol.* 22:102–10
51. Gullerova M, Proudfoot NJ. 2008. Cohesin complex promotes transcriptional termination between convergent genes in *S. pombe*. *Cell* 132:983–95
52. Guo C, Yoon HS, Franklin A, Jain S, Ebert A, et al. 2011. CTCF-binding elements mediate control of V(D)J recombination. *Nature* 477:424–30
53. Guo Y, Xu Q, Canzio D, Shou J, Li J, et al. 2015. CRISPR inversion of CTCF sites alters genome topology and enhancer/promoter function. *Cell* 162:900–10
54. Hadjur S, Williams LM, Ryan NK, Cobb BS, Sexton T, et al. 2009. Cohesins form chromosomal *cis*-interactions at the developmentally regulated *IFNG* locus. *Nature* 460:410–13
55. Handoko L, Xu H, Li G, Ngan CY, Chew E, et al. 2011. CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.* 43:630–38
56. Hark AT, Schoenherr CJ, Katz DJ, Ingram RS, Levorse JM, Tilghman SM. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. *Nature* 405:486–89
57. Heger P, Marin B, Bartkuhn M, Schierenberg E, Wiehe T. 2012. The chromatin insulator CTCF and the emergence of metazoan diversity. *PNAS* 109:17507–12

58. Heidari N, Phanstiel DH, He C, Grubert F, Jahanbanian F, et al. 2014. Genome-wide map of regulatory interactions in the human genome. *Genome Res.* 24:1905–17
59. Hnisz D, Weintraub AS, Day DS, Valton AL, Bak RO, et al. 2016. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* 351:1454–58
60. Hu J, Zhang Y, Zhao L, Frock RL, Du Z, et al. 2015. Chromosomal loop domains direct the recombination of antigen receptor genes. *Cell* 163:947–59
61. Hughes JR, Roberts N, McGowan S, Hay D, Giannoulatou E, et al. 2014. Analysis of hundreds of *cis*-regulatory landscapes at high resolution in a single, high-throughput experiment. *Nat. Genet.* 46:205–12
62. Imakaev MV, Fudenberg G, Mirny LA. 2015. Modeling chromosomes: beyond pretty pictures. *FEBS Lett.* 589:3031–36
63. Ing-Simmons E, Seitan VC, Faure AJ, Flicek P, Carroll T, et al. 2015. Spatial enhancer clustering and regulation of enhancer-proximal genes by cohesin. *Genome Res.* 25:504–13
64. Jin F, Li Y, Dixon JR, Selvaraj S, Ye Z, et al. 2013. A high-resolution map of the three-dimensional chromatin interactome in human cells. *Nature* 503:290–94
65. Kagey MH, Newman JJ, Bilodeau S, Zhan Y, Orlando DA, et al. 2010. Mediator and cohesin connect gene expression and chromatin architecture. *Nature* 467:430–35
66. Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi CF, et al. 2000. Functional association of CTCF with the insulator upstream of the *H19* gene is parent of origin-specific and methylation-sensitive. *Curr. Biol.* 10:853–56
67. Katainen R, Dave K, Pitkänen E, Palin K, Kivioja T, et al. 2015. CTCF/cohesin-binding sites are frequently mutated in cancer. *Nat. Genet.* 47:818–21
68. Kemp CJ, Moore JM, Moser R, Bernard B, Teater M, et al. 2014. CTCF haploinsufficiency destabilizes DNA methylation and predisposes to cancer. *Cell Rep.* 7:1020–29
69. Klenova EM, Nicolas RH, Paterson HF, Carne AF, Heath CM, et al. 1993. CTCF, a conserved nuclear factor required for optimal transcriptional activity of the chicken c-*myc* gene, is an 11-Zn-finger protein differentially expressed in multiple forms. *Mol. Cell. Biol.* 13:7612–24
70. Kung JT, Kesner B, An JY, Ahn JY, Cifuentes-Rojas C, et al. 2015. Locus-specific targeting to the X chromosome revealed by the RNA interactome of CTCF. *Mol. Cell* 57:361–75
71. Lavagnolli T, Gupta P, Hörmanseder E, Mira-Bontenbal H, Dharmalingam G, et al. 2015. Initiation and maintenance of pluripotency gene expression in the absence of cohesin. *Genes Dev.* 29:23–38
