27

Chromatin Insulators and Epigenetic Inheritance in Health and Disease

Jingping Yang and Victor G. Corces

Department of Biology, Emory University, Atlanta, Georgia, USA

27.1 Introduction

The genome of eukaryotic organisms is compacted in the nucleus in a manner that allows transmission of the genetic material between generations and access of the replication and transcription machinery to the DNA. Expression or repression of specific genes is accompanied by alterations in the local chromatin carried out by DNA methyl transferases, histone modifying enzymes, ATP-dependent chromatin remodeling complexes, and histone chaperones that replace canonical histone for specific variants. Some of these changes in chromatin structure can be transmitted from mother to daughter cells or from one organismal generation to the next. As a consequence, alterations in the structure of the 10 nm chromatin fiber can carry epigenetic information encoding a memory of specific transcriptional states. Less understood is whether this chromatin fiber is arranged into higher-order levels of organization and, if so, whether they contribute to the establishment of patterns of gene expression that can be epigenetically inherited. Results obtained in the last few years suggest an affirmative answer to these two questions.

Chromatin insulators are DNA-bound protein complexes that can mediate intra- and inter-chromosome interactions. In doing so, they can bring close together regulatory sequences located at large distances from each other. The effect on the expression of adjacent genes depends on the nature of the sequences brought into proximity, resulting in activation or repression of single genes or large chromosomal domains. In the process, insulators can then alter the status of the chromatin in these genes or domains and, as a consequence, modify patterns of epigenetic inheritance. It is now becoming apparent that insulator-mediated interactions contribute to the establishment of a three-dimensional organization of the DNA in the nucleus. Since this structure is directly related to gene expression, it follows that the organization of the genome within the nucleus is in part a determinant and in part a consequence of the transcriptional status of a

Toxicology and Epigenetics, First Edition. Edited by Saura C. Sahu.

cell. This organization may be cell-type specific and correlate with patterns of gene expression necessary for cell differentiation during development. Because this organization is both a cause and an effect of all other epigenetic modifications, it may represent a fingerprint of the epigenetic state of a specific cell. In other words, knowledge of the three-dimensional organization of the DNA of a specific cell may give us an understanding of its epigenetic potential.

Insulators are emerging as the main players involved in the establishment and/or maintenance of the three-dimensional organization of the genome. Here we review the structure and properties of these DNA/multi-protein complexes, their role in gene expression, and the mechanisms by which misregulation of their function leads to various human diseases.

27.2 Structure and organization of insulators

Insulators are DNA–protein complexes originally defined by their ability to block enhancer promoter interactions and/or to serve as barriers against the spreading of the silencing effects of heterochromatin. These properties are probably a subset of the large array of functions insulators can play in the genome as a consequence of their ability to mediate inter- and intra-chromosomal interactions (Phillips and Corces, 2009). Insulators were originally discovered in *Drosophila* and we will first discuss the characteristics of these regulatory elements in this organism.

There are several types of insulators in *Drosophila* that have been studied in detail, including the *scs* and *scs'* flanking the heat shock *hsp70* locus (Kellum and Schedl, 1992; Zhao *et al*., 1995), the *gypsy* insulator first found in the *gypsy* retrovirus (Geyer and Corces, 1992; Hoover *et al*., 1993), the *Fab-7*, *Fab-8*, and *Mcp* insulators located in the Bithorax Complex (BX-C) where they orchestrate the complex spatio-temporal expression of the three genes present in the locus (Gyurkovics *et al*., 1990; Karch *et al*., 1994), and the *SF1* insulator described in the Antennapedia Complex (ANT-C) (Belozerov *et al*., 2003). Each of these insulators consists of a DNA sequence and a specific DNA binding protein that interacts with this sequence (Figure 27.1a). In the case of the *scs* insulator, the DNA binding protein is Zeste-White 5 (ZW5), which is a zinc finger protein required for cell viability. Null mutations in the *zw5* gene are recessive lethal, but hypomorphic alleles display a variety of pleiotropic effects on wing, bristle, and eye development (Gaszner *et al*., 1999). The *scs'* sequences interact with a protein called Boundary Element Associated Factor 32 (BEAF-32) (Zhao *et al*., 1995). The C terminal region of the protein is involved in protein–protein interactions. Expression of a dominant negative form of BEAF-32 results in changes in chromosome structure and cell viability (Roy *et al*., 2007). The *gypsy* insulator (to which we will refer to as the Su(Hw) insulator) contains binding sites for the Suppressor of Hairy-wing (Su(Hw)) protein, which is a 12 zinc finger DNA binding protein. Mutations in the *su(Hw)* gene cause female sterility but do not result in lethality (Harrison *et al*., 1993). *Fab-8* and *Mcp* insulators harbor a sequence bound by the *Drosophila* homolog of mammalian CTCF (CCCTC binding factor) (dCTCF), which has 11 zinc fingers (Holohan *et al*., 2007). Mutations in *dCTCF* are lethal and show abdominal homeotic phenotypes (Gerasimova *et al*., 2007; Mohan *et al*., 2007). *Fab-7* may represent another class of insulators that use the GAGA factor (GAF) as a DNA binding protein, which also contains a BTB (for BR-C, ttk and bab) domain. Mutations in the *trl* gene, which encodes GAF, affect *Fab-7* insulator activity (Schweinsberg *et al*., 2004). In addition, GAF is present and required for the function of the *SF1* insulator found in the ANT-C (Belozerov *et al*., 2003) (Figure 27.1b).

Vertebrates appear to rely mostly on the widespread CTCF insulator (Figure 27.1a) (Wallace and Felsenfeld, 2007). CTCF is a highly conserved protein containing an 11 zinc finger central DNA binding domain, displaying close to 100% homology between mouse, chicken, and human, embedded within slightly more divergent N- and C-termini (Ohlsson *et al*., 2001). On the basis of its ability to bind to a wide range of

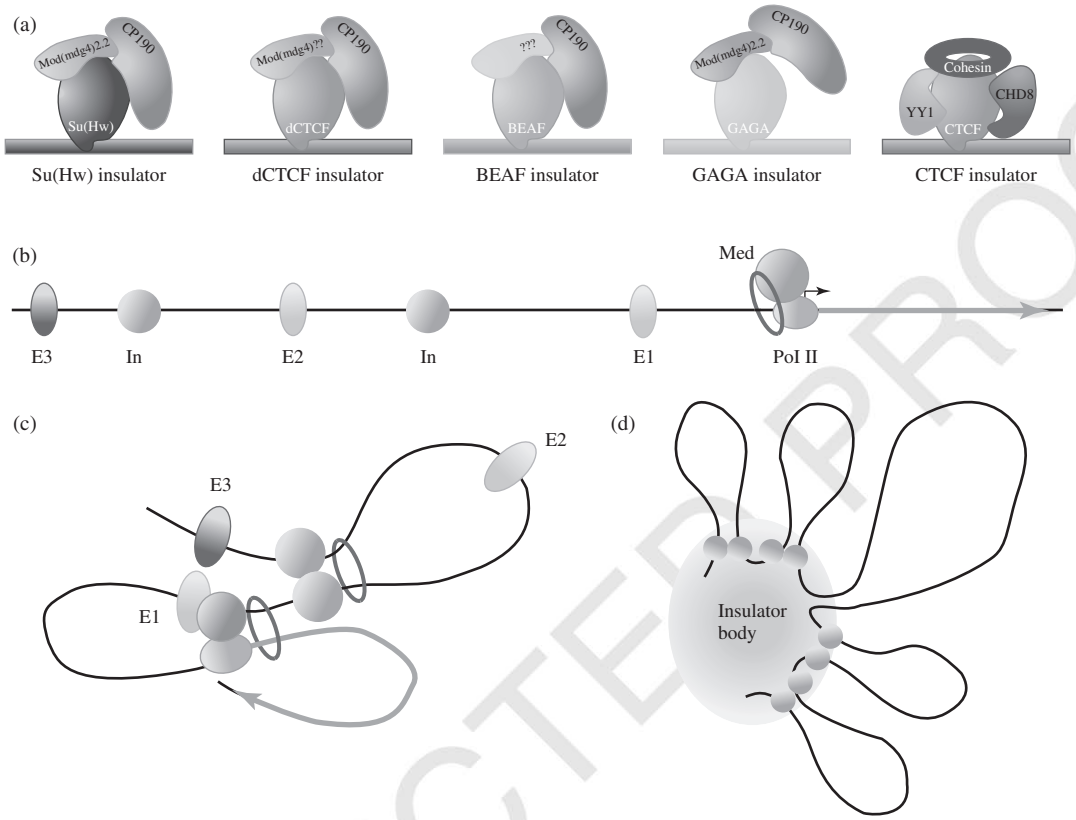

Figure 27.1 *(a) Diagram showing the structure of different* Drosophila *and vertebrate insulators. Each* Drosophila *insulator contains a different binding protein that may define its specific function. All insulators share the common protein CP190, although the role of this protein in the function of the GAGA insulator has not been demonstrated experimentally. In addition, all subclasses may also have one Mod(mdg4) isoform. The structure of the vertebrate CTCF insulator is indicated on the left. (b) Linear organization of a typical eukaryotic gene. The RNA coding region of the gene is represented by an arrow, the transcription complex and RNA polymerase II is shown by an oval, and the Mediator complex by a sphere. The cohesin complex is indicated by a ring; cohesin is also found at some enhancers (not shown). Enhancers in the upstream regulatory region of the gene are indicated by ovals of different shades. Insulators are represented by spheres. (c) Three-dimensional arrangement of the same region represented in (b). The most proximal enhancer (E1) contacts mediator and/or the transcription complex; cohesin stabilizes this interaction. Insulator elements, such as CTCF in vertebrates, contact each other to form a loop; this interaction is also mediated by cohesin. As a consequence of the formation of this loop, Enhancer E2 is unable to act on the promoter of the gene while enhancer E3 is brought close to the promoter to activate transcription. (d) Many insulator sites come together at one nuclear location to form insulator bodies. This arrangement is similar to that formed by Polycomb Response Elements (PREs) and Polycomb-Group (PcG) proteins, which come together at Pc bodies. (e) Organization of the* Drosophila *Bithorax Complex (BX-C), showing the linear arrangement of the three genes in the locus, which are indicated by arrows; ovals represent the transcription complex at the promoter of each gene whereas spheres represent PREs and associated proteins. (f) Interactions among the PREs and promoters of the genes result in a specific three-dimensional arrangement of the locus that results in repression of transcription. (g) Multiple Hox gene loci can be co-repressed and associate at nuclear locations termed Pc bodies*

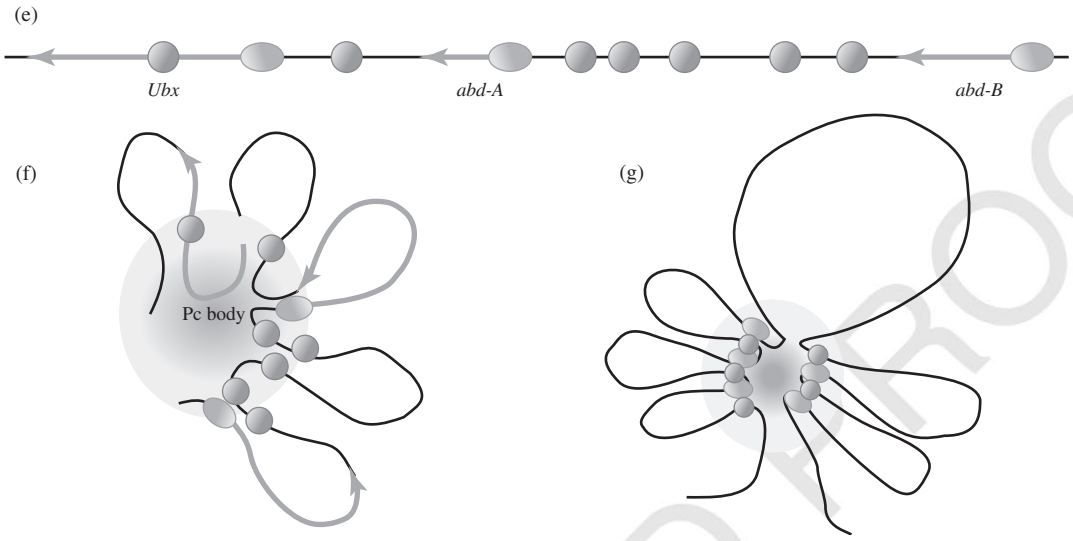

Figure 27.1 *(continued)*

variant sequences as well as specific co-regulatory proteins through combinatorial use of different zinc fingers, CTCF was originally described as a multifunctional factor (Filippova *et al*., 1996). This feature may explain the ability of vertebrate cells to fulfill all insulator functions with just one insulator whereas *Drosophila* cells require multiple insulators. CTCF homozygous knockout mice exhibit early embryonic lethality prior to implantation, highlighting the critical importance of CTCF in diverse cellular processes (Heath *et al*., 2008).

