Cell. Mol. Life Sci. 66 (2009) 1647 – 1660
1420-682X/09/101647-14
DOI 10.1007/s00018-009-8647-z

Cellular and Molecular Life Sciences

Review

CCCTC-binding factor: to loop or to bridge

J. Zlatanova[a, *] and P. Caiafa[b]

[a] Department of Molecular Biology, University of Wyoming, Laramie, WY 82071 (USA), Fax: +1 307 766 2892, e-mail: jordanka@uwyo.edu
[b] Department of Cellular Biotechnology and Hematology, Second Faculty of Medicine and Surgery, University "La Sapienza" and Pasteur Institute-Fondazione Cenci Bolognetti, Rome (Italy)

Received 10 October 2008; received after revision 11 December 2008; accepted 16 December 2008
Online First 14 January 2009

Abstract. Eukaryotic genomes have complex spatial organization in the nucleus. The factors and the mechanisms involved in this organization remain an enigma. Among the many proteins implicated in such a role, the ubiquitous Zn-finger protein CTCF stands out. Here we summarize the evidence placing CTCF in the enviable position of a master organizer of the genome. CTCF can form loops *in cis*, and can bridge sequences located on different chromosomes *in trans*. The thousands of CTCF binding sites, identified in recent genome-wide localization studies, and their distribution along the genome further support a crucial role of CTCF as a chromatin organizer.

Keywords. Genome organization, transcription, imprinting, boundary elements, X-chromosome inactivation, intra- and inter-chromosomal interactions.

Genome organization in the eukaryotic nucleus

Nowadays, there is a clear understanding that gene organization along the linear genomic sequences and the spatial organization of the genome in the nucleus are of paramount importance for gene regulation [1, 2]. Thus, housekeeping genes often cluster in gene-dense regions [3, 4], presumably benefiting from open chromatin domains over the cluster and collective positioning at nuclear zones of high transcription competence (transcription 'factories', see refs. [5, 6]). Developmentally regulated genes are also organized in (smaller) clusters, the best studied examples of which are the chicken and mammalian β-globin gene clusters. The individual genes in the cluster compete for physical contact with a shared upstream regulatory region, the locus control region (LCR), thus forming chromatin loops between individual promoters and LCR [7, 8]. These loops are developmentally regulated, with the LCR changing gene partners, depending on which gene is to be expressed at a given point in development. The contacts between the different regions of the chromatin fiber are believed to occur through random collisions and are then stabilized by protein binding [2].

The transcriptional effects of the intra-chromosomal gene organization are complemented by inter-chromosomal interactions. Recent genome-wide studies introduced modified chromosome conformation capture (3C) [9] techniques that allow unbiased discovery of interacting chromosome regions [10–14]. These so-called 4C methods (3C methods which use genome-wide microarray or sequencing analyses of the interacting regions) have recognized the existence of thousands of interactions between each target sequence and many DNA fragments scattered throughout the genome (reviewed in [15]). The majority of

* Corresponding author

interactions occur through particular DNA topologies that are very cell-specific. Such cell-to-cell differences in topology are expected to exist according to the suggested self-organization model of nuclear architecture [1, 16].

Finally, it is widely believed that each interphase chromosome (the chromatin fiber organizing the uninterrupted linear DNA molecule in each chromosome) occupies its own space in the nucleus. These so-called chromosome territories (Fig. 1) are dispersed over the nuclear volume, with some recognizable patterns of organization. For example, smaller and/or gene-rich chromosomes tend to localize more centrally in the nucleus, whereas larger and/or gene-poor chromosomes prefer more peripheral positions [17, 18]. (Note that some tissue-specific preferences for co-localization of individual chromosomes may also exist [19].) A polarized distribution of gene-dense vs. gene-poor chromatin has been convincingly demonstrated *within* individual chromosome territories, with the gene-dense regions located towards the nuclear interior [20–22].

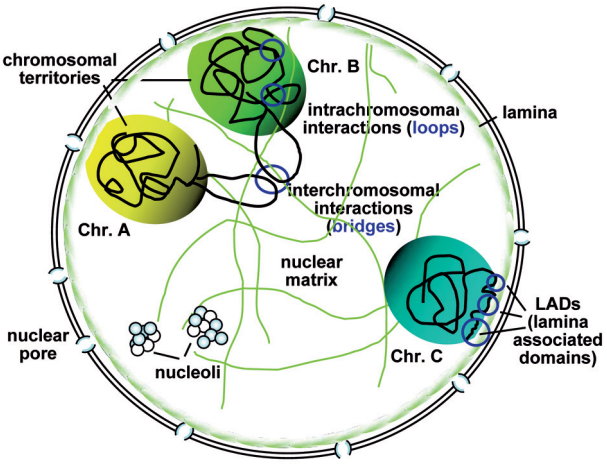

Figure 1. Genome organization within the eukaryotic nucleus. The linear double stranded DNA molecules that constitute individual interphase chromosomes occupy distinct portions of the nuclear volume, forming chromosome territories (denoted here in three different colors for three example chromosomes). The gradient shading of the three chromosomal territories reflects the radial distribution of the gene-rich vs. gene-poor regions (see text). The chromatin fiber within each territory can form individual loop domains, whereas domains making temporal excursions out of their respective territories can be bridged together, presumably to be coordinately regulated (through interactions with transcriptional factories, for example). A significant portion of the chromatin fiber is associated with the lamina structure in the so-called lamina-associated domains (LADs) that are transcriptionally repressed. Intra- and inter-chromosomal interactions also form distinct nucleolar structures that contain both the active and inactive copies of the tandemly-repeated ribosomal genes. A final recognizable structure that may have a role in the spatial and topological organization of the genome is the nuclear matrix, an insoluble meshwork of various 'skeletal' proteins.

Despite the fact that each chromosome occupies its own territory, chromatin fibers often 'play hooky', looping out of the territory to interact with escapee loops from other chromosomes. These 'external' (to the chromosome territories) loops are often gene-rich and the frequency with which they form tends to correlate with the transcriptional activity of the resident genes [23]. This behavior may be needed for the genes present in the loops to interact with transcription or replication factories [5, 6, 24]. It must be noted, however, that the existence of such pre-assembled factories that attract genes for transcription/replication is considered rather controversial [1, 15]. Indeed, such factories probably exist when the transcriptional machinery and all other required factors transiently gather on a transcribed gene; however, they may assemble one factor at a time in a stochastic collision process, rather than exist as preassembled stable entities waiting to be visited by genes [1].

Thus, it becomes increasingly clear that the formation of intra-chromosomal loops *in cis* and of bridges between chromatin sites *in trans* are dynamic processes that organize nuclear structure. In some cases, the functional significance of these interactions has been deciphered; however, the majority of interactions still await assignment of function.

When loops or bridges form, whether by a deterministic process or stochastically, they need to be stabilized, at least for the time of the function of the loop. How does stabilization happen? It has long been presumed that protein factor binding to the two DNA helices at the base of the loop is the mechanism of stabilization. Among the proven or presumed proteins are components of the insoluble nuclear matrix, Topoisomerase II, SATB1, transcription termination factors, lamina proteins, and last but not least, the transcription or replication factories. Marenduzzo et al. [24] have recently introduced the term "molecular ties" for such proteins, arguing for a major role of transcription factories as such ties.

Is there a universal abundant protein factor that can serve to form and/or stabilize loops and bridges? What would be the properties required of such a factor? How would its activities be regulated? In what follows, we will argue that CTCF may be playing such a universal role in topologically organizing the genome as a whole or some specific regions. We will also discuss the relationship of this organization to regulation of gene expression, whenever such data are available.