72. Le Dily F, Baù D, Pohl A, Vicent GP, Soronellas D, et al. 2014. Distinct structural transitions of chromatin topological domains correlate with coordinated hormone-induced gene regulation. *Genes Dev.* 28:2151–62
73. Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, et al. 2004. Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* 430:573–78
74. Lesterlin C, Ball G, Schermelleh L, Sherratt DJ. 2014. RecA bundles mediate homology pairing between distant sisters during DNA break repair. *Nature* 506:249–53
75. Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, et al. 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell* 148:84–98
76. Li L, Lyu X, Hou C, Takenaka N, Nguyen HQ, et al. 2015. Widespread rearrangement of 3D chromatin organization underlies Polycomb-mediated stress-induced silencing. *Mol. Cell* 58:216–31
77. Li W, Notani D, Ma Q, Tanasa B, Nunez E, et al. 2013. Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation. *Nature* 498:516–20
78. Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326:289–93
79. Ling JQ, Li T, Hu JF, Vu TH, Chen HL, et al. 2006. CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nf1*. *Science* 312:269–72
80. Lopez-Serra L, Kelly G, Patel H, Stewart A, Uhlmann F. 2014. The Scc2-Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. *Nat. Genet.* 46:1147–51
81. Losada A. 2014. Cohesin in cancer: chromosome segregation and beyond. *Nat. Rev. Cancer* 14:389–93
82. Lupianez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, et al. 2015. Disruptions of topological chromatin domains cause pathogenic rewiring of gene-enhancer interactions. *Cell* 161:1012–25

83. MacPherson MJ, Sadowski PD. 2010. The CTCF insulator protein forms an unusual DNA structure. *BMC Mol. Biol.* 11:101
84. Matthews AG, Kuo AJ, Ramón-Maiques S, Han S, Champagne KS, et al. 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 450:1106–10
85. Maurano MT, Wang H, John S, Shafer A, Canfield T, et al. 2015. Role of DNA methylation in modulating transcription factor occupancy. *Cell Rep.* 12:1184–95
86. Mazumdar C, Shen Y, Xavy S, Zhao F, Reinisch A, et al. 2015. Leukemia-associated cohesin mutants dominantly enforce stem cell programs and impair human hematopoietic progenitor differentiation. *Cell Stem Cell* 17:675–88
87. McCord RP, Nazario-Toole A, Zhang H, Chines PS, Zhan Y, et al. 2013. Correlated alterations in genome organization, histone methylation, and DNA-lamin A/C interactions in Hutchinson-Gilford progeria syndrome. *Genome Res.* 23:260–69
88. Merkenschlager M, Odom DT. 2013. CTCF and cohesin: linking gene regulatory elements with their targets. *Cell* 152:1285–97
89. Mifsud B, Tavares-Cadete F, Young AN, Sugar R, Schoenfelder S, et al. 2015. Mapping long-range promoter contacts in human cells with high-resolution capture Hi-C. *Nat. Genet.* 47:598–606
90. Minajigi A, Froberg JE, Wei C, Sunwoo H, Kesner B, et al. 2015. A comprehensive Xist interactome reveals cohesin repulsion and an RNA-directed chromosome conformation. *Science* 349:aab2276
91. Mishiro T, Ishihara K, Hino S, Tsutsumi S, Aburatani H, et al. 2009. Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J.* 28:1234–45
92. Mizuguchi T, Fudenberg G, Mehta S, Belton JM, Taneja N, et al. 2014. Cohesin-dependent globules and heterochromatin shape 3D genome architecture in *S. pombe*. *Nature* 516:432–35
93. Mullenders J, Aranda-Orgilles B, Lhoumaud P, Keller M, Pae J, et al. 2015. Cohesin loss alters adult hematopoietic stem cell homeostasis, leading to myeloproliferative neoplasms. *J. Exp. Med.* 212:1833–50