In addition to the DNA binding proteins, insulator activity also requires other components that interact with the DNA binding proteins and among themselves to form a large multi-protein complex (Figure 27.1a). In the Su(Hw) insulator, Su(Hw) interacts with Mod(mdg4)2.2 and CP190 (Gause *et al*., 2001; Ghosh *et al*., 2001). Mod(mdg4)2.2 does not bind to DNA directly but interacts with Su (Hw) through its carboxy-terminal domain. Mod(mdg4)2.2 also contains a BTB domain in the N-terminal region that mediates homo and hetero-multimerization with other insulator components. The *mod(mdg4)* gene encodes approximately 29 different isoforms that arise by alternative *cis*- and *trans*-splicing (Buchner *et al*., 2000; Labrador and Corces, 2003; Labrador *et al*., 2001). The different isoforms may interact with the DNA binding proteins of various *Drosophila* insulators. Null mutations in the gene result in lethality, but mutations affecting the Mod(mdg4)2.2 isoform are viable and show defects in Su(Hw) insulator function (Gerasimova *et al*., 1995). CP190 also contains a BTB domain as well as three zinc fingers and it interacts with both Su(Hw) and Mod(mdg4)2.2. CP190 binds DNA with low affinity and specificity but it does not interact directly with insulator sequences present in the *gypsy* retrovirus, where it is recruited through interactions with Su(Hw) and Mod(mdg4)2.2 instead. Mutations in the *CP190* gene are lethal (Pai *et al*., 2004).

In *Drosophila*, CP190 is a shared component of the different types of insulators (Figure 27.1a). It also interacts with dCTCF (Gerasimova *et al*., 2007; Mohan *et al*., 2007). Genome-wide mapping indicates colocalization of CP190 with dCTCF and BEAF-32 (Bartkuhn *et al*., 2009; Bushey *et al*., 2009). On the other hand, GAF does not appear to interact directly with CP190 but has been shown to interact with Su(Hw) and Mod(mdg4)2.2 (Melnikova *et al*., 2004). Since these two proteins can in turn interact with CP190, GAF insulators may act mechanistically like the other three types.

Interestingly, the Mod(mdg4) and CP190 proteins have not been conserved in vertebrates. Instead, the vertebrate CTCF protein interacts with cohesins, which have recently emerged as critical partners of CTCF in mediating chromosomal interactions (Figure 27.1a). Cohesins form a ring-shaped complex that holds chromatids together between the S and M phases of the cell cycle. In vertebrates, cohesins are present at CTCF target sequences in the genome and this localization depends on CTCF (Parelho *et al*., 2008; Rubio *et al*., 2008; Stedman *et al*., 2008; Wendt *et al*., 2008). These observations have led to the proposal that the ring-like structure of the cohesin complex can mediate interactions between distant genomic sites by a mechanism similar to that employed to maintain sister chromatids together. It has been recently suggested that CTCF recruits cohesins through interaction with the SA2 subunit of the cohesin complex (Xiao *et al*., 2011).

CTCF has been found to interact with other nuclear factors in addition to cohesin (Figure 27.1a). CTCF interacts in vivo with chromodomain helicase (CHD8), which is a member of the CHD family and has an SNF2-like (sucrose nonfermentable 2-like) helicase/ATPase domain, at the *H19/Igf2* Imprinting Control Region (ICR), the promoter regions of the *BRCA1* (breast cancer 1 susceptibility protein) and *c-myc* genes, and the 5′HS5 insulator of the *β-globin* locus in human cells (Ishihara *et al*., 2006). Like CTCF, CHD8 is also required for the insulator activity of the *H19/Igf2* ICR. Loss of CHD8 leads to expression of *Igf2* from the maternal and paternal alleles (Ishihara *et al*., 2006). CHD8 together with CTCF can affect epigenetic aspects of chromatin structure such as CpG DNA methylation and histone acetylation. Knockdown of CHD8 leads to hypermethylation of CpGs at CTCF binding sites in the promoter regions of the *BRCA1* and *c-myc* genes (Ishihara *et al*., 2006). CTCF can also interact *in vitro* with SIN3A, which recruits a histone deacetylase activity necessary for the silencing function of CTCF (Lutz *et al*., 2000); with transcription factors such as YY1, which is a required cofactor for CTCF in processes such as X chromosome inactivation (Donohoe *et al*., 2007); with YB1 in vitro and in vivo as a co-repressor at the *myc* promoter (Chernukhin *et al*., 2000); with Kaiso, which is required for activity of the chicken *HS4* insulator (Defossez *et al*., 2005); with Oct4 to control X-chromosome pairing during X-chromosome inactivation (Donohoe *et al*., 2009); and with thyroid hormone receptor (TR) response elements in the mouse *c-myc* and the human *APP* genes (Weth *et al*., 2010). Intriguingly, CTCF has been also found to interact with RNA polymerase II. The largest subunit of Pol II associates with CTCF target sites in vivo and this interaction is dependent on the presence of intact CTCF binding sequences. A single CTCF binding site in the *c-myc* gene, but not a mutant site deficient for CTCF binding, is sufficient to activate transcription from a promoterless reporter gene in stably transfected cells (Chernukhin *et al*., 2007). Although these results can be interpreted as CTCF recruiting Pol II to activate transcription, an alternative interpretation is that the association of CTCF with Pol II is a consequence of its interaction with cohesin and the Mediator complex (Kagey *et al*., 2011).

The extent of the association between CTCF and these different co-factors is not clear. It is possible that CTCF has context-dependent functions mediated by different protein partners. In *Drosophila*, the different insulators use unique DNA binding proteins for presumably different functions, but they share CP190, which may mediate insulator interactions through multimerization of this protein. In vertebrates, the most common insulator contains CTCF and cohesin, which mediate interactions between specific sites, while possibly using different cofactors for the various functions CTCF plays in various nuclear processes.

27.3 Insulators and chromatin architecture

27.3.1 Genome-wide distribution of insulator proteins

Chromatin immunoprecipitation followed by microarray hybridization (ChIP-chip) or next generation sequencing (ChIP-seq) indicates that Su(Hw), BEAF-32, and dCTCF bind thousands of sites through

the *Drosophila* genome (Bushey *et al*., 2009; Holohan *et al*., 2007; Jiang *et al*., 2009; Nègre *et al*., 2010). Many, but not all of these sites also contain the CP190 protein. The different insulators, represented by their individual DNA binding proteins, show distinct localization patterns relative to gene features, suggesting that they play separate functional roles in nuclear biology. CTCF and BEAF-32 show a distribution that is highly skewed toward the 5′ end of genes and is enriched in the first 200 bp just upstream of the transcription start site (TSS), while most Su(Hw)-binding sites are found further than 1 kb away. Most dCTCF and BEAF-32 sites −83 and 89%, respectively – are present at the 5′ end of genes that are highly expressed, whereas Su(Hw) associates with genes expressed at low levels. Genes containing dCTCF in the 200 bp region upstream of their TSS are mostly involved in development, whereas genes containing BEAF-32 in this region are mostly involved in metabolic processes. Both dCTCF and BEAF-32 are enriched near or at genes involved in cell cycle control, whereas Su(Hw)-containing genes show little significant clustering based on biological process (Bushey *et al*., 2009; Emberly *et al*., 2008; Jiang *et al*., 2009). Thus different insulator proteins appear to associate with genes involved in different cellular processes. In addition, dCTCF is enriched between annotated cis-regulatory elements and their nearest non-target promoter while BEAF-32 and Su(Hw) are significantly depleted (Nègre *et al*., 2010). Su(Hw) displays higher enrichment at the boundaries of chromosome rearrangement break points than the other proteins. Therefore, the specific association of these insulators with various genomic landmarks indicates their division of labor. Su(Hw) may demarcate independent functional domains and, within these domains, dCTCF may facilitate the targeting of regulatory elements to the correct genes.

Since *Drosophila* has several insulators showing distinct localization patterns with respect to genomic features, it is interesting to contrast this information with the genome wide localization of CTCF in vertebrates. The distribution of CTCF in mouse embryonic stem cells and in various human cell lines has been studied using ChIP-chip and ChIP-seq approaches (Barski *et al*., 2007; Chen *et al*., 2008; Cuddapah *et al*., 2009; Jothi *et al*., 2008; Kim *et al*., 2007). The number of sites uncovered varies depending on the approach, from 13 800 in IMR90 human fibroblasts using ChIP-chip (Kim *et al*., 2007) to 39 600 in mouse ES (Embryonic Stem) cells using ChIP-seq (Chen *et al*., 2008). In general, the distribution of CTCF sites on each chromosome correlates with gene density. However, different from transcription factors, CTCF binding sites are generally located an average of 48 kb away from promoters. Approximately 46% of the sites are located in intergenic regions, consistent with the classical role of insulators in establishing chromatin domains, whereas 20% of the sites display promoter proximal localization. A significant number of the sites fall within genes, with 22% in introns and 12% in exons in human fibroblast cells (Kim *et al*., 2007). Consistent with the classic role of insulators, CTCF depleted domains, which exhibit lower-than-average CTCF density, tend to include clusters of related gene families and genes that are transcriptionally coregulated. On the other hand, CTCF-enriched domains, which have higher than average CTCF binding, often have multiple alternative promoters (Kim *et al*., 2007). In addition, some CTCF sites are located at the transition of chromatin domains with different epigenetic status. For example, there are 793 CTCF sites flanking H3K27me3 domains in CD4+ T cells and this occupancy may be cell type specific (Cuddapah *et al*., 2009). CTCF can also bind 5–10 kb outside lamina-associated domains (LADs), which contain low gene densities and low expressing genes, representing a repressive chromatin environment; out of 1344 LADs found, 333 have CTCF binding at least at one side of the border (Guelen *et al*., 2008). Thus the distribution of CTCF agrees with its role in the establishment and/or maintenance of functional transcriptional domains.

The genomic distribution of insulators suggests that they can play a role in separating chromatin domains or facilitating the interactions between regulatory elements and their targets. Recent observations suggest that insulators may carry out these functions by mediating intra- and inter-chromosome interactions.