CTCF: the protein and its genome-wide distribution

CTCF is a ubiquitous, abundant nuclear protein with very diverse functions, including enhancer-blocking, X-chromosome inactivation, gene imprinting, and gene activation or repression (Table 1). CTCF structure can be subdivided in three distinct domains: an N-terminal region, a central domain containing 11 zinc-fingers, and a C-terminal region (reviewed in ref. [53]). The three distinct domains provide interaction platforms for various proteins and contain sites for distinct posttranslational modifications: the N-terminus is poly(ADP-ribosyl)ated [54], whereas the C domain contains several sites for phosphorylation by protein kinase CK2 [55]. The protein is characterized by a relatively uniform nuclear distribution in interphase, with additional binding sites at the periphery of the nucleolus. CTCF also binds to the nuclear matrix, indicating a possible functional connection between CTCF-dependent insulator elements and the nuclear matrix [56].

CTCF became a major research focus when it was realized that it is critically involved in enhancer-blocking functions in higher eukaryotes. Felsenfeld's laboratory identified CTCF binding to specific sites in the β-globin locus and in the imprinted *Igf2/H19* locus [26, 34] and proved that CTCF binding was necessary for the enhancer-blocking functions at both loci. Importantly, CTCF binds only when the binding site is unmethylated, and thus differentiates between the differentially methylated maternal and paternal alleles of the *Igf2/H19* locus. CTCF is the only recognized protein in higher eukaryotes that, by binding to specific sites, exerts insulator functions [57–59].

A major step forward in our understanding of possible CTCF functions came from genome-wide localization studies which combined chromatin immunoprecipitation (ChIP) with analysis of the precipitated DNA by microarrays (ChIP-on-chip) [60], or by sequencing (ChIP-seq) [61]. The ChIP analyses were complemented by a computational approach [62]. A detailed description of the genome-wide studies is presented in Table 2, which also covers major findings from 3C and 4C approaches (see below). The results of these fascinating studies revealed unexpected features of CTCF localization genome-wide. First, the number of CTCF binding sites is enormous (between 14,000 and 20,000). The sites are generally correlated with genes. The genome-wide distribution, according to Barski et al. [61], is 41% in intergenic regions, 31% in transcribed regions and 28% within 2 kb of transcription start sites (note, however, that the other two studies found CTCF binding sites to be generally far from promoters). Almost 14,000 genomic regions are flanked by CTCF on both sides, forming CTCF-pair-defined domains (CPD) of average size of 210 kb, each containing 2.5 genes on the average [60]. Simple math makes us realize that practically all genes in the human genome are organized in these domains. These "average" numbers may be hiding a more complex reality, since the CTCF domains actually fall into three categories: (i) average domains, (ii) very large domains that contain clusters of co-regulated genes, and (iii) domains that contain numerous CTCF sites over relatively short distances; these occur in genes containing multiple promoters and may be involved in promoter choice.

The results from another recent genome-wide study reveal an intriguing relationship between CTCF presence in the human genome and domains that bind to lamin B1 [64] (Table 2). Close to 1,400 very large (median size ~550 kb) chromatin domains are bound to lamin B1; these so-called lamina-associated domains (LADs) contain few genes in a repressive chromatin environment. In many cases CTCF binding sites occur just outside the LAD domains, the CTCF peak centering at 5–10 kb outside the LAD borders. The CTCF peak does not coincide with the peak of promoter enrichment.

Thus, the general picture that emerges from these studies is as follows: 40% of the genome is sequestered into LADs that provide a repressive environment to the few genes present. CTCF is excluded from these domains. The rest of the genome contains thousands of CTCF binding sites, generally correlated with genes, probably far from promoters, and often displaying non-random distribution (impoverished over clusters of co-regulated genes and enriched over multiple-promoter genes).

Earlier transfection experiments [65] reported tethering of CTCF binding sites to the periphery of the nucleolus. The nucleolar localization was attributed to the direct interaction between CTCF and the abundant nucleolar protein nucleophosmin/B23. Nucleolar localization of CTCF was also reported in cells overexpressing CTCF; the transfer of CTCF to the nucleolus was dependent on poly(ADP-ribosyl)ation and inhibited ribosomal gene transcription [66]. CTCF binding sites with insulator function in enhancer-blocking assays were also identified in control regions of rDNA loci in *Xenopus* [26]. Although the detailed molecular mechanisms behind the nucleolar function of CTCF remain to be established, it is clear that the nucleolus presents a chromatin environment for abundant CTCF binding.

All these data point to an important, maybe crucial, role of CTCF in organizing the genome. The genome-wide data, as revealing as they are, open a plethora of questions about the functional significance of CTCF binding. Below, we will describe in some detail the

Table 1. Involvement of CTCF in diverse cellular functions

Function in	CTCF Role(s)	Key references
Locus Control Regions (LCRs) in clusters of developmentally regulated genes	β-globin cluster in chicken: Enhancer-blocking insulator function at both 5' and 3' DNase hypersensitive sites; 5'HS site 4 serves as paradigm of enhancer-blocking insulators;	[25, 26]
	β-globin clusters in mouse and humans: role in organizing constitutive loops in progenitor non-expressing cells; LCR-gene contacts needed for expression are established within these loops during differentiation into erythroid cells with the help of TFs	[27]
X-chromosome inactivation	Essential for X-chromosome pairing that occurs just before the onset of X-chromosome inactivation;	[28–30]
	CTCF, in complex with Yy1, transactivates transcription of *Tsix*, the binary switch controlling *Xist* transcription;	[31]
	Forms boundaries between escape and inactivated gene on the inactive X-chromosome;	[32]
	Possible role in formation of two chromatin loops at the X-inactivation center	[33]
Gene imprinting at the *Igf1/H19* and other imprinted loci	Controls the imprinted repression of *Igf2* in the maternal allele through binding to the unmethylated Imprinting Control Region (ICR)	[34–37]
	Protects ICR at the maternal allele of the *Igf2/H19* imprinted locus against DNA methylation (in somatic cells);	[38–40]
	Regulates asynchronous replication of the imprinted *Igf2/H19* alleles	[41]
Regulation of gene expression	*c-myc* gene expression (mechanism of repression controversial);	[42, 43]
	Activates the amyloid β-protein precursor gene in response to growth factors;	[44]
	Represses chicken lysozyme gene through interaction with a silencer;	[45]
	Creates a boundary against spreading of DNA methylation in the promoter of *BRCA1* gene;	[46]
	Controls the stable expression of human *Rb* gene by protecting its CpG islands from methylation;	[47]
	Forms a methylation-dependent insulator at CTG repeats contributing to differential expression of neighboring genes in *DM1* locus (involved in myotonic dystrophy); controls antisense transcription;	[48, 49]
	Regulates co-expressed genes through loop formation (*MHCII* genes);	[50, 51]
	Activates PARP-1	[52]

Table 2. Genome-wide analysis of CTCF binding site localization and of nuclear organization of genes/regions containing such sites