94. Murayama Y, Uhlmann F. 2015. DNA entry into and exit out of the cohesin ring by an interlocking gate mechanism. *Cell* 163:1628–40
95. Nagano T, Lubling Y, Stevens TJ, Schoenfelder S, Yaffe E, et al. 2013. Single-cell Hi-C reveals cell-to-cell variability in chromosome structure. *Nature* 502:59–64
96. Nakahashi H, Kwon KR, Resch W, Vian L, Dose M, et al. 2013. A genome-wide map of CTCF multivalency redefines the CTCF code. *Cell Rep.* 3:1678–89
97. Narendra V, Rocha PP, An D, Raviram R, Skok JA, et al. 2015. CTCF establishes discrete functional chromatin domains at the *Hox* clusters during differentiation. *Science* 347:1017–21
98. Nasmyth K. 2001. Disseminating the genome: joining, resolving, and separating sister chromatids during mitosis and meiosis. *Annu. Rev. Genet.* 35:673–745
99. Nasmyth K, Haering CH. 2009. Cohesin: its roles and mechanisms. *Annu. Rev. Genet.* 43:525–58
100. Nativio R, Wendt KS, Ito Y, Huddleston JE, Uribe-Lewis S, et al. 2009. Cohesin is required for higher-order chromatin conformation at the imprinted *IGF2-H19* locus. *PLOS Genet.* 5:e1000739
101. Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, et al. 2013. Organization of the mitotic chromosome. *Science* 342:948–53
102. Nichols MH, Corces VG. 2015. A CTCF code for 3D genome architecture. *Cell* 162:703–5
103. Nora EP, Dekker J, Heard E. 2013. Segmental folding of chromosomes: a basis for structural and regulatory chromosomal neighborhoods? *BioEssays* 35:818–28
104. Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* 485:381–85
105. Pant V, Kurukuti S, Pugacheva E, Shamsuddin S, Mariano P, et al. 2004. Mutation of a single CTCF target site within the *H19* imprinting control region leads to loss of *Igf2* imprinting and complex patterns of de novo methylation upon maternal inheritance. *Mol. Cell. Biol.* 24:3497–504
106. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* 132:422–33
107. Peric-Hupkes D, Meuleman W, Pagie L, Bruggeman SW, Solovei I, et al. 2010. Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* 38:603–13
108. Phillips JE, Corces VG. 2009. CTCF: master weaver of the genome. *Cell* 137:1194–211

109. Pope BD, Ryba T, Dileep V, Yue F, Wu W, et al. 2014. Topologically associating domains are stable units of replication-timing regulation. *Nature* 515:402–5
110. Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159:1665–80
111. Remeseiro S, Cuadrado A, Gómez-López G, Pisano DG, Losada A. 2012. A unique role of cohesin-SA1 in gene regulation and development. *EMBO J.* 31:2090–102
112. Rowley JD. 1973. A new consistent chromosomal abnormality in chronic myelogenous leukaemia identified by quinacrine fluorescence and Giemsa staining. *Nature* 243:290–93
113. Rubio ED, Reiss DJ, Welcsh PL, Disteche CM, Filippova GN, et al. 2008. CTCF physically links cohesin to chromatin. *PNAS* 105:8309–14
114. Saldaña-Meyer R, González-Buendía E, Guerrero G, Narendra V, Bonasio R, et al. 2014. CTCF regulates the human p53 gene through direct interaction with its natural antisense transcript, Wrap53. *Genes Dev.* 28:723–34
115. Sanborn AL, Rao SSP, Huang S-C, Durand NC, Huntley MH, et al. 2015. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *PNAS* 112:E6456–65
116. Sandhu KS, Li G, Poh HM, Quek YL, Sia YY, et al. 2012. Large-scale functional organization of long-range chromatin interaction networks. *Cell Rep.* 2:1207–19
117. Sanyal A, Lajoie BR, Jain G, Dekker J. 2012. The long-range interaction landscape of gene promoters. *Nature* 489:109–13
118. Schmidt D, Schwalie PC, Ross-Innes CS, Hurtado A, Brown GD, et al. 2010. A CTCF-independent role for cohesin in tissue-specific transcription. *Genome Res.* 20:578–88