27.3.2 Intra-chromosomal interactions and insulator function in *drosophila*

Immunoflourescence microscopy using antibodies to *Drosophila* Su(Hw), Mod(mdg4)2.2, dCTCF, and CP190 shows the presence of these proteins in a punctate pattern in the nucleus. These structures, called insulator bodies, are preferentially located around the nuclear periphery, and it has been suggested that they represent sites where several individual insulator sites coalesce as a consequence of intra- and/or inter-chromosomal interactions (Figure 27.1d). The morphology of the insulator bodies is disrupted by mutations in lamin, the main component of the nuclear lamina, and various insulator components (Gerasimova *et al*., 2000, 2007; Pai *et al*., 2004). The fact that different DNA-binding insulator proteins colocalize at insulator bodies suggests that the various *Drosophila* insulators are able to interact with each other. In support of this conclusion, ChIP-chip analyses indicate that Su(Hw), BEAF-32, and CTCF overlap at 9–24% of sites where only the DNA consensus sequence for one of the proteins is present, suggesting interactions between two or more different insulators at these sites (Bushey *et al*., 2009). Interactions between insulator sites have been visualized by fluorescence in situ hybridization (FISH), showing that two individual Su(Hw) insulator sites come together to form a loop. Insertion of an additional insulator between the original two Su(Hw) sites leads to the formation of two smaller loops (Byrd and Corces, 2003). Using Chromosome Conformation Capture (3C) (Dekker *et al*., 2002), it has been shown that a *Drosophila* insulator containing the dCTCF and CP190 proteins is induced at the *Eip75B* gene to form a loop with a pre-existing insulator after cells are treated with the steroid hormone ecdysone. This loop prevents an ecdysone enhancer from activating transcription of genes that are located outside of the loop and not regulated by this hormone (Figure 27.1c) (Wood *et al*., 2011). 3C has also been used to show that two Su(Hw) insulators can interact and loop out the intervening sequences to bring an upstream Polycomb (Pc) response element (PRE) close to a downstream promoter. When one insulator is deleted, the PRE cannot associate with the promoter and histone H3 trimethylated in Lys27 (H3K27me3) present at the PRE is lost at the promoter region after the interaction between the two sequences is disrupted (Comet *et al*., 2011). The *Fab-7*, *Fab-8*, or *Mcp* insulators, which use dCTCF or GAF as the DNA binding protein, have also been found to mediate intra-chromosomal interactions. The *Fab-7* and *Mcp* insulators target the *abdominal-B* (*abd-B*) and *Antennapedia* (*Antp*) genes, which are located approximately 10 Mb apart in chromosome 3R (Figure 27.1e). These two loci colocalize in nuclei of cells in which both genes are repressed (Figure 27.1g), but deletion of *Fab-7* or *Mcp* results in a reduction of the interaction and colocalization, suggesting an important role for these two insulators in the interaction (Bantignies *et al*., 2011). The *Fab-7* and *Fab-8* insulators have also been shown to interact with a CTCF site located in the *abd-B* promoter region by testing the expression of a reporter gene (Kyrchanova *et al*., 2008). Consisting with this result, the *abd-B* promoter and the *Mcp*, *Fab-7*, and *Fab-8* elements have been found to cluster in S2 cells or fly head tissue, where *abd-B* is repressed, but not in tissues where *abd-B* is expressed (Figure 27.1f) (Cleard *et al*., 2006; Lanzuolo *et al*., 2007). The insulator sequences present in these regulatory elements, rather than other potential regulatory elements such as PREs, are responsible for these interactions (Li *et al*., 2011). These results suggest a general role for insulators in mediating intra-chromosomal interactions in order to modulate different transcriptional regulatory processes.

27.3.3 Intra-chromosomal interactions and insulator function in vertebrates

Vertebrate insulators also mediate intra-chromosomal interactions to facilitate complex transcription processes. At the *H19/Igf2* locus, the ICR immediately upstream of *H19* contains CTCF sites that are crucial for the imprinted expression of these two genes (Figure 27.2a,b). Results from 3C experiments demon-

strate that these CTCF sites can mediate allele-specific chromosome interactions that seem to control the accessibility of the *Igf2* promoter to the shared enhancer (Figure 27.2c,d). On the maternal allele of both mouse and human cells the ICR is not methylated and CTCF binds to DNA. In the mouse, the ICR interacts with CTCF sites present in the upstream differentially methylated region (DMR1) and downstream MAR3 sites that flank the *Igf2* gene. The three-dimensional structure formed as a consequence of these interactions keeps *Igf2* in an enclosed domain (Figure 27.2d) (Kurukuti *et al*., 2006; Murrell *et al*., 2004). In human cells, the ICR interacts with a CTCF Downstream Site (CTCF DS) located downstream of the shared enhancer; this interaction creates a loop that encloses the enhancer (Nativio *et al*., 2009). In both mouse and human maternal chromosomes the enhancer is unable to interact with the *Igf2* promoter. The mouse 142* allele has a mutated CTCF binding site in the ICR, and in mice that inherit this allele from their mother, the ICR interacts with DMR2 instead of DMR1 and MAR3; this is accompanied by

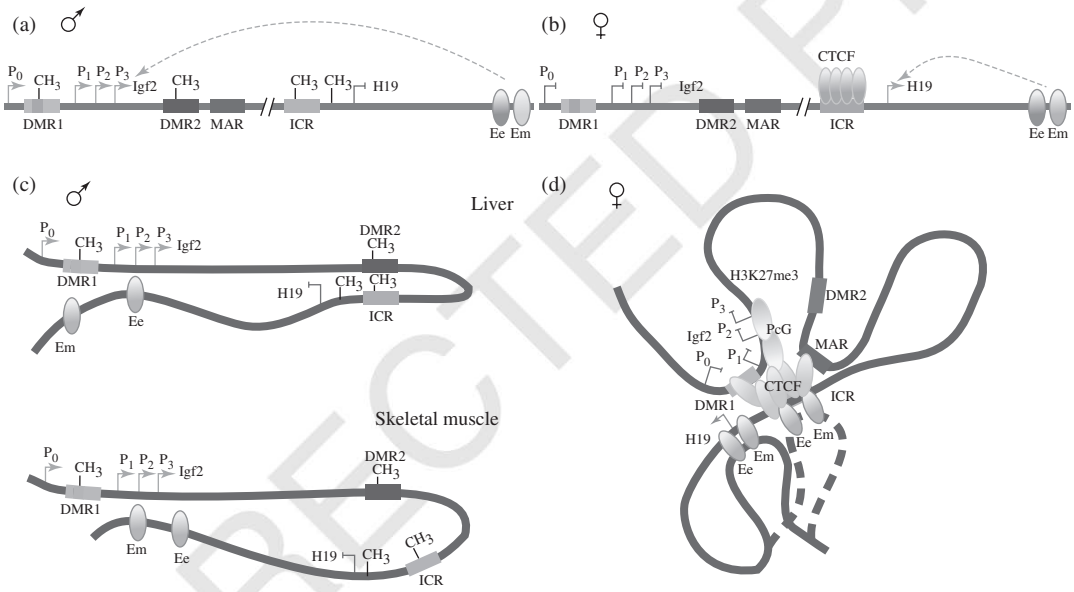

Figure 27.2 *(a,b) Allele-specific interactions at the mouse* H19/Igf2 *imprinted locus. The maternally expressed non-coding* H19 *gene is located approximately 90 kb downstream from the gene encoding Insulin-like growth-factor 2 (*Igf2*) that is expressed exclusively from the paternal allele. The imprinting control region (ICR) upstream of* H19 *contains four CTCF binding sites and is essential for regulation of the entire locus. Differentially methylated regions (DMRs), such as DMR1 upstream of* Igf2 *promoters (P1, 2, 3) and DMR2 within* Igf2*, act in concert to regulate allele-specific expression patterns from the downstream enhancers (Ee and Em). CH_3 indicates methylated DNA and ovals represent enhancers. (c,d) Diagram of three-dimensional arrangements of the chromatin mediated by CTCF in maternal and paternal alleles. (e) Linear representation of the mouse β-*globin *locus. Four globin genes (solid arrows) are located within a cluster of olfactory receptor genes (open arrows). Developmentally regulated globin expression (εγ and βh1 in primitive erythroid cells; β-major, and β-minor in definitive erythroid cells) is regulated by a series of* cis*-acting regulatory elements surrounding the locus. An upstream locus control region (LCR) containing six DNase I-hypersensitive sites is required for high-level transcription. Three CTCF binding sites have been identified upstream (5′HS85, 5′62/60, and 5′HS5) and one 20 kb downstream (3′ HS1) of the gene. Black arrows indicate hypersensitive sites. (f,h). Diagrams illustrating lineage-specific CTCF-mediated interactions and* globin *gene expression profiles in erythroid progenitors (f), definitive erythroid cells (g), and non-erythroid brain cells (h)*

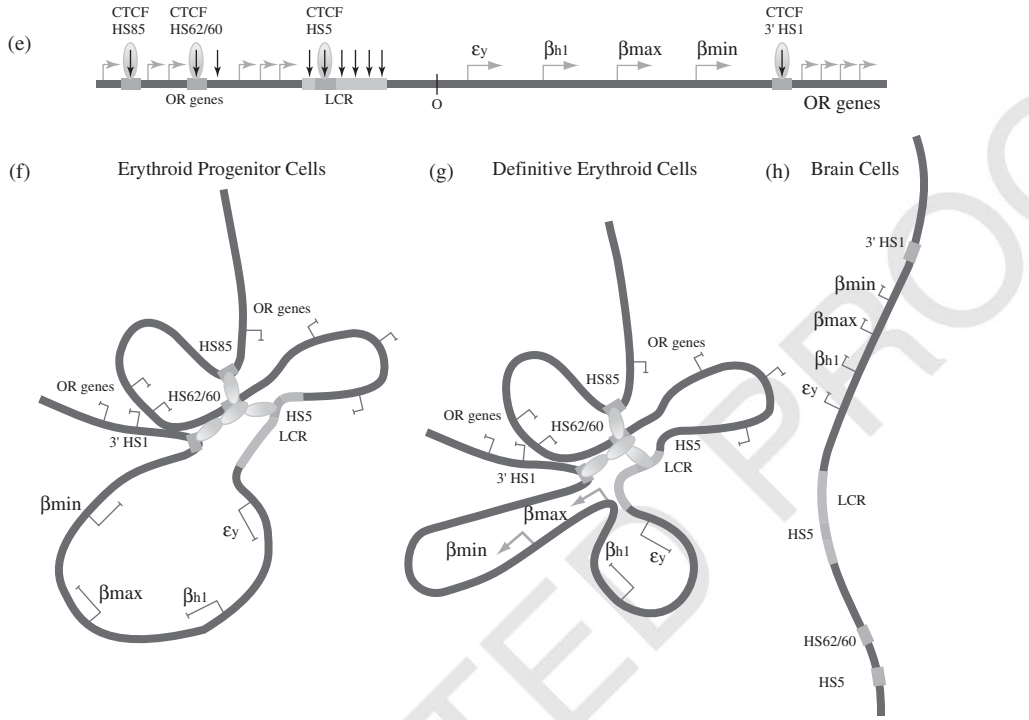

Figure 27.2 *(continued)*

biallelic expression of *Igf2*. On the other hand, there is no effect if 142* is paternally inherited (Kurukuti *et al*., 2006; Pant *et al*., 2003). On the paternal chromosome, the ICR is methylated and CTCF does not bind. In mouse cells the ICR now interacts with the DMR2 site located downstream of *Igf2* and allows the enhancer to interact with the *Igf2* promoter (Figure 27.2c) (Kurukuti *et al*., 2006; Murrell *et al*., 2004) whereas in human cells the ICR cannot interact with the CTCF DS and allows *Igf2* expression as well (Nativio *et al*., 2009). These results suggest that imprinted expression of the *H19/Igf2* locus is dependent on a specific three-dimensional organization of the region mediated by interactions between CTCF sites.

Other examples of insulator-mediated regulation of gene expression include the cytokine interferon-γ (IFNG) locus, which contains three conserved CTCF binding sites located upstream, downstream, and within the gene 1.5 kb from the TSS in human and mouse cells. The CTCF sites display cell type-specific interactions and also interact with the enhancers in the locus in specialized T helper 1 (Th1) cells, although the enhancers do not have CTCF binding sites (Hadjur *et al*., 2009; Sekimata *et al*., 2009). Both the three-dimensional conformation of the locus and robust *IFNG* expression in Th1 cells are dependent on the presence of CTCF (Sekimata *et al*., 2009).

In addition to enhancer promoter interactions, CTCF also mediates gene to gene communication and facilitates their co-regulation. Interactions between the Insulin (*INS*) gene and the *SYT8* gene, which is located over 300 kb away, depend on CTCF. Increase of the interaction by glucose is accompanied by a rise in *SYT8* expression, which is important for insulin secretion in islets cells (Xu *et al*., 2011). Insulin synthesis and secretion from pancreatic β-cells are tightly regulated and their deregulation causes diabetes. The interaction between CTCF sites brings an enhancer located adjacent to the INS gene close to the promoter of the *SYT8* gene and may be important for the coregulation of the two genes and functional cooperation

between the two proteins in islet cells (Figure 27.1c). At the Major Histocompatibility Class II (MHC-II) locus, the XL9 element contains CTCF sites and is located in the intergenic region between the MHC-II genes *HLA-DRB1* and *HLA-DQA1*. 3C experiments detect an interaction between XL9 and proximal promoter elements of these two MHC-II genes. These interactions also depend on the Class II transactivator (CIITA), which can associate with CTCF (Majumder *et al*., 2008). The chromatin conformation mediated by the association between CIITA and CTCF is required for transcription of the MHC-II genes. Loss of CTCF binding due to methylation of the DNA leads to the loss of expression from *HLA-DQ* in Laz221 cells, which were obtained from a patient with acute lymphocytic leukemia (Majumder and Boss, 2011).