Site studied	Method	Major findings	Relevance to CTCF	Reference
CTCF binding sites genome wide	Creation of a CTCF library by cloning DNA fragments ChIPed with C-terminal CTCF antibody (mouse fetal liver)	>200 new target sites identified in a wide variety of genes and presumably in heterochomatin domains		[63]
H19 ICR (imprinting control region)	Associated Chromosome Trap (modification of Chromosome Conformation Capture, 3C) and Fluorescence *in situ* hybridization (FISH) applied to mouse bone marrow fibroblast cell line	The *H19* ICR maternal allele on chromosome 7 interacts with the paternal allele of an intergenic region between imprinted genes *Wsb1* and *Nf1* on chromosome 11 (these genes are paternally expressed); Loss of contact results in decrease of *Wsb1* and *Nf1* expression, suggesting that *Wsb1* and *Nf1* are recruited to a transcription factory through CTCF	Both interacting alleles contain bound CTCF; Knock-down of CTCF or deletion of maternal ICR abolish interaction	[12]
H19 ICR	Circular Chromosome Conformation Capture (4C) followed by cloning and sequencing of amplified fragments	114 unique sequences from all chromosomes interact with ICR, with some preference for intra-chromosome interactions; Imprinted loci are strongly overrepresented in the 4C library, with contacts affecting transcriptional status *in trans*; Some contacts are primarily with the maternally-inherited ICR; Interacting DNA regions change during ES differentiation and also differ between embryoid bodies and neonatal liver	Physical proximity of sites depends on intact CTCF target sites	[13]
β-globin locus in active (fetal liver) and inactive (fetal brain) state; *Rad23a* (a housekeeping gene)	(Circular) Chromosome Conformation Capture followed by microarray analysis (3C-on-chip, termed 4C); the chip contains probes located within 100 bp from the sites of initial restriction endonuclease cutting of cross-linked DNA and covers seven complete mouse chromosomes	Active β-globin locus preferentially interacts with other transcribed loci on the same chromosome; *Rad23a* also interacts with active regions; Inactive β-globin locus interacts with other transcriptionally silent loci	Contacts to hypersensitive site 2 of ICR are identified: CTCF involvement not directly addressed	[14]
CTCF binding sites in primary human fibroblasts	ChIP-on-chip; A series of 38 arrays that contain a total of 14.6 million 50-mer oligonucleotides positioned every 100 bp along the non-repeated sequence of the human genome; validation of results by conventional ChIP on randomly selected binding sites	13,804 CTCF sites correlate with genes but are generally very far from promoters (average distance ~48 kb); 13,766 genomic regions are flanked by CTCF on both sides (CPD, CTCF-pair-defined domains): average size of CPDs 210 kb with 2.5 genes inside; Two types of distinct domains: CTCF-depleted (contain clusters of co-regulated genes) and CTCF-enriched (contain genes with multiple promoters); A consensus CTCF-binding motif defined (18% of sites do not contain the consensus); Most sites are also occupied in other cell types, with a fraction subject to cell-type-dependent regulation		[60]

Table 2 *(Continued)*

Site studied	Method	Major findings	Relevance to CTCF	Reference
Search for regulatory motifs in conserved non-coding elements (CNEs) in the human genome	Computational approaches to create a data set of ~830,000 CNEs (conserved over 12 mammalian genomes) and then create a catalog of enriched sequence motifs; CTCF and RFX1 (Regulatory Factor X) sites confirmed by *in vitro* affinity capture from HeLa nuclear extract	Only 16 known regulatory elements among the 223 discovered motifs; Three of the top 50 most highly enriched motifs contained CTCF binding sites, totaling 14,987 sites; CTCF sites closely follow genes distribution but are far from promoters; Comparison with expression patterns of divergent co-regulated gene pairs suggests that the majority of CTCF sites function as insulators; A consensus CTCF-binding motif defined (resembles the one defined in [55])		[62]
20 histone methylations, H2A.Z replacement variant, Pol II and CTCF	ChIPed DNA analyzed by Solexa 1G Genome Analyzer and mapped on human genome (ChIP-seq) ($CD4^+$T cells)	20,262 CTCF sites identified (41 % in intergenic regions, 31 % in transcribed regions, 28 % within 2 kb of TSS); Sites are enriched in H3K4me1/2/3, H3K9me and in H2A.Z		[61]
Lamina- Associated Domains (LADs) in human genome	DamID mapping in cultured human lung fibroblasts (chimeric DNA adenine methyltransferase-lamin B1 protein introduced in cells, followed by amplification of methylated DNA and high-density microarray probing of human chromosome 4 and then the entire human genome	Identified 1,344 LADs (median size 553 kb, 40 % of the genome) containing gene-poor regions in repressive chromatin environment; Computationally discovered that 22 % of LADs have CTCF sites on one side, and 2 % on both sides; the peak of CTCF sites centers at 5–10 kb outside the LAD borders and does not coincide with the peak of promoter enrichment	CTCF sites show specific localization outside of LADs, consistent with insulator function	[64]

gene systems studied so far to illustrate the role of CTCF in forming loops and bridges, and the functional significance of these loops and bridges, wherever such data are available.

CTCF in the organization of loops *in cis*

Despite the numerous reports of CTCF binding to various loci, only in few cases do we know that CTCF binding is involved in loop formation *in cis*. Below, we will present a few well characterized examples.

The β-globin locus

CTCF involvement in enhancer-blocking functions was first recognized in studies of the chicken β-globin gene cluster [26]. Subsequently, it was reported that this function is conserved in the mouse and human clusters. Fig. 2 depicts the map of the mouse locus, which is embedded in a large heterochromatic region containing numerous olfactory receptor genes on both sites of the four developmentally-regulated globin genes and the locus control region [7, 27]. The locus contains two CTCF-dependent enhancer-blocking elements, located at DNase I hypersensitive sites at its 5'- and 3'-borders. Two additional CTCF sites have been identified, deep in the 5' cluster of olfactory receptor genes. 3C analysis has detected CTCF-dependent interactions among all four sites in both progenitor (non-expressing) and erythroid (expressing) cells, with the intervening chromatin fiber looping out. No such interactions are detected in brain cells, where the locus adopts a seemingly linear conformation [7]. Thus, chromatin loops form following commitment to erythroid differentiation, and then exist constitutively, irrespectively of the actual expression status of the globin genes. The contacts between the LCR and the promoters of the individual genes to be expressed are established later during differentiation and depend on binding of transcription factors (TFs), such as the erythroid Krüppel-like factor (EKLF) and the globin transcription factor 1 (GATA-1) (Fig. 2A). It should be noted that the functional significance of these CTCF-mediated loops remains obscure, since earlier work has demonstrated that the presence of these sites is dispensable for high-level β-globin transcription [27, 68].

A recent series of papers from de Laat's group has significantly contributed to our understanding of the function of the LCR in loop formation. Noordermeer et al. [69] targeted the human β-globin LCR in two opposite orientations to a transcriptionally active gene-dense region of the mouse genome. The locus resided mostly at the edge of its chromosome territory and LCR integration caused the region to occasionally escape from its territory. Since there was a distinct effect on transcription of individual neighboring genes on either side of the integrated locus, and, in some cases, this effect was orientation-dependent, it was concluded that nuclear repositioning was not sufficient to increase transcription. Rather, it was the extent of loop formation between LCR and individual genes that determined the transcriptional effects of the integrated element. It was also demonstrated that looping within the β-globin locus and the long-range interactions *in trans* (bridging) persist following inhibition of Pol II transcription by drugs [67]. These data led to the recognition of three distinct states of the locus (Fig. 2B). Importantly, once the locus organization (both in terms of loops *in cis* and bridges *in trans*) and the chromatin modification status are established during differentiation, they are maintained whether or not Pol II is bound and the gene is actively transcribed. In relevance to our topic, it is certain that at least the constitutive loops established *in cis* at the onset of differentiation are CTCF-mediated. In view of other data (see below), we suspect that the bridges may also be CTCF-dependent.

The major histocompatibility gene cluster

Another interesting example of the role of CTCF in forming intra-chromosomal loops *in cis* comes from work on the organization and expression of two genes in the major histocompatibility class II cluster [50, 51]. The two genes studied, *HLA-DRB1* and *HLA-DQA1*, are divergently transcribed in a co-regulated manner. A region situated in the intergenic region, *XL9*, is a binding site for CTCF (Fig. 3A); deletion of this site diminishes expression of the genes. 3C analysis demonstrated loop formation between *XL9* and the proximal promoter regions of the two genes. CTCF collaborated with two other proteins to form the loops: RFX (a transcription factor) and CIITA (a transcriptional coactivator), known to interact with conserved promoter sequences YXW. Figure 3A depicts the two possible states of the locus; transcriptional activity is only observed upon loop formation.