119. Schoenfelder S, Furlan-Magaril M, Mifsud B, Tavares-Cadete F, Sugar R, et al. 2015. The pluripotent regulatory circuitry connecting promoters to their long-range interacting elements. *Genome Res.* 25:582–97
120. Seitan VC, Faure AJ, Zhan Y, McCord RP, Lajoie BR, et al. 2013. Cohesin-based chromatin interactions enable regulated gene expression within preexisting architectural compartments. *Genome Res.* 23:2066–77
121. Seitan VC, Hao B, Tachibana-Konwalski K, Lavagnolli T, Mira-Bontenbal H, et al. 2011. A role for cohesin in T cell receptor rearrangement and thymocyte differentiation. *Nature* 476:467–71
122. Shrimali S, Srivastava S, Varma G, Grinberg A, Pfeifer K, Srivastava M. 2012. An ectopic CTCF-dependent transcriptional insulator influences the choice of Vβ gene segments for VDJ recombination at TCRβ locus. *Nucleic Acids Res.* 40:7753–65
123. Shukla S, Kavak E, Gregory M, Imashimizu M, Shutinoski B, et al. 2011. CTCF-promoted RNA polymerase II pausing links DNA methylation to splicing. *Nature* 479:74–79
124. Skibbens RV. 2010. Buck the establishment: reinventing sister chromatid cohesion. *Trends Cell Biol.* 20:507–13
125. Sleutels F, Soochit W, Bartkuhn M, Heath H, Dienstbach S, et al. 2012. The male germ cell gene regulator CTCFL is functionally different from CTCF and binds CTCF-like consensus sites in a nucleosome composition-dependent manner. *Epigenetics Chromatin* 5:8
126. Sofueva S, Yaffe E, Chan WC, Georgopoulou D, Vietri Rudan M, et al. 2013. Cohesin-mediated interactions organize chromosomal domain architecture. *EMBO J.* 32:3119–29
127. Soh YM, Bürmann F, Shin HC, Oda T, Jin KS, et al. 2015. Molecular basis for SMC rod formation and its dissolution upon DNA binding. *Mol. Cell* 57:290–303
128. Solovei I, Kreysing M, Lanctôt C, Kösem S, Peichl L, et al. 2009. Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. *Cell* 137:356–68
129. Soshnikova N, Montavon T, Leleu M, Galjart N, Duboule D. 2010. Functional analysis of CTCF during mammalian limb development. *Dev. Cell* 19:819–30
130. Splinter E, de Wit E, Nora EP, Klous P, van de Werken HJ, et al. 2011. The inactive X chromosome adopts a unique three-dimensional conformation that is dependent on Xist RNA. *Genes Dev.* 25:1371–83
131. Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, et al. 2006. CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus. *Genes Dev.* 20:2349–54
132. Stanhope-Baker P, Hudson KM, Shaffer AL, Constantinescu A, Schlissel MS. 1996. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity in vitro. *Cell* 85:887–97

133. Stark WM, Sherratt DJ, Boocock MR. 1989. Site-specific recombination by Tn3 resolvase: topological changes in the forward and reverse reactions. *Cell* 58:779–90
134. Stedman W, Kang H, Lin S, Kissil JL, Bartolomei MS, Lieberman PM. 2008. Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J.* 27:654–66
135. Stegle O, Teichmann SA, Marioni JC. 2015. Computational and analytical challenges in single-cell transcriptomics. *Nat. Rev. Genet.* 16:133–45
136. Suter DM, Molina N, Gatfield D, Schneider K, Schibler U, Naef F. 2011. Mammalian genes are transcribed with widely different bursting kinetics. *Science* 332:472–74
137. Symmons O, Uslu VV, Tsujimura T, Ruf S, Nassari S, et al. 2014. Functional and topological characteristics of mammalian regulatory domains. *Genome Res.* 24:390–400
138. Tang Z, Luo OJ, Li X, Zheng M, Zhu JJ, et al. 2015. CTCF-mediated human 3D genome architecture reveals chromatin topology for transcription. *Cell* 163:1611–27