CTCF-mediated chromatin loops also seem to demarcate active or repressive domains. At the Apolipoprotein (APO) locus, CTCF sites interact with each other in Hep3B cells and lead to the formation of two transcribed loops with *APOC3/APOA4/APOA5* in one loop and *APOA1* in the other. Consistent with this model based on 3C data, CTCF knockdown leads to a decrease of *APOC3/APOA4/APOA5* but an increase of *APOA1* expression (Mishiro *et al*., 2009). Similarly, the human Hox genes have been shown to be spatially clustered when they are silenced via a specific three-dimensional architecture of the locus, as is also the case for *Drosophila*. CTCF is present at the sites of contact and appears to be responsible for the establishment and/or maintenance of this architecture (Ferraiuolo *et al*., 2010). At the HOXA locus, the CTCF site flanking *HOXA9* mediates a loop that separates *HOXA9-HOXA13* from the other *HOXA* genes and keeps repressive histone modifications within the domain defined by this loop in differentiated cells human lung fibroblasts (IMR90) (Kim *et al*., 2011).

The role of CTCF in creating a specific three-dimensional arrangement of the DNA to regulate gene expression has been also extensively studied at the β-globin locus, where there are several CTCF sites that are conserved in mice and humans. These sites are located within two DNase I hypersensitive regions that flank the β-globin locus: the 5′HS5 located in the Locus Control Region (LCR) and the downstream 3′HS1; additional CTCF sites are located further away at either end of the locus Figure 27.2e). The 5′HS5 and 3′HS1 sites interact with each other as well as with the more distal CTCF sites flanking the locus, and these interactions are severely weakened in the absence of the CTCF protein. The interactions between 5′HS5 and 3′HS1 are cell type specific but they are not required for proper transcription of the *β-globin* genes. These interactions exist in both erythroid cells, in which the *β-globin* genes are transcribed, and in erythroid progenitor and fibroblast cells, in which the genes are not expressed, but not in non-expressing brain cells (Figure 27.2g) (Hou *et al*., 2010; Splinter *et al*., 2006; Tolhuis *et al*., 2002). Mutation of 3′HS1 destabilizes the long-range interactions among these CTCF sites but mutation or deletion of 3′HS1 or 5′HS5 can neither affect the expression kinetics nor levels of the *β-globin* genes in erythroid cells (Farrell *et al*., 2000; Splinter *et al*., 2006). Nevertheless, a number of CTCF sites present outside of the β-globin locus interact with each other but not with 5′HS5 or 3′HS1, and some of these interactions are both cell type specific and transcription relevant. These interactions can only be detected in either erythroid cells or non-erythroid cells (Figure 27.2f,h). Knockdown of CTCF, which leads to a global reduction of both groups of interactions, can negatively affect *β-globin* gene transcription (Hou *et al*., 2010). It is possible that these interactions are required for *β-globin* expression, although it is also possible that the observed changes in transcription are due to indirect effects of CTCF knockdown. Therefore, it appears that CTCF-mediated interactions involving 5′HS5, 3′HS1, and other CTCF sites in the adjacent region create a three-dimensional organization of the *β-globin* locus before the commitment to the erythroid lineage that is not necessary or sufficient for the activation of the *globin* genes. The fact that CTCF-mediated interactions among these sites are not observed in brain cells suggests that this organization may play an earlier role during differentiation after the commitment to ectodermal or mesodermal fates. New interactions established later during erythroid differentiation and encompassing a different set of more distally located CTCF sites may be involved in the establishment of a global architecture that is actually more directly responsible for the expression of the β-globin locus (Figure 27.2f,h).

CTCF-mediated interactions also regulate non-transcription processes through the formation of loops. For example, V(D)J recombination is highly regulated during B cell development. Igh rearrangement in pro-B cells begins with D_H to J_H rearrangement followed by rearrangement of a V_H gene segment to D_HJ_H. In mice, there are more than 100 V_H genes spanning a 2.5 Mb region whereas the J_H genes occupy a 2 kb region. How do all the V genes access the small J cluster in the Igh loci? ChIP-seq experiments in pro-B cells reveals about 60 CTCF sites located throughout the V_H region and two clusters within the other parts of Igh locus (Degner *et al*., 2011). CTCF plays a critical role in this process through its interaction with cohesin, whose localization changes in a cell lineage-specific manner (Degner *et al*., 2009). Three-dimensional measurement of chromatin compaction in pro-B cells using FISH indicates that reduction of CTCF binding results in a decrease in Igh locus compaction, suggesting that CTCF probably brings V_H close to the D_H-J_H region. There are two additional clusters of CTCF sites, one is next to D_H and the other at the 3′ regulatory region of J_H. They interact strongly in pre-pro-B cells and pro-B cells, but only minimally in murine embryonic fibroblasts, creating a distinct domain before D_H-J_H recombination. These interactions depend on CTCF. 3C also reveals the interaction of these sites with an intronic enhancer (Eμ) required for the antisense transcription of D_H (Degner *et al*., 2011). Antisense transcription through the D_H locus precedes D_H-J_H rearrangement and has been proposed to make the D_H region accessible for subsequent rearrangement. Thus, the interaction excludes V_H from the domain and prevents the antisense transcription as well as involvement of V_H before D_H-J_H rearrangement. Knocking down of CTCF confirms the blocking effect of CTCF on the interaction between the Eμ enhancer and the V_H genes, as antisense transcription from V_H is increased when CTCF is downregulated (Degner *et al*., 2011). In conclusion, CTCF mediated intra-chromosomal interactions regulate V_H-D_H-J_H recombination in a spatial and temporal manner.

Chromatin Interaction Analysis by Paired-End Tag (ChIA-PET) has been used to map 1480 *cis* interactions mediated by CTCF in mouse ES cells (Handoko *et al*., 2011). Analysis of these interactions reveals five different types of CTCF-mediated chromatin loops that separate chromatin domains with different epigenetic modifications (Figure 27.3). Clustering of histone modifications in and around these chromatin loops results in the classification of five distinct patterns. Category I is defined by the presence of active histone modifications such as H3K4me1, H3K4me2, and H3K36me3 inside the loops, whereas the repressive marks H3K9me3, H3K20me3, or H3K27me3 are depleted inside but present outside of the loops. Category II loops show the opposite distribution of histone modifications, with extensive H3K9me3, H3K20me3, and H3K27me3 in the loops, indicating the formation of repressive domains inside of the loops. These two types of interactions may create independent domains for different regulation of gene expression as is the case in the APO locus discussed above. Category III loops are suggested to form hubs for enhancer and promoter activities. These loops show enrichment of enhancer marks H3K4me1 and H3K4me2 inside of the loop and the promoter mark H3K3me3 at the end of the loop. The interaction between CTCF sites forming these loops could bring enhancers closer to the target promoters for transcription activation. Genes with enhancers located more than 10 kb away are significantly upregulated in ES cells versus neuronal stem cells if they fall into loops than those that do not. Category IV loops show opposite chromatin states flanking the end of chromatin loops, but do not exhibit any specific pattern of histone modifications within the loops. Category V loops do not show any specific signatures. The function of the last two types of loops are not yet clear (Handoko *et al*., 2011).

27.3.4 Insulator-mediated inter-chromosomal interactions

Insulators can mediate not only intra-chromosome but also interactions between sequences located in different chromosomes. Insertion of the *Fab-7* insulator at a specific chromosome location in transgenic *Drosophila* results in increased interactions between two sequences located in different chromosomes (Bantignies *et al*., 2003). In the nuclei of cells from wild type flies, the *abd-B* gene in chromosome 3R is

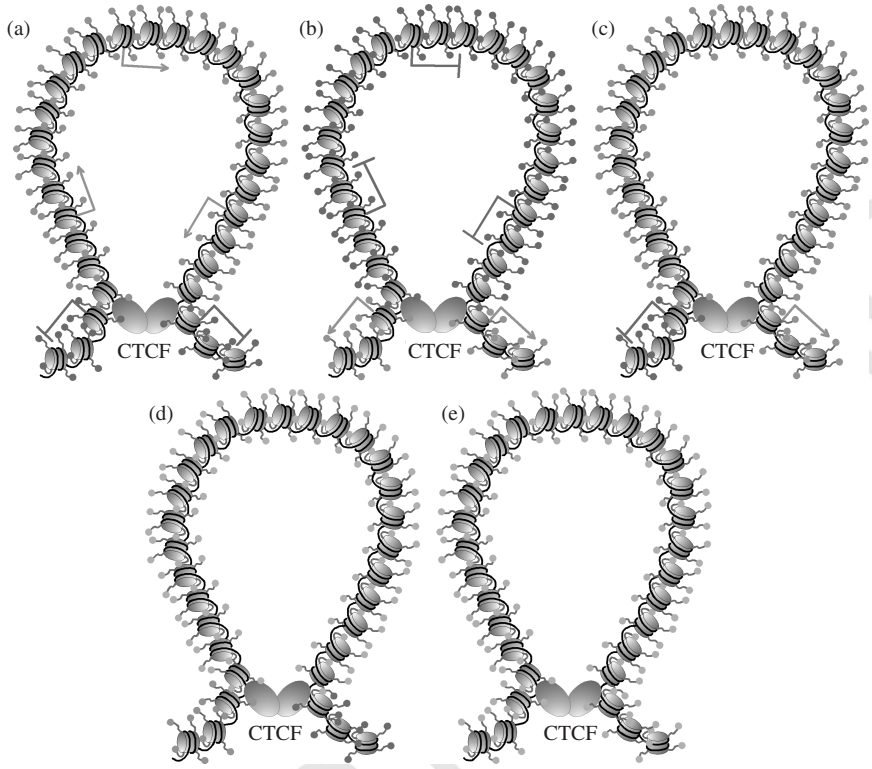

Figure 27.3 *Domains created by interactions between CTCF insulators in mouse embryonic stem cells. Actively transcribed genes are represented by an arrow and silenced genes by two perpendicular lines; nucleosomes and the histone tails are represented in gray, with active histone modifications indicated as light gray spheres and repressive modifications as dark gray spheres. DNA is represented in black and CTCF as an oval. (a) CTCF forms a loop to separate a domain containing active histone modifications and transcribed genes from repressive marks and silenced genes. (b) CTCF forms a loop to separate a domain containing repressive histone modifications and silenced genes from active marks and transcribed genes. (c) CTCF forms a loop containing nucleosomes enriched in mono- and dimethylated H3K4, and trimethylated H3K4 at the boundaries of the loops, whereas the active transcription modification H3K36me3 and repressive H3K27me3 mark are observed outside the loops on opposite sides. (d) A fourth class of loops formed by CTCF lack specific histone modifications, while active H3K4 methylation marks are observed in one side and repressive H3K9, H3K20, and H3K27 methylation modifications are present in the other side. (e) The rest of the loops formed by CTCF do not show specific chromatin modifications*

found close to the *scalloped* (*sd*) gene present in the X chromosome at a frequency of 7% by FISH. When a copy of *Fab-7* is inserted at the *sd* locus, the two loci colocalize in 23% of the nuclei. If the endogenous *Fab-7* element is mutated in this transgenic strain, the frequency of interaction is reduced to that of wild type (Bantignies *et al*., 2003). Insertion of *Mcp* also leads to more frequent interactions between sites in the *Drosophila* genome (Vazquez *et al*., 2006).