The mouse imprinted *Igf2/H19* locus

The mouse *Igf2/H19* locus is the best characterized imprinted locus, in which the mutually exclusive transcription of the two residing genes, *Igf2* and *H19*, is determined by the differential DNA methylation patterns present on the maternal and the paternal alleles (see Fig. 3B, its legend, and Table 1 for more details). The imprinting control region (ICR) that governs the allele-specific expression of the two genes is unmethylated on the maternal allele and binds CTCF [34]. 3C analysis has demonstrated

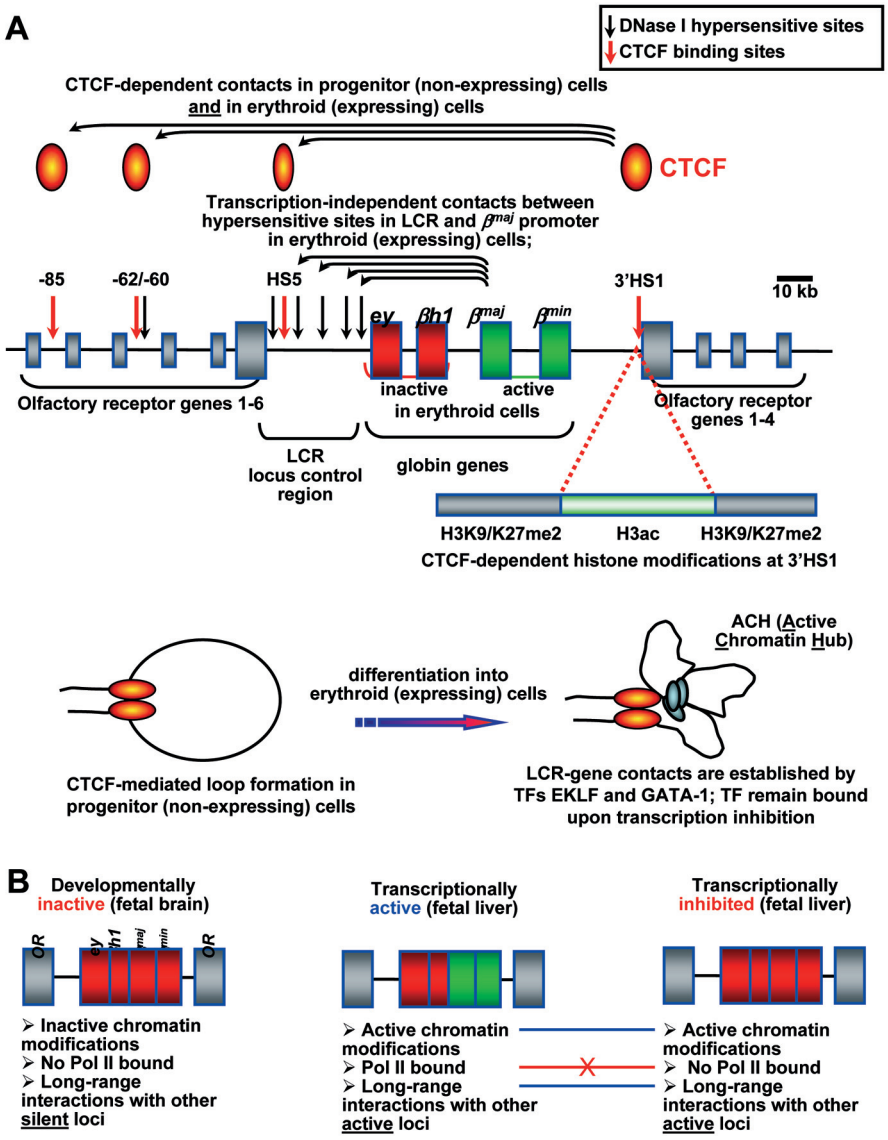

Figure 2. The CTCF-mediated loop structure in the mouse β-globin gene locus and its relation to transcription. (*A*) Schematic depicts the developmentally regulated β-globin gene cluster and its locus control region (LCR) which encompasses several DNase I hypersensitive sites (HS). Sites HS5 and 3'HS1 flank the entire locus and function as enhancer-blocking elements. The locus is embedded in the highly compacted chromatin structure of the silent olfactory receptor genes. CTCF binds four sites in the locus: the two enhancer-blocking elements HS5 and 3'HS1 and two more sites further upstream in the olfactory receptor gene cluster. These sites contact each other in both progenitor and erythroid cells [bottom schematic in *(A)*]. Further smaller loops between the LCR and the promoters of individual globin genes are established with the help of transcription factors (TFs) during erythroid differentiation. These contacts are, however, also transcription-independent, since the TFs remain bound upon transcription inhibition. Thus, additional unidentified mechanism(s) must be involved in the actual control of transcription. (*B*) Schematics of the globin gene locus in a developmentally inactive state (left), a transcriptionally active state (middle), and a transcriptionally inhibited state (right). Note the differences among the three states in chromatin modifications, Pol II binding, and long-range interactions *in trans.* Figure based mainly on [7, 8, 27, 67].

contacts between the ICR and DMR1, another differentially methylated region, on the maternal allele, with the intervening DNA looping out [37] (Fig. 3B). This topology protects the promoter of the *Igf2* gene from interaction with the two enhancer elements, 3' from the *H19* gene. CTCF bound to ICR 'glues' the two DNA regions to stabilize the loop.

The X-chromosome inactivation center

In mammals, gene dosage compensation between XX females and XY males occurs through a random inactivation of one of the two X-chromosomes. A recent important study addresses the higher-order chromatin organization at the X-inactivation center before, at the onset of, and post X-chromosome inactivation [33]. The inactivation process is complex