139. Tark-Dame M, Jerabek H, Manders EMM, Heermann DW, van Driel R. 2014. Depletion of the chromatin looping proteins CTCF and cohesin causes chromatin compaction: insight into chromatin folding by polymer modelling. *PLOS Comput. Biol.* 10:e1003877
140. Tsujimoto Y, Finger LR, Yunis J, Nowell PC, Croce CM. 1984. Cloning of the chromosome breakpoint of neoplastic B cells with the t(14;18) chromosome translocation. *Science* 226:1097–99
141. van de Corput MP, de Boer E, Knoch TA, van Cappellen WA, Quintanilla A, et al. 2012. Super-resolution imaging reveals three-dimensional folding dynamics of the β-globin locus upon gene activation. *J. Cell Sci.* 125:4630–39
142. van Steensel B. 2011. Chromatin: constructing the big picture. *EMBO J.* 30:1885–95
143. Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, et al. 2015. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep.* 10:1297–309
144. Viny AD, Ott CJ, Spitzer B, Rivas M, Meydan C, et al. 2015. Dose-dependent role of the cohesin complex in normal and malignant hematopoiesis. *J. Exp. Med.* 212:1819–32
145. Wan LB, Pan H, Hannenhalli S, Cheng Y, Ma J, et al. 2008. Maternal depletion of CTCF reveals multiple functions during oocyte and preimplantation embryo development. *Development* 135:2729–38
146. Wang H, Maurano MT, Qu H, Varley KE, Gertz J, et al. 2012. Widespread plasticity in CTCF occupancy linked to DNA methylation. *Genome Res.* 22:1680–88
147. Wendt KS, Yoshida K, Itoh T, Bando M, Koch B, et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* 451:796–801
148. Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, et al. 2013. Master transcription factors and Mediator establish super-enhancers at key cell identity genes. *Cell* 153:307–19
149. Wijchers PJ, Krijger PH, Geeven G, Zhu Y, Denker A, et al. 2016. Cause and consequence of tethering a subTAD to different nuclear compartments. *Mol. Cell* 61:461–73
150. Wu LJ, Errington J. 1994. *Bacillus subtilis* SpoIIIE protein required for DNA segregation during asymmetric cell division. *Science* 264:572–75
151. Xiao T, Wallace J, Felsenfeld G. 2011. Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol. Cell. Biol.* 31:2174–83
152. Xu C, Corces VG. 2016. Towards a predictive model of chromatin 3D organization. *Semin. Cell Dev. Biol.* In press. doi: 10.1016/j.semcdb.2015.11.013
153. Xu N, Donohoe ME, Silva SS, Lee JT. 2007. Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. *Nat. Genet.* 39:1390–96
154. Yan J, Enge M, Whitington T, Dave K, Liu J, et al. 2013. Transcription factor binding in human cells occurs in dense clusters formed around cohesin anchor sites. *Cell* 154:801–3
155. Yang F, Deng X, Ma W, Berletch JB, Rabaia N, et al. 2015. The lncRNA *Firre* anchors the inactive X chromosome to the nucleolus by binding CTCF and maintains H3K27me3 methylation. *Genome Biol.* 16:52
156. Zhang N, Pati D. 2009. Handcuff for sisters: a new model for sister chromatid cohesion. *Cell Cycle* 8:399–402
157. Zhang Y, Wong C-H, Birnbaum RY, Li G, Favaro R, et al. 2013 . Chromatin connectivity maps reveal dynamic promoter-enhancer long-range associations. *Nature* 504:306–10

158. Zhao H, Sifakis EG, Sumida N, Millán-Ariño L, Scholz BA, et al. 2015. PARP1- and CTCF-mediated interactions between active and repressed chromatin at the lamina promote oscillating transcription. *Mol. Cell* 59:984–97
159. Zhao Z, Tavoosidana G, Sjölinder M, Göndör A, Mariano P, et al. 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.* 38:1341–47
160. Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, et al. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *PNAS* 111:996–1001