In mouse cells the *H19* ICR on chromosome 7 has been shown to interact with genes on different chromosomes, including *Wsb1/Nf1* on chromosome 11, *Abcg2* on chromosome 6, and *Osbpl1a* on chromosome 18 (Ling *et al*., 2006; Zhao *et al*., 2006). The majority of these inter-chromosomal interactions take place primarily with the maternally inherited *H19* ICR. Knockdown of CTCF or mutation of CTCF

sites on the *H19* ICR in the maternal chromosome abrogates these associations, suggesting that these interactions are CTCF-dependent (Ling *et al*., 2006; Zhao *et al*., 2006). Interestingly, imprinted loci are strongly overrepresented among the regions involved in inter-chromosome interactions with the *H19* ICR. The clustering of these imprinting genes by inter-chromosomal interactions to form an 'imprinting interactome' may facilitate the regulation of these genes in *trans*. When the CTCF site is mutated in the maternal *H19* ICR, the interactions are abrogated and the transcription of these imprinted genes is also altered; expression from the paternal chromosome of the *Wsb1/Nf1* gene is reduced by 50-fold, expression of the *Impact* gene is reduced by 50% and transcription of *Osbpl1a* is increased by 2.5-fold (Ling *et al*., 2006; Zhao *et al*., 2006). In addition, the asynchronous replication timing of these genes in spermatogonia is switched from late to early (Sandhu *et al*., 2009). Therefore, CTCF mediated inter-chromosomal interactions are important for the nonallelic regulation of the epigenetic status of multiple genes in *trans*. One important advantage for such regulation is that the interactions can be reprogrammed during germline development, when the epigenetic states of imprinted domains are reprogrammed. In the testis, when the maternal allele is turned to the paternal mode, the interactome can be observed in spermatogonia, but not in spermatocytes and round spermatids, where the reprogramming is complete (Sandhu *et al*., 2009).

Experiments described in the previous section using ChIA-PET to map the CTCF interactome also detect 336 *trans* interactions in mouse ES cells (Handoko *et al*., 2011). In mammals, CTCF also mediates inter-chromosomal interactions that results in the homologous pairing of the 15q11-q13 locus, which is associated with autism (Meguro-Horike *et al*., 2011), and pairing between the two X chromosomes in female cells, which is important for X-chromosome inactivation. To achieve the mutually exclusive designation of active X and inactive X, it is necessary for the two X chromosomes to communicate in *trans* through homologous pairing. Pairing depends on a 15-kb region within the Tsix and Xite loci. ChIP experiments show that CTCF binds both Tsix and Xite elements in female embryonic stem cells. CTCF knockdown reduces the frequency of X-X pairing to background levels in wild type embryonic stem cells and embryoid bodies (Xu *et al*., 2007).

Taken together, these results suggest that insulators can mediate long-range intra- and inter-chromosome interactions at different loci throughout the genome either by interactions with other insulator sites or association with other factors. The result of these interactions is a specific three-dimensional arrangement of the chromatin that can have different effects on chromosome biology, depending on the nature of the sequences brought together by CTCF.

27.3.5 Insulators and the nuclear matrix

Observations described in the preceding sections suggest that insulators mediate intra- and inter-chromosomal interactions to regulate various nuclear processes. In doing so, insulators establish a specific three-dimensional arrangement of the chromatin fiber in the nucleus. In addition to interactions among different DNA sequences, insulator may also mediate the association of chromatin with the nuclear lamina. In *Drosophila*, it has been shown that two loci containing the Su(Hw) insulator localize more frequently at the periphery of the nucleus and close to each other. Without the Su(Hw) insulators, the two loci locate inside of the nucleus and separate from each other (Gerasimova *et al*., 2000). Although insulator bodies, where multiple insulators come together, are present throughout the nucleus, they seem to localize preferentially in the nuclear periphery. *Drosophila* insulator proteins fractionate with components of the nuclear lamina, suggesting a direct or indirect interaction between the two (Byrd and Corces, 2003). Furthermore, the dTopors protein, which is located in the nuclear periphery, interacts with both Mod(mdg4)2.2 and lamin, and mutations in the *lamin Dm0* gene result in disruption of this localization and insulator activity (Capelson and Corces, 2005).

It appears that interaction with a nuclear substrate is also required for proper function of the CTCF insulator in vertebrates. For example, it has been shown that CTCF interacts with nucleophosmin, a nuclear matrix protein that is located in the nucleolus and associates with CTCF at the chicken *HS4* insulator in the β-globin locus. Consistent with this distribution, transgenic copies of reporter genes containing the chicken *HS4* insulator show preferential localization to the outer part of the nucleolus, and this localization depends on the presence of an intact CTCF binding site (Yusufzai *et al*., 2004). In addition to nucleophosmin, other nuclear matrix proteins such as lamin may also be important for CTCF function. In human cells, many of the lamin B1-associated domains (LADs) in the genome are flanked by CTCF (Guelen *et al*., 2008). Lamin A also works together with CTCF to establish nucleolar localization. The D4Z4 human subtelomeric repeats localize telomeres to the nuclear periphery and display insulator activity in transgenic cell lines. CTCF and Lamin A both bind to D4Z4 and are required for proper perinuclear localization. Reduction of either protein results in a decrease in the localization of telomeres at the nuclear periphery and impairs insulator function (Ottaviani *et al*., 2009a,b).

These results suggest that insulators mediate interactions among DNA sequences and nuclear structures such as the nuclear lamina and the nucleolus. It is not clear whether the role of these interactions is to bring the DNA to specific nuclear compartments or to simply attach these sequences to a substrate in order to maintain the topological state of insulator-mediated loops.

27.4 Regulation of insulator function

Insulator proteins are present at numerous sites in the genome and they mediate long-range chromatin interactions that regulate gene expression and play important roles during development or differentiation. Therefore, the cell must possess mechanisms to regulate insulator function in order to guarantee proper execution of these nuclear processes (Figure 27.4).

One strategy used by the cell may be to control the interaction between insulator proteins and their DNA target sequences. For example, 40–60% of CTCF sites are common among $CD4^+$ T, HeLa, and Jurkat cells (Cuddapah *et al*., 2009) and the rest of the sites occupied by CTCF vary between these two cell types. Similarly, around 30% of CTCF sites are different in human IMR90 fibroblasts and U937 erythroid progenitor cells (Kim *et al*., 2007). In *Drosophila*, 18% of Su(Hw), 18% of dCTCF, and 11% of BEAF sites are specific for embryonic Kc cells compared to differentiated neural Mbn2 cells (Bushey *et al*., 2009). In vertebrates, the best understood mechanism to control CTCF occupancy is through DNA methylation of CpG dinucleotides within and around the CTCF binding site. This has been well studied in the *H19/Igf2* locus, where CTCF only binds to the unmethylated maternal allele but does not bind to the methylated paternal allele (Figure 27.2a,b) (Bell and Felsenfeld, 2000; Hark *et al*., 2000). CTCF binding to methylation-free sites is also documented for other genes, including both imprinted and non-imprinted loci, such as *Rasgrf1*, *KvDMR1*, *GRB10*, *INK/ARF*, and *DM1* (Filippova *et al*., 2001; Fitzpatrick *et al*., 2007; Hikichi *et al*., 2003; Rodriguez *et al*., 2010; Yoon *et al*., 2005). Therefore, regulation of CTCF binding via DNA methylation may be a general strategy to control the function of this insulator.

A second mechanism to regulate insulator function that is shared by *Drosophila* and vertebrates is the use of proteins that can compete for binding to the same DNA sequence. In *Drosophila*, DREF (DNA replication-related element-binding factor) has been characterized as a transcription factor that shares binding sites with BEAF. It is possible that DREF regulates BEAF binding through competition for the same DNA sequences (Hart *et al*., 1999). Vertebrate cells have a homolog of CTCF named CTCFL or BORIS (Brother of the Regulator of Imprinted Sites). These two proteins share homology of the central 12 zinc finger DNA binding domain and they can recognize the same DNA binding sequences (Loukinov *et al*., 2002). BORIS is normally present in the testis during germ-line development, where it can stimulate

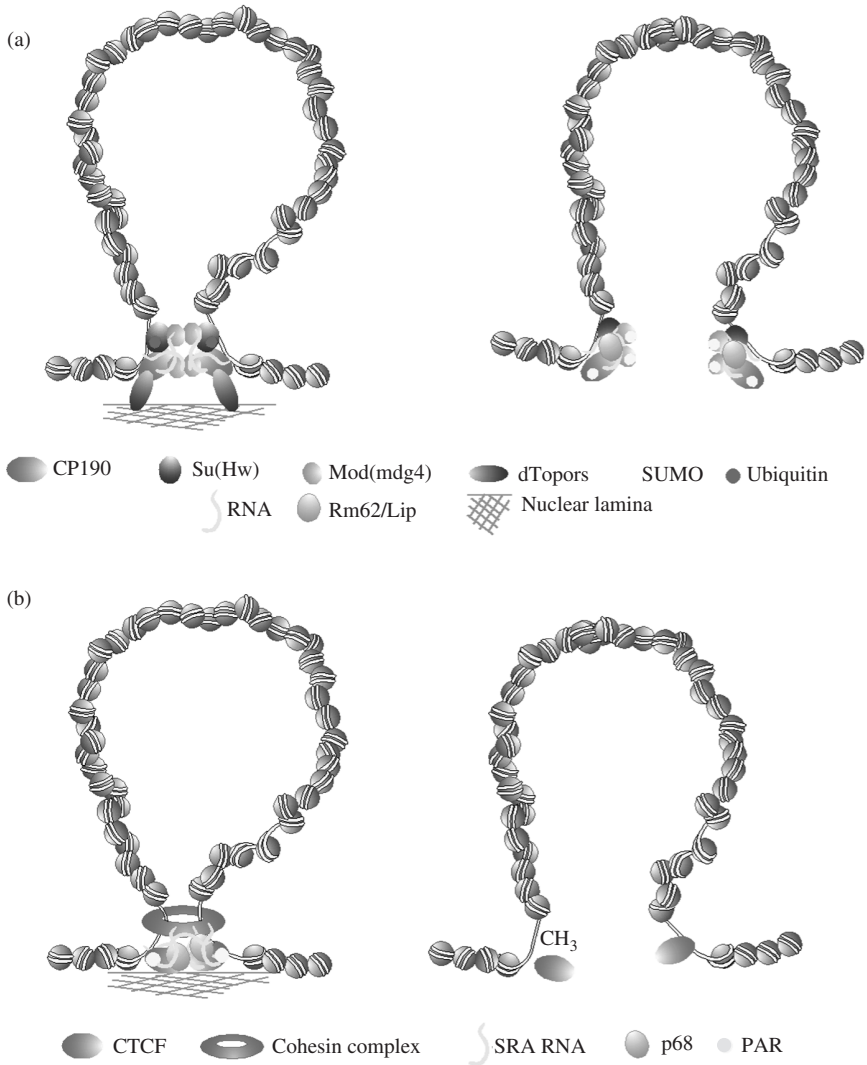

Figure 27.4 *Similarities in the regulatory mechanisms of insulator function in* Drosophila *and vertebrates. (a) Regulation of insulator activity in* Drosophila. *The left panel shows two active insulators coming together to make a functional loop and the right panel displays two inactive insulators unable to form a loop. At an active insulator site, dTopors is present, Rm62/Lip is not present, Su(Hw) is ubiquitinated, Mod(mdg4)2.2 and CP190 are not sumoylated, and dTopors serves as a bridge to the nuclear lamina/matrix. At inactive insulator sites, dTopors is absent and Su(Hw) is not ubiquitinated, whereas Mod(mdg4)2.2 and CP190 are sumoylated. Rm62/Lip is present and bound to RNA. Under these conditions, the two insulator sites cannot interact to form a loop. Absence of dTopors also precludes interactions with the nuclear lamina. (b) Regulation of insulator activity in vertebrates. The left panel shows two active CTCF insulators coming together to make a functional loop and the right panel displays two inactive CTCF insulators unable to form a loop. At an active insulator site, CTCF is present and poly(ADP-ribosyl)ated and, directly or indirectly, interacts with the nuclear matrix/lamina; in addition, the SRA RNA and p68 are present while the cohesin complex brings together the two DNA molecules forming the base of the loop. At inactive insulator sites, CTCF is either not bound because of the DNA is methylated or it is bound but not modified by poly(ADP-ribosyl)ation; cohesin, SRA RNA, and/or p68 are absent. Under these conditions the two insulator sites cannot interact to form a loop*

the histone-methyltransferase activity of PRMT7 (Protein arginine N-methyltransferase 7) and contribute to the efficient DNA methylation of BORIS/CTCF sites in male germ line (Jelinic *et al*., 2006). Therefore, BORIS can play a role in writing the epigenetic marks that will be read by CTCF. Usually, BORIS is expressed in a mutually exclusive manner with CTCF and only influences CTCF function by its effects in the germ line (Loukinov *et al*., 2002). However, BORIS can be aberrantly expressed under particular conditions; for example, reduction of CTCF in normally BORIS-negative human fibroblasts results in derepression of BORIS expression (Renaud *et al*., 2007). Abnormally expressed BORIS may affect the activity of CTCF by various mechanisms. First, BORIS can compete with CTCF for binding sites and lead to alterations in gene expression; although BORIS shares the zinc finger domain with CTCF, the two proteins are different in the N- and C-terminal regions that constitute approximately two-thirds of the full length amino acid sequence of these proteins and could recruit different functional partners. For example, the transcription factor Sp1 can interact with BORIS but not CTCF (Hong *et al*., 2005; Kang *et al*., 2007). Second, the two proteins differ in their ability to bind methylated DNA. In contrast to CTCF, the binding of BORIS to its target site is methylation-independent in vitro and methylation-preferential in vivo at the *H19/Igf2* ICR (Nguyen *et al*., 2008). As a consequence, BORIS could bind to methylated CTCF/BORIS sites, where CTCF is not normally present, and affect the expression of adjacent genes. Third, aberrantly expressed BORIS together with ubiquitously present PRMT7 might change the methylation status of previously non-methylated sites and evict CTCF. The binding of BORIS and the different transcription factors recruited to genes harboring the CTCF/BORIS sites could lead to totally different expression patterns from those observed when only CTCF is present in the cell (Hong *et al*., 2005). Not surprisingly then, aberrant expression of BORIS can lead to disease and cancer as we discuss in the following.