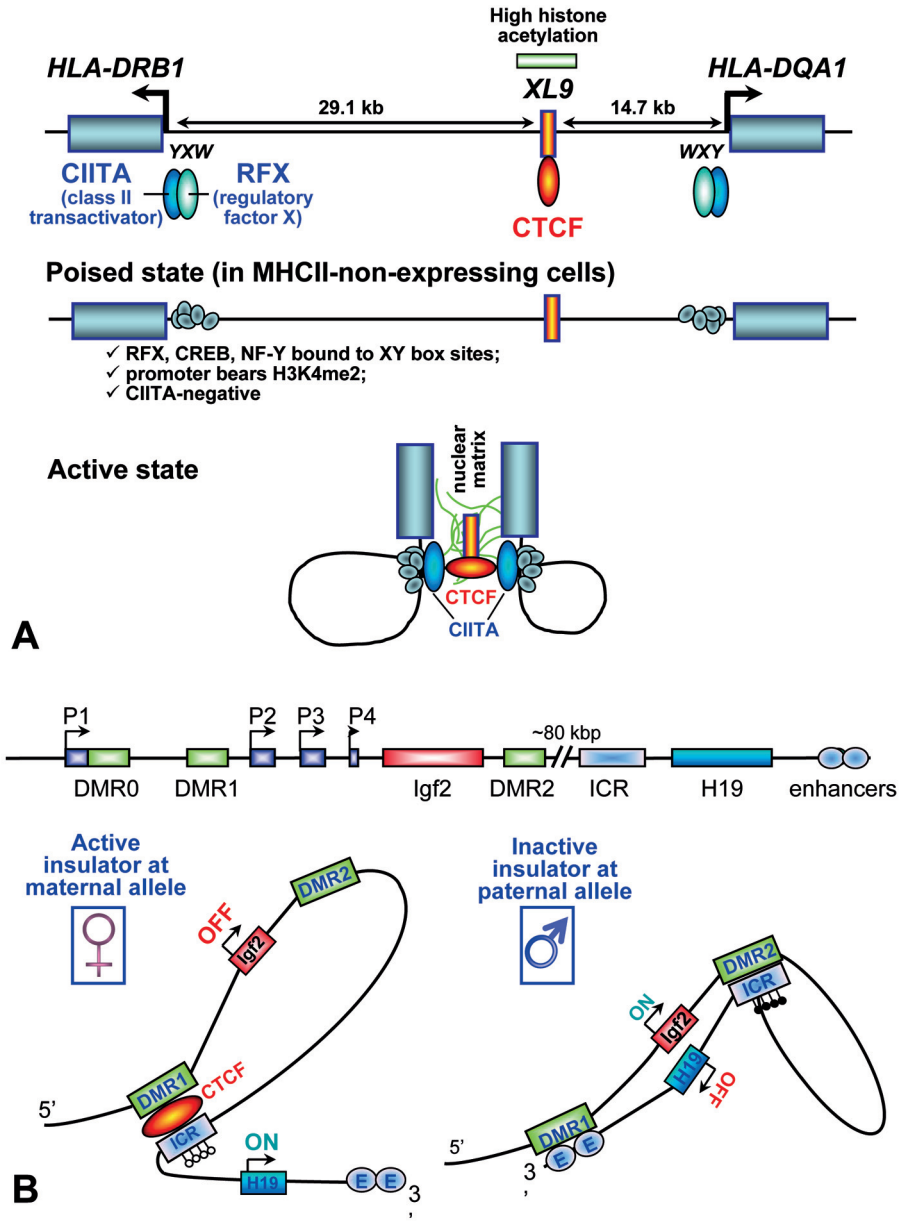

Figure 3. CTCF-mediated intrachromosomal loops. (*A*) Loop formation between two MHCII genes, *HLA-DRB1* and *HLA-DQA1*, and the intergenic CTCF-binding site *XL9* [46]. In MHCII-non-expressing cells, the proximal gene promoters that contain boxes YXW are bound to several TFs, but CIITA, the class II transactivator, is unavailable. This configuration keeps the genes in a poised state. When the genes are activated, CIITA binds the proximal promoters and interacts with CTCF and RFX to form loops that make transcription possible. (*B*) Differential loop formation in the maternal and paternal alleles of the *Igf2/H19* imprinted locus is responsible for the transcriptional status of the two genes in the locus. Top: map of the region, with the *Igf2* and *H19* genes and the differentially methylated DNA regions (DMR0, 1, 2, and the Imprinting Control Region, ICR) involved in the regulation of the locus. The arrows above the blue boxes designate alternative promoters. In the maternal allele the loop is mediated by CTCF binding to the unmethylated ICR and to another differentially methylated region, DMR1, upstream of the *Igf2* gene. This conformation creates an active insulator that precludes the utilization of the two enhancers downstream of the *H19* gene by the *Igf2* promoter: hence the *Igf2* gene is inactive. In the paternal allele, ICR is methylated and cannot bind CTCF. As a result, the enhancer-blocking activity of ICR is lost, and a different special conformation is formed, now allowing *Igf2* expression. Figure in (*B*) based on [37].

and occurs through at least three genetically separable stages: (i) "counting" of the X-chromosome/autosome ratio to ensure the inactivation of only one of the two X-chromosomes; (ii) "choice" of the chromosome to be inactivated, and (iii) actual inactivation process, initiated by coating of the designated inactive chromosome by the non-coding Xist RNA [30, 70, 71]. CTCF has been implicated in the initial pairing of the two X-chromosomes through their X-inactivation centers (see below), in the "choice" decision, as well as in the inactivation process *per se*.

A partial map of the X-inactivation center is presented in Fig. 4A and a brief description is provided in the figure legend. Using hypersensitive site mapping and 3C methodology, Tsai et al. [33] identified two independent loop domains, one between *Xite* and *Tsix*, the other one between *Xist* and another non-coding gene, *Jpx/Enox*, not previously implicated in X-chromosome inactivation (Fig. 4B). The interactions are both developmentally- and sex-specific and help to explain the transcriptional activities of the different non-coding elements, and their mutual interplay during the process of inactivation (for more information, see legend to Fig. 4B). The work of Tsai et al. [33] did not explicitly address the role of CTCF in the recognized dynamics of the conformational transitions in the domain. The authors, however, consider CTCF as the major candidate for mediating the *cis*-interactions. Indeed, they point out that in addition to the CTCF binding sites recognized in the inactivation

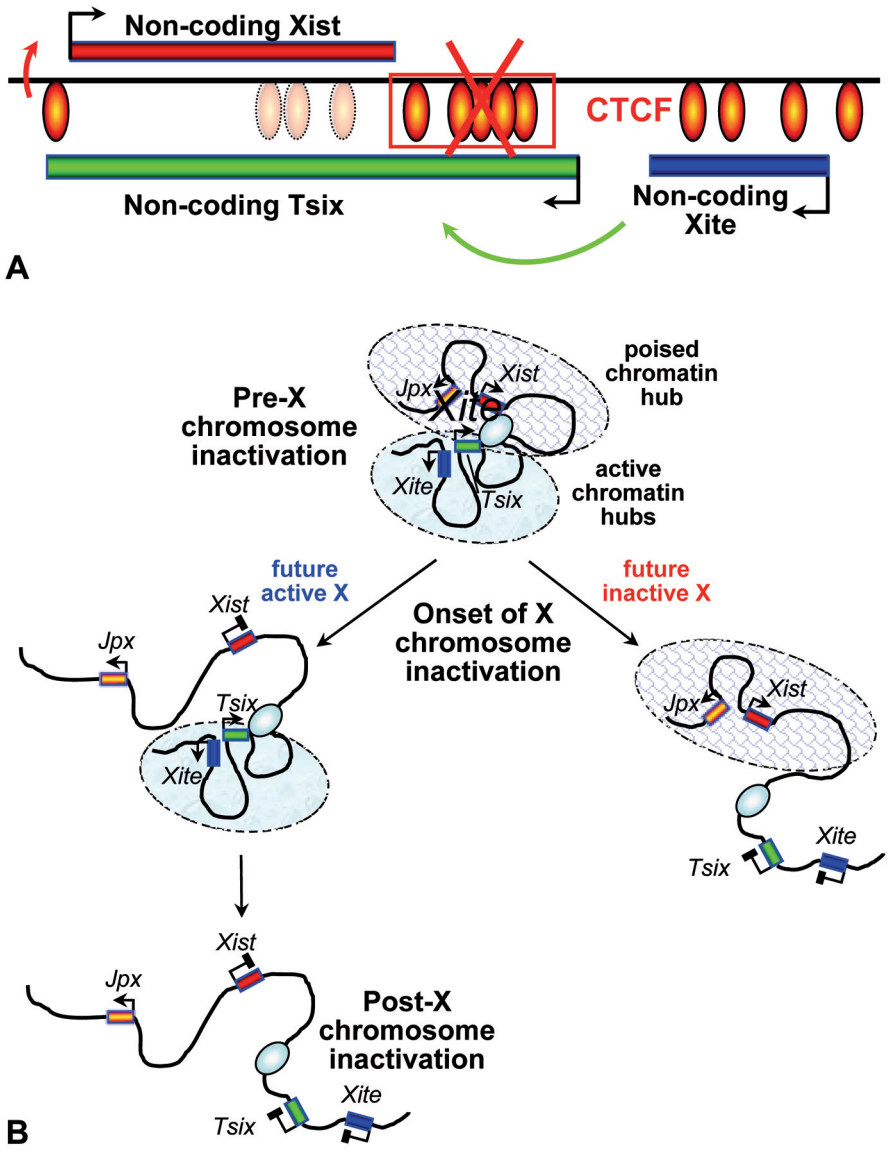

Figure 4. Loop formation at the X-chromosome inactivation center (Xic) controls its activity. (*A*) Map of the three genes that specify the sequence of three non-coding transcripts, Xite, Tsix, and Xist, known to be involved in the function of Xic. Following the choice of which of the two X-chromosomes will be inactivated, Xite activates the antisense transcription of Tsix on the future active chromosome. This in turn blocks the production of Xist transcript (see green and red arrows). On the future inactive X, Xite and Tsix are down-regulated, which leads to induction of Xist transcription and the onset of the actual inactivation process. This binary switch for Xist controls the inactivation process. The map also shows the CTCF binding sites occupied by the protein. (*B*) Dynamic changes in the chromatin fiber conformation at Xic accompany the transitions from pre-inactivation through onset of inactivation to actual inactivated states. In the pre-state, two loops are formed, one between *Xite* and *Tsix*, the other one between *Xist* and a region containing the *Jpx* gene further upstream. The *Xite/Tsix* loop is in an active conformation, whereas the *Xist/Jpx* is in a poised state. At the onset of inactivation, only one of these loops persists, the *Xite/Tsix* loop on the future active X, and the *Xist/Jpx* loop on the future inactive X. Following the establishment of X-inactivation, all loops are lost. Figure based on [31, 33].

center, there is a strong CTCF binding site in the proximal promoter of *Xist* in human and mouse (earlier identified by Pugacheva et al. [72]), exactly in the region of contacts between *Xist* and *Jpx*. In addition, although some of the major CTCF sites (crossed out in the schematic) do not participate in the loop formation, there are other sites close by that can perform this function.

CTCF in the organization of bridges *in trans*

As mentioned in the previous section, CTCF has been implicated in the initial pairing of the two X-chromosomes through their X-inactivation centers [28, 29]. This pairing is CTCF-dependent and may involve additional 'pairing factors', as depicted in Fig. 5. Interestingly, pairing also depends on transcription through an unknown mechanism. In X-chromosome pairing, bridging occurs between two homologous chromosomes; however, there is no conceptual impediment to CTCF-mediated bridge formation involving any two chromosomes. Indeed, 3C and 4C studies (see Table 2) demonstrate that, at least in some cases, the inter-chromosomal interactions are CTCF-dependent [12, 13].

Does CTCF delineate chromatin regions of distinct epigenetic modifications?

Many studies have addressed the issue of whether CTCF binding also creates boundaries between distinct chromatin domains that differ in their DNA

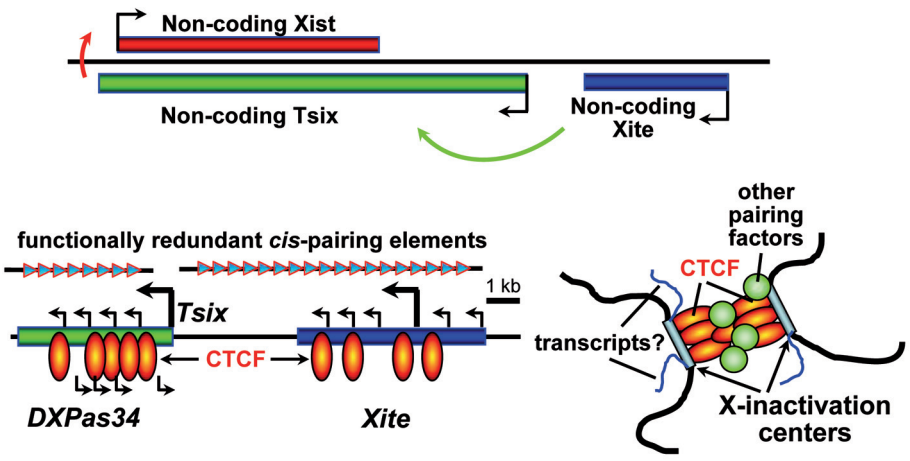

Figure 5. CTCF-mediated X-chromosome pairing (bridging). Top: schematic of a portion of Xic implicated in the homologous pairing between the two X-chromosomes (pairing is a prerequisite for the binary switch in *Xist* expression). Bottom left: blow-up of a portion of the region showing CTCF binding sites over the *Xite* and *Tsix* regions. Bottom right: CTCF-mediated bridge-formation between the inactivation centers of the two homologous X-chromosomes. Figure based predominantly on [29].

methylation and/or posttranslational histone modifications. In some cases this seems to be the case. Thus, for example, two CTCF sites flank the CpG methylation boundary in the promoter region of the breast cancer 1 gene (*BRCA1*) [46]. The *Rb* gene proximal promoter (a CpG island) is unmethylated and enriched in 'active' histone marks, H3ac and H3K4me2; this chromatin environment may be created by the CTCF molecule bound at that site [47]. The four CTCF binding sites at the *c-myc* insulator element (MINE) and in the promoter of the *c-myc* gene also seem to define a region of hyper-acetylation and H3K9 hypo-methylation over the gene region [43]. The single positioned nucleosome flanked by two CTCF binding sites in the 3'-untranslated region of the distrophia myotonica-protein kinase gene (*DMPK*) provides an interesting example of a very local chromatin environment. This gene region contains numerous CTG repeats (up to 100 in healthy individuals) which stably position a nucleosome that contains H3K9me2 and binds HP1γ [48, 49]. Interestingly, CTG repeat expansion in disease leads to loss of CTCF binding and heterochromatization of the entire region. What seems to be certain is that the presence of bound CTCF protects its binding site from DNA methylation, at least at imprinting control regions [38–40].

A careful review of the studies of individual gene systems failed to reveal a clear, recognizable, general pattern of chromatin modifications. The genome-wide study of Barski et al. [61] does, however, recognize such a pattern: when >20,000 CTCF binding sites in the human genome are aligned, several histone modifications (H3K4me1/me2/me3, H3K9me1, and H3K27me1) are all enriched over these sites, albeit to a different degree. Of note, H3K4 methylation is generally considered an 'active' mark, whereas methylation of K9 and K27 on the same histone is linked to gene inactivity. Why 'active' and 'inactive' histone methylation marks peacefully coexist over CTCF binding sites remains to be determined. The genome-wide study [61] has also identified a high level of enrichment of the histone replacement variant H2A.Z [73]. H2A.Z has been earlier identified as a CTCF partner in affinity purification experiments [65]. The functional significance of this partnership is, at present, not clear. It is possible that nucleosomes marked by H2A.Z are recognized by CTCF, as part of a (sequence-independent) mechanism for CTCF recruitment to specific chromatin regions.

Two recent papers may shed some light on how chromatin modification patterns are established. Miles et al. [74] studied the epigenetic profile of the human β-globin locus in transgenic mice and found that the chromatin status tightly correlates with the level of intergenic transcription. An unexpectedly large transcript (initiated at a distance of >250 kb from the locus, in the nearby silent olfactory receptor genes) and several smaller transcripts from within the LCR and the regions directly flanking the active globin genes have been identified. Thus, Miles et al. [74] suggest that "controlling the time of intergenic transcription may be a strategy adopted to modify specified domains". A very similar conclusion – that the transcriptional status, not the ICR (the CTCF binding site), determines the histone modification pattern at the imprinted *H19* locus – has been reached by the Bartolomei laboratory [75]. It remains to be seen whether the role of transcription in defining chromatin modifications will turn to be the general rule rather than the exception. If this would be the case, then clearly CTCF functions in a way that does not involve direct effects on chromatin structure.

Concluding remarks

The recent focus on CTCF, originally identified as a negative transcription factor, comes from the realization that the protein is involved in insulator functions in multiple gene systems. The recently published genome-wide localization studies revealed an unexpectedly large number of CTCF binding sites along the genome. These observations suggest that CTCF may be playing a universal role in topologically organizing the genome as a whole, by its ability to form intra-chromosomal loops and inter-chromosomal bridges. While the topological function of CTCF seems beyond doubt, its relationship to gene transcription and chromatin epigenetic modifications is less clear. In many cases, the CTCF-mediated loop formation is constitutive, and does not correlate with transcription. In other cases, however, CTCF may directly participate in the regulation of transcription, through mechanisms that may not involve loops or bridges (Table 1). Stunningly, CTCF interacts directly with RNA polymerase II (Pol II) [76]; the functional consequences of this interaction may be profound and need to be further investigated.

Acknowledgements. JZ is supported in part by NSF grant 0504239; PC is partially financed by Ministero della Salute and Ministero dell'Istuzione, dell'Università e della Ricerca, Italy.