Since CP190 or cohesin are required for inter-insulator interactions in *Drosophila* and vertebrates, respectively, the recruitment of these components may control the activity of the various insulators in these organisms (Figure 27.4). In *Drosophila*, 17% of CP190 sites present in Kc cells and 14% in Mbn2 cells were found to be cell type-specific (Bushey *et al*., 2009). After heat shock, most of CP190 disassociates with the chromatin while the localization of the DNA binding proteins is not affected (Oliver *et al*., 2011; Wood *et al*., 2011). In vertebrates, depletion of cohesin components without affecting CTCF expression can also lead to disruption of chromatin interactions and changes in the expression of genes under CTCF control (Hadjur *et al*., 2009; Hou *et al*., 2010; Mishiro *et al*., 2009; Nativio *et al*., 2009). In the *Igh* locus, although the binding of CTCF is not altered during B cell differentiation, the recruitment of cohesin is cell type specific (Degner *et al*., 2009). The binding of cohesin, in turn, is antagonized by OCT4 which is also a cofactor of CTCF. Loss of OCT4 in neuronal progenitor cells results in loading of cohesin and chromosome looping, which contributes to heterochromatin partitioning and selective gene activation across the *HOXA* locus (Kim *et al*., 2011). Intriguingly, OCT4 facilitates CTCF function in X chromosome pairing (Donohoe *et al*., 2009). It is possible that cohesin and OCT4 associate with CTCF to mediate intra- or inter-chromosome interactions, respectively.

The recruitment of cofactors to the DNA binding insulator proteins may be regulated through post translational modification. dTopors, in addition to serving as an attachment point for insulators to the nuclear lamina, it has E3 ubiquitin ligase activity. This activity is required for proper insulator function (Figure 27.4a). Its substrate has not been clearly identified but Su(Hw) is a likely candidate, since over-expression of dTopors enzymatic activity reverses the effect of *mod(mdg4)* mutations on the ability of Su(Hw) to interact with chromatin (Capelson and Corces, 2005). In addition, modification of Mod(mdg4)2.2 and CP190 by sumoylation inhibits insulator function. Disruption of the SUMO conjugation pathway improves the enhancer-blocking function of a partially active insulator, indicating that SUMO modification acts to negatively regulate the activity of the Su(Hw) insulator. Sumoylation does not affect the ability of CP190 or Mod(mdg4)2.2 to bind chromatin (Capelson and Corces, 2006). Interestingly, dTopors inhibits sumoylation of Mod(mdg4)2.2 and CP190. Therefore, this protein may have a double effect on

insulator function by ubiquitinating some insulator components and inhibiting the sumoylation of others. O-glycosylation of BEAF can be detected in *Drosophila* embryonic cells. However, it is not clear whether glycosylation is required for the function of this insulator (Pathak *et al*., 2007).

In vertebrates, CTCF can also be covalently modified by poly(ADP-ribosyl)ation (Figure 27.4b). CTCF interacts with Poly(ADP-ribose) polymerase (PARP)-1 both in vivo and in vitro (Guastafierro *et al*., 2008). Both proteins colocalize at the *H19/Igf2* ICR and overlap genome wide at more than 140 CTCF target sites in the mouse genome, including both imprinted and non-imprinted loci (Yu *et al*., 2004). Poly(ADP-ribosyl)ation does not affect the ability of CTCF to bind DNA but it is required for its insulator function. For example, mutation of the poly(ADP-ribosyl)ation site in CTCF or treatment with the PARP inhibitor ABA (3-aminobenzamide), can compromise imprinting of the *H19/Igf2* locus, resulting in biallelic expression of *Igf2*, but CTCF binding to the *H19/Igf2* ICR is not affected (Farrar *et al*., 2010; Yu *et al*., 2004). In addition to Poly(ADP-ribosyl)ation of CTCF, PARP-1 can also modify itself. The modified PARP-1 can inhibit DNMT1 activity and control DNA methylation patterns (Guastafierro *et al*., 2008). This may in turn control the binding of CTCF. Thus poly(ADP-ribosyl)ation may affect the function of CTCF by modulating its interaction with DNA and with other proteins (Figure 27.4b). CTCF can be also phosphorylated at the C-terminus by the protein kinase CK2 (Casein kinase 2). Mutation of the phosphorylated residue in CTCF does not alter its nuclear localization or DNA binding in vitro, but enhances repression of the *c-myc* promoters. A phospho-mimetic mutant behaves in the opposite fashion (El-Kady and Klenova, 2007; Klenova *et al*., 2001). Thus, phosphorylation may affect the interaction of CTCF with other proteins and modulate its activity.

Insulator activity can be also regulated by the RNAi machinery. The *Drosophila* Su(Hw) insulator contains an RNA component whose proper expression or assembly requires components of the RNAi machinery. The activity of this insulator decreases in the presence of mutations in components of the RNAi machinery while insulator function is restored by mutations in RNA helicase Rm62. These observations have led to a model suggesting that insulator bodies contain RNA whose synthesis requires RNAi proteins. Rm62 may interact with this RNA to decrease insulator function (Figure 27.4a) (Lei and Corces, 2006). In vertebrates, the RNA helicase p68 (DDX5) and its associated noncoding RNA, steroid receptor RNA activator (SRA), bind to CTCF, and are both essential in vivo for insulator function at the *H19/Igf2* ICR. Contrary to the inhibitory role of the *Drosophila* RNA helicase Rm62 on insulator function, p68 is needed to positively regulate insulator function in vertebrate cells (Figure 27.4b). Knockdown of p68 decreases the insulator activity of the *H19/Igf2* ICR and increases the expression of *Igf2* as previously observed in CTCF knockdown cells (Yao *et al*., 2010). In *Drosophila* cells, Rm62 interacts with CP190, while in vertebrates p68 is required for the recruitment of cohesin. In spite of the opposite effect of the two helicases on insulator function, the conservation of this mechanism between insects and vertebrates is striking.

27.5 Insulators and the external/internal cellular environment

27.5.1 Response to hormones

CTCF sites are present immediately adjacent to thyroid hormone response elements. At the chicken lysozyme and human *c-myc* regions, CTCF sites can act in the typical insulator role and mediate enhancer blocking in the absence of thyroid hormone. However, treatment with thyroid hormone abrogates this activity (Lutz *et al*., 2003). The CTCF insulator can thus ensure that the adjacent genes are not transcriptionally affected by enhancers or silencers until the cell needs to respond to thyroid hormone. In *Drosophila*, treatment with the steroid hormone 20-hydroxyecdysone induces binding of CP190 to a CTCF site present in the *Eip75B* gene. This new CP190 site increases the interaction with a pre-existing CP190 site located in the 5′ region of the gene, forming a chromatin loop. This loop prevents the interaction between the

ecdysone receptor binding site (enhancer) and the promoter of genes located 5′ to *Eip75B*. RNAi treatment of CP190 causes a increase in the interaction (Wood *et al*., 2011). In the mouse, 1,25-dihydroxyvitamin D_3 can induce the recruitment of CBP (CREB-binding protein) and transcription of the *Receptor Activator of NF-κB Ligand* (*RANKL*) gene, which encodes a protein important during osteoclastogenesis. The CBP-mediated increase in acetylation of H3K4 in the *RANKL* region is extended to sequences located between two CTCF/cohesin sites, suggesting that insulators create domains necessary for proper hormone response (Martowicz *et al*., 2011). Consistent with this idea, computational analysis of estrogen receptor (ER) sites combined with expression profiles and genome-wide localization maps of CTCF also indicates that CTCF partitions the human genome into distinct ER regulatory domains. CTCF domains encompassing genes whose transcription decreases after estrogen stimulation have a different distribution of ER binding sites than domains containing genes whose transcription increases (Chan and Song, 2008). Experimental results confirm this observation. There are two CTCF sites surrounding the well defined ER target *TFF1* locus. The two sites are separated by 40 kb but cluster in the nuclear space in a manner that depends on CTCF. Although the interaction does not correlate with estrogen treatment and transcription, the *TFF1* gene cannot respond to estrogen without this interaction, and the entire locus displays a heterochromatin-like structure, as is the case in estrogen non-responsive breast cancer cells (Zhang *et al*., 2010). Therefore, the chromatin architecture mediated by CTCF demarcates active versus inactive regions, building a framework of adjacent chromosome territories for proper hormone response.

27.5.2 Insulators and viral expression

Insulators not only regulate chromatin structure in eukaryotic cells, but they can also organize the genome of viruses. CTCF sites have been identified in the double stranded DNA genome of various human herpes viruses like Epstein–Barr virus (EBV), herpes simplex virus 1, and Kaposi's sarcoma-associated herpes virus (KSHV) (Chau *et al*., 2006; Chen *et al*., 2007; Stedman *et al*., 2008). Functional studies of the role of insulators in viral biology have focused on aspects of the latency cycle such as repression, activation, and insulation of latent transcripts, which can vary depending on the host cell or tumor type.

Genome wide comparisons of epigenetic modifications among different latency types reveal that CTCF binds at several key regulatory nodes in the genome of EBV (Tempera *et al*., 2010). The establishment and maintenance of EBV latent infection requires distinct viral gene expression programs that are determined largely by promoter selection. For example, the Q promoter (Qp) determines the type I expression pattern and the C promoter (Cp) determines type III gene expression. In type I latency Qp is active while in type III latency Qp is epigenetically silenced and Cp is active (Chau *et al*., 2006; Tempera *et al*., 2010, 2011). CTCF can bind between the OriP enhancer and the Cp and the Qp promoters. CTCF binding at Qp is required for stable maintenance of the EBV episome in 293 cells (Chau *et al*., 2006; Tempera *et al*., 2010). Recently, 3C experiments have shown that CTCF establishes different chromatin architectures between the OriP enhancer and the Qp or Cp promoters. OriP is in close proximity to Qp in type I latency, and to Cp in type III latency. Mutations in the CTCF binding site located at Qp disrupt the interaction between OriP and Qp and lead to the activation of Cp transcription. Mutation of the CTCF binding site between OriP and Cp, as well as siRNA of CTCF, eliminates both OriP-associated loops (Tempera *et al*., 2011). It is possible that the interaction between the two CTCF sites brings Qp close to OriP and compete out Cp, while Cp could come adjacent to the enhancer without the interaction as a default structure. If this is true, it can explain the observation that binding of CTCF between OriP and Cp does not correlates with Cp activity (Salamon *et al*., 2009).