1 Misteli, T. (2007) Beyond the sequence: cellular organization of genome function. Cell 128, 787–800.
2 de Laat, W. and Grosveld, F. (2007) Inter-chromosomal gene regulation in the mammalian cell nucleus. Curr. Opin. Genet. Dev. 17, 456–464.
3 Caron, H., van Schaik, B., van der Mee, M., Baas, F., Riggins, G., van Sluis, P., Hermus, M.C., van Asperen, R., Boon, K., Voute, P.A., Heisterkamp, S., van Kampen, A. and Versteeg, R. (2001) The human transcriptome map: clustering of highly expressed genes in chromosomal domains. Science 291, 1289–1292.
4 Lercher, M. J., Urrutia, A. O. and Hurst, L. D. (2002) Clustering of housekeeping genes provides a unified model of gene order in the human genome. Nat. Genet. 31, 180–183.
5 Iborra, F. J., Pombo, A., Jackson, D. A. and Cook, P. R. (1996) Active RNA polymerases are localized within discrete transcription "factories' in human nuclei. J. Cell. Sci. 109 (Pt 6), 1427–1436.
6 Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J. A., Lopes, S., Reik, W. and Fraser, P. (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. Nat. Genet. 36, 1065–1071.
7 Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F. and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol. Cell 10, 1453–1465.
8 Palstra, R. J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F. and de Laat, W. (2003) The beta-globin nuclear compartment in development and erythroid differentiation. Nat. Genet. 35, 190–194.
9 Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. (2002) Capturing chromosome conformation. Science 295, 1306–1311.
10 Lomvardas, S., Barnea, G., Pisapia, D. J., Mendelsohn, M., Kirkland, J. and Axel, R. (2006) Interchromosomal interactions and olfactory receptor choice. Cell 126, 403–413.
11 Wurtele, H. and Chartrand, P. (2006) Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. Chromosome Res. 14, 477–495.
12 Ling, J. Q., Li, T., Hu, J. F., Vu, T. H., Chen, H. L., Qiu, X. W., Cherry, A. M. and Hoffman, A. R. (2006) CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. Science 312, 269–272.
13 Zhao, Z., Tavoosidana, G., Sjolinder, M., Gondor, A., Mariano, P., Wang, S., Kanduri, C., Lezcano, M., Sandhu, K.S., Singh, U., Pant, V., Tiwari, V., Kurukuti, S. and Ohlsson, R. (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. Nat. Genet. 38, 1341–1347.
14 Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B. and de Laat, W. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). Nat. Genet. 38, 1348–1354.
15 de Laat, W. (2007) Long-range DNA contacts: romance in the nucleus? Curr. Opin. Cell. Biol. 19, 317–320.
16 Misteli, T. (2001) The concept of self-organization in cellular architecture. J. Cell Biol. 155, 181–185.
17 Bolzer, A., Kreth, G., Solovei, I., Koehler, D., Saracoglu, K., Fauth, C., Muller, S., Eils, R., Cremer, C., Speicher, M. R. and Cremer, T. (2005) Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol. 3, e157.
18 Cremer, T. and Cremer, C. (2001) Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat. Rev. Genet. 2, 292–301.
19 Roix, J. J., McQueen, P. G., Munson, P. J., Parada, L. A. and Misteli, T. (2003) Spatial proximity of translocation-prone gene loci in human lymphomas. Nat. Genet. 34, 287–291.
20 Scheuermann, M. O., Tajbakhsh, J., Kurz, A., Saracoglu, K., Eils, R. and Lichter, P. (2004) Topology of genes and non-transcribed sequences in human interphase nuclei. Exp. Cell Res. 301, 266–79.
21 Kupper, K., Kolbl, A., Biener, D., Dittrich, S., von Hase, J., Thormeyer, T., Fiegler, H., Carter, N. P., Speicher, M. R., Cremer, T. and Cremer, M. (2007) Radial chromatin positioning is shaped by local gene density, not by gene expression. Chromosoma 116, 285–306.
22 Federico, C., Cantarella, C. D., Di Mare, P., Tosi, S. and Saccone, S. (2008) The radial arrangement of the human chromosome 7 in the lymphocyte cell nucleus is associated with chromosomal band gene density. Chromosoma 117, 399–410.
23 Volpi, E. V., Chevret, E., Jones, T., Vatcheva, R., Williamson, J., Beck, S., Campbell, R. D., Goldsworthy, M., Powis, S. H., Ragoussis, J., Trowsdale, J. and Sheer, D. (2000) Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. J. Cell Sci. 113 (Pt 9), 1565–76.
24 Marenduzzo, D., Faro-Trindade, I. and Cook, P. R. (2007) What are the molecular ties that maintain genomic loops? Trends Genet. 23, 126–133.
25 Chung, J. H., Whiteley, M. and Felsenfeld, G. (1993) A 5' element of the chicken beta-globin domain serves as an insulator in human erythroid cells and protects against position effect in Drosophila. Cell 74, 505–514.
26 Bell, A. C., West, A. G. and Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell 98, 387–396.
27 Splinter, E., Heath, H., Kooren, J., Palstra, R. J., Klous, P., Grosveld, F., Galjart, N. and de Laat, W. (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev. 20, 2349–2354.

28 Bacher, C. P., Guggiari, M., Brors, B., Augui, S., Clerc, P., Avner, P., Eils, R. and Heard, E. (2006) Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. Nat. Cell. Biol. 8, 293–299.

29 Xu, N., Donohoe, M. E., Silva, S. S. and Lee, J. T. (2007) Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. Nat. Genet. 39, 1390–1396.

30 Erwin, J. A. and Lee, J. T. (2008) New twists in X-chromosome inactivation. Curr. Opin. Cell Biol. 20, 349–355.

31 Donohoe, M. E., Zhang, L. F., Xu, N., Shi, Y. and Lee, J. T. (2007) Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. Mol. Cell 25, 43–56.

32 Filippova, G. N., Cheng, M. K., Moore, J. M., Truong, J. P., Hu, Y. J., Nguyen, D. K., Tsuchiya, K. D. and Disteche, C. M. (2005) Boundaries between chromosomal domains of X inactivation and escape bind CTCF and lack CpG methylation during early development. Dev. Cell 8, 31–42.

33 Tsai, C. L., Rowntree, R. K., Cohen, D. E. and Lee, J. T. (2008) Higher order chromatin structure at the X-inactivation center via looping DNA. Dev. Biol. 319, 416–425.

34 Bell, A. C. and Felsenfeld, G. (2000) Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. Nature 405, 482–485.

35 Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M. and Tilghman, S. M. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. Nature 405, 486–489.

36 Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C. F., Wolffe, A., Ohlsson, R. and Lobanenkov, V. V. (2000) Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. Curr. Biol. 10, 853–856.

37 Murrell, A., Heeson, S. and Reik, W. (2004). Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. Nat. Genet. 36, 889–893.

38 Pant, V., Mariano, P., Kanduri, C., Mattsson, A., Lobanenkov, V., Heuchel, R. and Ohlsson, R. (2003) The nucleotides responsible for the direct physical contact between the chromatin insulator protein CTCF and the H19 imprinting control region manifest parent of origin-specific long-distance insulation and methylation-free domains. Genes Dev. 17, 586–590.

39 Schoenherr, C. J., Levorse, J. M. and Tilghman, S. M. (2003) CTCF maintains differential methylation at the Igf2/H19 locus. Nat. Genet. 33, 66–69.

40 Szabo, P. E., Tang, S. H., Silva, F. J., Tsark, W. M. and Mann, J. R. (2004) Role of CTCF binding sites in the Igf2/H19 imprinting control region. Mol. Cell. Biol. 24, 4791–4800.

41 Bergstrom, R., Whitehead, J., Kurukuti, S. and Ohlsson, R. (2007) CTCF regulates asynchronous replication of the imprinted H19/Igf2 domain. Cell Cycle 6, 450–454.

42 Klenova, E. M., Nicolas, R. H., Paterson, H. F., Carne, A. F., Heath, C. M., Goodwin, G. H., Neiman, P. E. and Lobanenkov, V. V. (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, 7612–7624.

43 Gombert, W. M., Farris, S. D., Rubio, E. D., Morey-Rosler, K. M., Schubach, W. H. and Krumm, A. (2003) The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain. Mol. Cell. Biol. 23, 9338–9348.

44 Vostrov, A. A. and Quitschke, W. W. (1997) The zinc finger protein CTCF binds to the APBbeta domain of the amyloid beta-protein precursor promoter. Evidence for a role in transcriptional activation. J. Biol. Chem. 272, 33353–33359.