Understanding the role of CTCF in life cycle of viruses may provide targets for new antiviral therapies. It has been demonstrated that the expression pattern of EBV is dependent on the differentiation stage of the infected B cells (Babcock *et al*., 2000). It is possible that the same regulatory processes that control CTCF

occupancy or function in the cells also work in viruses. In addition to the two sites described previously, CTCF binds at other sites in the genome of EBV, not only around the regulatory elements of latent genes but also near lytic genes, suggesting that CTCF may contribute to virulence through additional chromatin interactions (Holdorf *et al*., 2011).

27.6 Insulators and disease

From the discussion in the previous sections it is obvious that insulators play a role in mediating intra- and inter-chromosomal interactions in order to regulate enhancer function and other aspects of transcription and nuclear biology. It is therefore likely that defects in insulator function will lead to abnormalities in gene expression and disease. Here, we discuss the direct involvement of insulators in various human syndromes.

27.6.1 Neurological diseases caused by alteration of CTCF function

CTCF plays a role in neurodegenerative disorders such as myotonic dystrophy (DM1), spinocerebellar ataxia 7 (SCA7), SCA2, dentatorubral-pallidoluysian atrophy (DRPLA), and Huntington's disease (HD). These diseases are associated with expansion of trinucleotide repeats in the affected region. CTCF sites are present at one or both sides of the CTG/CAG repeat-containing region and it has been suggested that CTCF can regulate gene expression in these loci through control of noncoding transcription (Filippova *et al*., 2001). In the *DM1* gene, the two CTCF sites flanking the CTG repeats can restrict the antisense transcription and the heterochromatin state of the repeats without affecting the nearby *DMPK* and *SIX5* genes, whose silencing contributes to myotonic myopathy, cataracts, and cardiac-conduction defects in the disease (Cho *et al*., 2005; Filippova *et al*., 2001). In the case of the *ataxin-7* gene, CTCF is required for the antisense transcription of noncoding RNA from SCAANT1 to repress sense transcription of *ataxin-7* (Sopher *et al*., 2011). In individuals affected by these diseases, the expansion of the repeats evicts CTCF, and leads to misregulation of the disease associated genes (Cho *et al*., 2005). CTCF can also regulate the stability of the trinucleotide repeats. In transgenic mice carrying a SCA7 genomic fragment with a CTCF binding site, mutation, or methylation of the DNA sequence at the binding site can promote repeat expansion (Libby *et al*., 2008). However, the mechanism by which CTCF affects the stability of the repeats is not clear. A recent report suggests that CTCF can contribute to repeat stability through regulation of DNA replication. CTCF sites located between a replication origin and the repeats can slow or pause the progression of the replication machinery and enable safe passage of the replication fork. Without CTCF, the procession of the replication fork allows for slippage, hairpin formation, fork reversal, and other replication errors (Cleary *et al*., 2010). Thus, the interaction between the repeat and CTCF binding may affect repeat instability and abnormal transcription of the locus.

In addition to neurodegenerative diseases, CTCF also affects other neurological processes. For example, CTCF binds at the promoter region of the *BDNF* gene, which is critical for neuronal function. Lowering levels of Nicotinamide Adenine Dinucleotide (NAD) in mouse primary cortical neurons increases the methylation at the locus. The methylation triggers the dissociation of CTCF and subsequent silencing of *BDNF* (Chang *et al*., 2010). This may explain how age or nutrition-associated reduction of NAD levels contribute to cognitive impairment. CTCF may also play a role in neurological processes through a more general mechanism. CTCF binds at the promoter of the N-acetylglucosaminyltransferase IX (*GnT-IX*) gene, which encodes a brain-specific glycosyltransferase that synthesizes branched O-mannose glycan. CTCF is required for the transcription of *GnT-IX*, as RNAi treatment of CTCF suppresses *GnT-IX* expression (Kizuka *et al*., 2011). Glycosylation is a major post translational modification for many proteins and plays an important role in neural plasticity in mammals, suggesting that CTCF may affect these processes by regulating the expression of the *GnT-IX* gene.

27.6.2 Imprinting-related diseases caused by changes in CTCF function

CTCF can regulate the expression and epigenetic features of imprinted genes both in *cis* and in *trans* through intra- or inter-chromosomal interactions as discussed previously. Almost all imprinted genes identified to date can be classified as regulators of embryonic growth, placental growth, or adult metabolism (Jelinic and Shaw, 2007). It is then not surprising that loss of imprinting (LOI) can lead to various human diseases including Beckwith-Wiedemann syndrome (BWS), Silver-Russell syndrome (SRS), Angelman's syndrome, immunodeficiency syndrome (ICF), Rett's syndrome, Albright hereditary osteodystrophy, and hydatidiform mole (Jelinic and Shaw, 2007).

Hyper- and hypomethylation at the *H19/Igf2* ICR result in reciprocal changes in CTCF binding, *H19/Igf2* expression, and the two contrasting growth disorders BWS and SRS. It was recently found that CTCF mediates opposing chromatin conformations at the BWS and SRS loci. In addition, the chromatin landscapes are also different among cells from control individuals and patients with the two syndromes. In lymphoblastoid cells from control individuals, the repressive marks H3K9me3 and H4K20me3 associate with the methylated paternal ICR allele whereas the bivalent H3K4me2/H3K27me3 mark, H3K9ac, and CTCF associate with the non-methylated maternal allele. In patient-derived cell lines, the maternal versus paternal asymmetric distribution of these epigenetic marks is lost. H3K9me3 and H4K20me3 become biallelic in BWS and H3K4me2, H3K27me3, and H3K9ac become biallelic in SRS (Nativio *et al*., 2011).

Although the role of CTCF on imprinting has only been studied in detail at the *H19/Igf2* locus, similar CTCF-associated mechanisms appear to also play a role in imprinting at the *Rasgrf1, DLK1/GTL2*, *Wsb1/Nf1*, and *KvDMR* loci (Fitzpatrick *et al*., 2007; Ling *et al*., 2006; Yoon *et al*., 2005). This suggests a widespread function of CTCF in the regulation of genomic imprinting and, as a consequence, the development of imprinting-associated human diseases.

27.6.3 Alteration of CTCF function results in cancer

Missregulation of imprinting can also lead to cancer. The earliest and most common alteration observed in human cancers is LOI, which has been documented in 100% of chronic myeloid leukemia, 80% of ovarian tumors, 70% of Wilm's tumors, 66% of colorectal cancer, 56% of Barrett's esophagus, 50% of renal-cell carcinomas, 50% of esophageal cancer, 47–85% of lung adenocarcinomas, and 30% of meningiomas (Fitzpatrick *et al*., 2007; Ling *et al*., 2006; Yoon *et al*., 2005). In addition to causing cancer due to its effects on imprinting, alterations in CTCF function can also lead to cancer due to its direct regulatory role on oncogenes and tumor suppressor genes.

CTCF affects the expression of genes such as human telomerase reverse transcriptase (*hTERT*) or *myc* that regulate cell cycle processes important for cell growth, differentiation, and apoptosis. An appropriate balance between these processes is important for normal development, whereas an imbalance can lead to tumor development. hTERT, a catalytic subunit and limiting factor for telomerase activity, which is important during the cell cycle. CTCF binds to sequences of the *hTERT* gene located in the first two exons and represses its expression. CTCF only binds to *hTERT* in cells where the gene is not transcribed but not in telomerase-expressing cells. Knockdown of CTCF derepresses *hTERT* gene expression in normal telomerase-negative cells (Renaud *et al*., 2005). *hTERT* is a general cancer susceptibility locus, and its abnormal expression may disturb appropriate control of the cell cycle and induce aberrant cell growth (Johnatty *et al*., 2010). CTCF also binds constitutively to two regulatory elements close to the *c-myc* P2 promoter, the c-*myc* insulator element (MINE or CTCF-N) located upstream of the P2 promoter, and the CTCF-A site located immediately downstream of the P2 promoter (Filippova *et al*., 1996; Gombert *et al*., 2003; Gombert and Krumm, 2009). Binding of CTCF can affect the expression of *c-myc*, although the molecular mechanism underlying this effect is not well understood. The MINE element at the *c-myc* promoter has been shown to have CTCF-dependent enhancer blocking activity and CTCF knockdown

leads to an increase in c-myc protein levels (Gombert *et al*., 2003; Gombert and Krumm, 2009; Torrano *et al*., 2005). The c-myc protein is important for cells during the transition from proliferation to differentiation. Ectopic expression of CTCF in K562 cells leads to growth retardation and promotion of differentiation associated with increased *c-myc* expression, whereas CTCF knockdown significantly inhibits differentiation with reduced *c-myc* transcription (Torrano *et al*., 2005). Thus, the loss of CTCF function may disturb the balance between cell growth, differentiation, and apoptosis due to its effects on the expression of cell growth regulators.

The expression of other tumor suppressor genes, such as the *Retinoblastoma* gene (*Rb*), also depends on CTCF. Mutations and deletions of the *Rb* gene have been associated with a number of inherited malignancies. CTCF binds in vitro and in vivo to the human *Rb* promoter and this binding is required for *Rb* expression. When the CTCF binding site is removed or mutated, expression of a reporter gene decreases (De La Rosa-Velázquez *et al*., 2007). Some other genes encoding regulators of the cell cycle, such as *p19(ARF)*, *p16(INK4a)*, *PLK*, *BRCA1*, *p53*, and *p27* are also growth suppressors frequently silenced in cancer whose expression is also controlled by CTCF (Filippova, 2007). For example, *BRCA1* is expressed in normal cells but it is silenced in some cancer cells. CTCF binds at the promoter region of *BRCA1* only in expressing cells but not in tumor cells in which *BRCA1* is silenced (Darci *et al*., 2004; Xu *et al*., 2010). The *INK4B–ARF–INK4A* (*INK/ARF*) locus contains three tumor suppressor genes that are kept silenced by DNA methylation in different types of cancer. The *p16 (INK4a)* tumor suppressor gene is a frequent target of epigenetic inactivation in cancers such as breast, lung, colorectal, and multiple myeloma. CTCF binds upstream of the *p16 (INK4a)* promoter and the absence of binding is associated with silencing of *p16 (INK4a)* expression in breast cancer and multiple myeloma cells. Moreover, ablation of CTCF protein function from *p16 (INK4a)*-expressing cells by shRNA results in epigenetic changes in the *p16 (INK4a)* promoter and loss of transcription (Rodriguez *et al*., 2010; Witcher and Emerson, 2009). Conditional over-expression of CTCF in B cells enhances expression of *p27*, *p21*, *p53*, and *p19 (ARF)*, followed by inhibition of cell growth and induction of apoptosis, while knockdown of CTCF results in inhibition of these genes (Qi *et al*., 2003). These results suggest that loss of CTCF function can lead to the silencing of growth suppressor genes and contribute to carcinogenesis. Whether these effects are due to changes in the three-dimensional organization of the DNA or more local effects on transcription is unclear at this time.

27.6.4 Other mechanism that alter CTCF function also lead to disease states

Alterations of CTCF function that lead to disease and cancer can have both genetic and epigenetic origins. The CTCF gene maps to the cancer-associated human chromosome locus 16q22.1, which is the smallest overlap region of a variety of deletions found in breast, prostate, ovarian, and Wilm's tumors. Since CTCF null mutations are lethal, most CTCF mutations involved in human disease appear to cluster in the zinc finger domain of the protein. Instead of causing a complete loss of function, mutations in this domain alter the binding ability of CTCF. For example, one of the mutations identified abrogates CTCF binding to a subset of target sites in certain genes involved in the regulation of cell proliferation (*c-myc*, *ARF*, *PIM1*, *PLK*, and *Igf2*) but does not alter binding to other regions of the genome, including the β-globin insulator, the lysozyme silencer, or the APP promoter (Filippova *et al*., 2002). The selective loss of CTCF function caused by mutations in the zinc finger region may be tolerated because it does not affect cell viability, but results instead in transformation to a malignant phenotype. Since only selective changes of CTCF function can be tolerated in cells, it is not surprising that CTCF mutations are infrequently discovered in cancer or disease states. Instead, it is possible that viable mutations leading to changes of CTCF function could affect the regulatory roles of this protein rather than its structure. These alterations could be epigenetic, gene specific, and relatively tolerable. Most CTCF-mediated cancers may occur through

aberrant methylation of CTCF binding sites, which results in loss of CTCF binding to the DNA, as is the case in the imprinting related diseases. This alteration of methylation at CTCF sites has also been observed at non-imprinted genes. For example, in mice lung tumors, CpG sites harboring CTCF binding sequences are hypermethylated at the *INK/ARF* locus, resulting in the absence of CTCF binding and reduced expression of the tumor suppressor genes at this locus (Rodriguez *et al*., 2010; Tam *et al*., 2003). In some human tumors, the promoter of the *BRCA1* gene is methylated, CTCF is evicted, and *BRCA1* is silenced (Darci *et al*., 2004; Xu *et al*., 2010).