45 Lutz, M., Burke, L. J., LeFevre, P., Myers, F. A., Thorne, A. W., Crane-Robinson, C., Bonifer, C., Filippova, G. N., Lobanenkov, V. and Renkawitz, R. (2003) Thyroid hormone-regulated enhancer blocking: cooperation of CTCF and thyroid hormone receptor. Embo J. 22, 1579–1587.

46 Butcher, D. T., Mancini-DiNardo, D. N., Archer, T. K. and Rodenhiser, D. I. (2004) DNA binding sites for putative methylation boundaries in the unmethylated region of the BRCA1 promoter. Int. J. Cancer 111, 669–678.

47 De La Rosa-Velazquez, I. A., Rincon-Arano, H., Benitez-Bribiesca, L. and Recillas-Targa, F. (2007) Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. Cancer Res. 67, 2577–2585.

48 Filippova, G. N., Thienes, C. P., Penn, B. H., Cho, D. H., Hu, Y. J., Moore, J. M., Klesert, T. R., Lobanenkov, V. V. and Tapscott, S. J. (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. Nat. Genet. 28, 335–343.

49 Cho, D. H., Thienes, C. P., Mahoney, S. E., Analau, E., Filippova, G. N. and Tapscott, S. J. (2005) Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. Mol. Cell 20, 483–489.

50 Majumder, P., Gomez, J. A. and Boss, J. M. (2006) The human major histocompatibility complex class II HLA-DRB1 and HLA-DQA1 genes are separated by a CTCF-binding enhancer-blocking element. J. Biol. Chem. 281, 18435–18443.

51 Majumder, P., Gomez, J. A., Chadwick, B. P. and Boss, J. M. (2008) The insulator factor CTCF controls MHC class II gene expression and is required for the formation of long-distance chromatin interactions. J. Exp. Med. 205, 785–798.

52 Guastafierro, T., Cecchinelli, B., Zampieri, M., Reale, A., Riggio, G., Sthandier, O., Zupi, G., Calabrese, L. and Caiafa, P. (2008) CCCTC-binding factor activates PARP-1 affecting DNA methylation machinery. J. Biol. Chem. 283, 21873–21880.

53 Ohlsson, R., Renkawitz, R. and Lobanenkov, V. (2001) CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet. 17, 520–527.

54 Yu, W., Ginjala, V., Pant, V., Chernukhin, I., Whitehead, J., Docquier, F., Farrar, D., Tavoosidana, G., Mukhopadhyay, R., Kanduri, C., Oshimura, M., Feinberg, A.P., Lobanenkov, V., Klenova, E. and Ohlsson, R. (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. Nat. Genet. 36, 1105–1110.

55 Klenova, E.M., Chernukhin, I.V., El-Kady, A., Lee, R.E., Pugacheva, E.M., Loukinov, D.I., Goodwin, G.H., Delgado, D., Filippova, G.N., Leon, J., Morse, H.C., 3rd, Neiman, P.E. and Lobanenkov, V.V. (2001) Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF. Mol. Cell. Biol. 21, 2221–2234.

56 Dunn, K. L., Zhao, H. and Davie, J. R. (2003) The insulator binding protein CTCF associates with the nuclear matrix. Exp. Cell. Res. 288, 218–223.

57 Bell, A. C., West, A. G. and Felsenfeld, G. (2001) Insulators and boundaries: versatile regulatory elements in the eukaryotic. Science 291, 447–450.

58 Gaszner, M. and Felsenfeld, G. (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. Nat. Rev. Genet. 7, 703–713.

59 Wallace, J. A. and Felsenfeld, G. (2007) We gather together: insulators and genome organization. Curr. Opin. Genet. Dev. 17, 400–407.

60 Kim, T. H., Abdullaev, Z. K., Smith, A. D., Ching, K. A., Loukinov, D. I., Green, R. D., Zhang, M. Q., Lobanenkov, V. V. and Ren, B. (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. Cell 128, 1231–1245.

61 Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. Cell 129, 823–837.

62 Xie, X., Mikkelsen, T. S., Gnirke, A., Lindblad-Toh, K., Kellis, M. and Lander, E. S. (2007) Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites. Proc. Natl. Acad. Sci. U S A 104, 7145–7150.

63 Mukhopadhyay, R., Yu, W., Whitehead, J., Xu, J., Lezcano, M., Pack, S., Kanduri, C., Kanduri, M., Ginjala, V., Vostrov, A., Quitschke, W., Chernukhin, I., Klenova, E., Lobanenkov, V. and Ohlsson, R. (2004) The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. Genome Res. 14, 1594–1602.
64 Guelen, L., Pagie, L., Brasset, E., Meuleman, W., Faza, M. B., Talhout, W., Eussen, B. H., de Klein, A., Wessels, L., de Laat, W. and van Steensel, B. (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. Nature 453, 948–951.
65 Yusufzai, T. M., Tagami, H., Nakatani, Y. and Felsenfeld, G. (2004) CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. Mol. Cell 13, 291–298.
66 Torrano, V., Navascues, J., Docquier, F., Zhang, R., Burke, L.J., Chernukhin, I., Farrar, D., Leon, J., Berciano, M.T., Renkawitz, R., Klenova, E., Lafarga, M. and Delgado, M.D. (2006) Targeting of CTCF to the nucleolus inhibits nucleolar transcription through a poly(ADP-ribosyl)ation-dependent mechanism. J. Cell. Sci. 119, 1746–1759.
67 Palstra, R. J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B. and de Laat, W. (2008) Maintenance of long-range DNA interactions after inhibition of ongoing RNA polymerase II transcription. PLoS ONE 3, e1661.
68 Bender, M. A., Byron, R., Ragoczy, T., Telling, A., Bulger, M. and Groudine, M. (2006) Flanking HS-62.5 and 3' HS1, and regions upstream of the LCR, are not required for beta-globin transcription. Blood 108, 1395–1401.
69 Noordermeer, D., Branco, M. R., Splinter, E., Klous, P., van Ijcken, W., Swagemakers, S., Koutsourakis, M., van der Spek, P., Pombo, A. and de Laat, W. (2008) Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. PLoS Genet. 4, e1000016.
70 Avner, P. and Heard, E. (2001) X-chromosome inactivation: counting, choice and initiation. Nat. Rev. Genet. 2, 59–67.
71 Clerc, P. and Avner, P. (2006) Random X-chromosome inactivation: skewing lessons for mice and men. Curr. Opin. Genet. Dev. 16, 246–253.
72 Pugacheva, E. M., Tiwari, V. K., Abdullaev, Z., Vostrov, A. A., Flanagan, P. T., Quitschke, W. W., Loukinov, D. I., Ohlsson, R. and Lobanenkov, V. V. (2005) Familial cases of point mutations in the XIST promoter reveal a correlation between CTCF binding and pre-emptive choices of X chromosome inactivation. Hum. Mol. Genet. 14, 953–965.
73 Zlatanova, J. and Thakar, A. (2008) H2A.Z: view from the top. Structure 16, 166–179.
74 Miles, J., Mitchell, J. A., Chakalova, L., Goyenechea, B., Osborne, C. S., O'Neill, L., Tanimoto, K., Engel, J. D. and Fraser, P. (2007) Intergenic transcription, cell-cycle and the developmentally regulated epigenetic profile of the human beta-globin locus. PLoS ONE 2, e630.
75 Verona, R. I., Thorvaldsen, J. L., Reese, K. J. and Bartolomei, M. S. (2008) The transcriptional status but not the imprinting control region determines allele-specific histone modifications at the imprinted H19 locus. Mol. Cell. Biol. 28, 71–82.
76 Chernukhin, I., Shamsuddin, S., Kang, S.Y., Bergstrom, R., Kwon, Y.W., Yu, W., Whitehead, J., Mukhopadhyay, R., Docquier, F., Farrar, D., Morrison, I., Vigneron, M., Wu, S.Y., Chiang, C.M., Loukinov, D., Lobanenkov, V., Ohlsson, R. and Klenova, E. (2007) CTCF interacts with and recruits the largest subunit of RNA polymerase II to CTCF target sites genome-wide. Mol. Cell. Biol. 27, 1631–1648.