In addition to DNA methylation, other insulator regulatory mechanisms discussed above could lead the occurrence of cancer and disease. For example, testicular and ovarian tumor cells exhibit the same methylation profile as normal cells, but the BORIS protein, which recognizes the same DNA binding sites as CTCF, is expressed abnormally. Both CTCF and BORIS bind to DNA in cancer cells but BORIS inhibits the function of CTCF and leads to expression of *hTERT* (Renaud *et al*., 2010). Telomerase activity is not detectable in most somatic cells of adult humans but is found in highly proliferative cells, such as germ cells and stem cells, and 85–95% of cancers. The expression of BORIS in normal cells is sufficient to allow *hTERT* transcription and to extend their lifespan (Renaud *et al*., 2010). Reciprocal binding of CTCF and BORIS has also been observed at the *NY-ESO-1* promoter and leads to derepression of this gene in lung tumors (Hong *et al*., 2005). BORIS also binds to the *Rb2/p130* promoter in H69 lung cancer cells but not normal MRC-5 lung fibroblasts cells. Ectopic overexpression of BORIS in MRC-5 reduces *Rb2/p130* expression (Fiorentino *et al*., 2011). BORIS is aberrantly expressed in 71% (41 of 58 cases) of breast tumors. High levels of BORIS correlate with high levels of progesterone receptor (PR) and ER. The link between BORIS and PR/ER was further confirmed by the ability of BORIS to activate the promoters of the *PR* and *ER* genes in reporter gene assays (D'Arcy *et al*., 2008). Defects in other regulatory mechanisms that alter CTCF function can also lead to cancer. For example, alterations in the poly(ADP-ribosyl)ation pathway results in the absence of CTCF poly(ADP-ribosyl)ation and reduction of *p16 (INK4a)* and *Rassf1A* expression (Witcher and Emerson, 2009).

Understanding of the different mechanisms that alter CTCF function has opened new possibilities in the design of treatments for cancer or other diseases. For example, Trichostatin A (TSA), an inhibitor of histone deacetylase activity, is a well-known antitumor agent that effectively and selectively induces arrest of tumor growth and apoptosis. *hTERT* appears to be one of the primary targets for TSA-induced apoptosis in cancer cells. TSA induces demethylation of CpGs present in the binding site of CTCF on the *hTERT* promoter, leading to its repression (Choi *et al*., 2010). CTCF can contribute to the regulation of a variety of genes whose proper expression is required for normal cell differentiation. Both genetic and epigenetic changes of CTCF function can lead to the miss-expression of these genes, resulting in the development of a malignant phenotype.

27.7 Concluding remarks

Insulators mediate intra- and inter-chromosomal interactions to facilitate regulation of gene expression. By facilitating physical contacts between distant regulatory sequences, insulators can affect a variety of nuclear processes that range between V(D)J recombination and transcription regulation. For example, insulators can prevent enhancers from acting on specific promoters or they can instruct specific enhancers to interact with the promoters of particular genes. Insulators can also establish chromatin domains that compartmentalize different histone modifications in specific regions of the genome. By virtue of these properties, insulators can affect epigenetic information encoded in the form of transcription memory. Since this epigenetic information responds to signals coming from the outside of the cell, insulators could be important regulatory elements that establish and/or maintain patterns of epigenetic information. The pattern

of nuclear organization determined by physical contacts among insulator elements may be both a cause and a consequent of all other forms of epigenetic memory. As our understanding of the role of these sequences in nuclear processes increases, it is possible that insulators become important targets to manipulate the epigenetic status of cells in health and disease.

Acknowledgments

Work in the authors' laboratory is supported by U.S. Public Health Service Award GM35463 from the National Institutes of Health.

References

Babcock, G.J., Hochberg, D. and Thorley-Lawson, D.A. (2000) The expression pattern of epstein-barr virus latent genes in vivo is dependent upon the differentiation stage of the infected B cell. *Immunity*, **13**, 497–506.

Bantignies, F.D.R., Grimaud, C., Lavrov, S. *et al*. (2011) Polycomb-dependent regulatory contacts between distant hox loci in Drosophila. *Cell*, **144**, 214–226.

Bantignies, F.D.R., Roure, V., Comet, I. *et al*. (2003) Inheritance of polycomb-dependent chromosomal interactions in Drosophila. *Genes Dev.*, **17**, 2406–2420.

Barski, A., Cuddapah, S., Cui, K. *et al*. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823–837.

Bartkuhn, M., Straub, T., Herold, M. *et al*. (2009) Active promoters and insulators are marked by the centrosomal protein 190. *EMBO J.*, **28**, 877–888.

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

Belozerov, V.E., Majumder, P., Shen, P. and Cai, H.N. *et al*. (2003) A novel boundary element may facilitate independent gene regulation in the Antennapedia complex of Drosophila. *EMBO J.*, **22**, 3113–3121.

Buchner, K., Roth, P., Schotta, G. *et al*. (2000) Genetic and molecular complexity of the position effect variegation modifier mod(mdg4) in Drosophila. *Genetics*, **155**, 141–157.

Bushey, A.M., Ramos, E. and Corces, V.G. *et al*. (2009) Three subclasses of a Drosophila insulator show distinct and cell type-specific genomic distributions. *Genes Dev.*, **23**, 1338–1350.

Byrd, K. and Corces, V.G. (2003) Visualization of chromatin domains created by the gypsy insulator of Drosophila. *J. Cell Biol.*, **162**, 565–574.

Capelson, M. and Corces, V.G. (2005) The ubiquitin ligase dTopors directs the nuclear organization of a chromatin insulator. *Mol. Cell*, **20**, 105–116.

Capelson, M. and Corces, V.G. (2006) SUMO conjugation attenuates the activity of the gypsy chromatin insulator. *EMBO J.*, **25**, 1906–1914.

Chan, C.S. and Song, J.S. (2008) CCCTC-Binding factor confines the distal action of estrogen receptor. *Cancer Res.*, **68**, 9041–9049.

Chang, J., Zhang, B., Heath, H. *et al*. (2010) Nicotinamide adenine dinucleotide (NAD)-regulated DNA methylation alters CCCTC-binding factor (CTCF)/cohesin binding and transcription at the BDNF locus. *Proc. Natl. Acad. Sci.*, **107**, 21836–21841.

Chau, C.M., Zhang, X.-Y., McMahon, S.B. and Lieberman P.M. *et al*. (2006) Regulation of epstein-barr virus latency type by the chromatin boundary factor CTCF. *J. Virol.*, **80**, 5723–5732.

Chen, Q., Lin, L., Smith, S. *et al*. (2007) CTCF-dependent chromatin boundary element between the latency-associated transcript and ICP0 promoters in the herpes simplex virus type 1 genome. *J. Virol.*, **81**, 5192–5201.

Chen, X., Xu, H., Yuan, P. *et al*. (2008) Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell*, **133**, 1106–1117.

Chernukhin, I.V., Shamsuddin, S., Kang, S.Y. *et al*. (2000) Physical and functional interaction between two pluripotent proteins, the Y-box DNA/RNA-binding factor, YB-1, and the multivalent zinc finger factor, CTCF. *J. Biol. Chem.*, **275**, 29915–29921.

Chernukhin, I., Shamsuddin, S., Robinson, A.F. *et al*. (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.

Cho, D.H., Thienes, C.P., Mahoney, S.E. *et al*. (2005) Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Mol. Cell*, **20**, 483–489.

Choi, J.H., Min, N.Y., Park, J. *et al*. (2010) TSA-induced DNMT1 down-regulation represses hTERT expression via recruiting CTCF into demethylated core promoter region of hTERT in HCT 116. *Biochem. Biophys. Res. Commun.*, **391**, 449–454.

Cleard, F., Moshkin, Y., Karch, F. and Maeda, R.K. *et al*. (2006) Probing long-distance regulatory interactions in the Drosophila melanogaster bithorax complex using Dam identification. *Nat. Genet.*, **38**, 931–935.

Cleary, J.D., Tome, S., Lopez-Castel, A. *et al*. (2010) Tissue-and age-specific DNA replication patterns at the CTG/CAG-expanded human myotonic dystrophy type 1 locus. *Nat. Struct. Mol. Biol.*, **17**, 1079–1087.

Comet, I., Schuettengruber, B., Sexton, T. and Cavalli, G. (2011) A chromatin insulator driving three-dimensional Polycomb response element (PRE) contacts and Polycomb association with the chromatin fiber. *Proc. Natl. Acad. Sci.*, **108**, 2294–2299.

Cuddapah, S., Jothi, R., Schones, D.E. *et al*. (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.

Darci, T.B., Debora, N.M.-D., Trevor K.A. and David, I.R. *et al*. (2004) DNA binding sites for putative methylation boundaries in the unmethylated region of the *BRCA* promoter. *Int. J. Cancer*, **111**, 669–678.

D'Arcy, V., Pore, N., Docquier, F. *et al*. (2008) BORIS, a paralogue of the transcription factor, CTCF, is aberrantly expressed in breast tumours. *Br. J. Cancer*, **98**, 571–579.

Defossez, P.A., Kelly, K.F., Filion, G.J.P. *et al*. (2005) The human enhancer blocker CTC-binding factor interacts with the transcription factor Kaiso. *J. Biol. Chem.*, **280**, 43017–43023.

Degner, S.C., Verma-Gaur, J., Wong, T.P. *et al*. (2009) Cutting edge: developmental stage-specific recruitment of cohesin to CTCF sites throughout immunoglobulin loci during B lymphocyte development. *J. Immunol.*, **182**, 44–48.

Degner, S.C., Wong, T.P., Jankevicius, G. and Feeney, A.J. *et al*. (2011) CCCTC-binding factor (CTCF) and cohesin influence the genomic architecture of the Igh locus and antisense transcription in pro-B cells. *Proc. Natl. Acad. Sci.*, **108**, 9566–9571.

Dekker, J., Rippe, K., Dekker, M. and Kleckner, N. *et al*. (2002) Capturing chromosome conformation. *Science*, **295**, 1306–1311.

De La Rosa-Velázquez, I.A., Rincón-Arano, H., Benítez-Bribiesca, L. and Recillas-Targa, F. *et al*. (2007) Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer Res.*, **67**, 2577–2585.

Donohoe, M.E., Silva, S.S., Pinter, S.F. *et al*. (2009) The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. *Nature*, **460**, 128–132.

Donohoe, M.E., Zhang, L.-F., Xu, N. *et al*. (2007) Identification of a Ctcf cofactor, Yy1, for the X chromosome binary switch. *Mol. Cell*, **25**, 43–56.

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

Emberly, E., Blattes, R., Schuettengruber, B. *et al*. (2008) BEAF regulates cell-cycle genes through the controlled deposition of H3K9 methylation marks into its conserved *Dual-Core* binding sites. *PLoS Biol.*, **6**, e327.

Farrar, D., Rai, S., Chernukhin, I. *et al*. (2010) Mutational analysis of the poly(ADP-Ribosyl)ation sites of the transcription factor CTCF provides an insight into the mechanism of its regulation by poly(ADP-Ribosyl)ation. *Mol. Cell Biol.*, **30**, 1199–1216.

Farrell, C.M., Grinberg, A., Huang, S.P. *et al*. (2000) A large upstream region is not necessary for gene expression or hypersensitive site formation at the mouse β-globin locus. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 14554–14559.

Ferraiuolo, M.A., Rousseau, M., Miyamoto, C. *et al*. (2010) The three-dimensional architecture of Hox cluster silencing. *Nucleic Acids Res.*, **38**, 7472–7484.

Filippova, G.N. (2007) Genetics and epigenetics of the multifunctional protein CTCF, in *Current Topics in Developmental Biology*, vol. **80** (ed. P.S. Gerald), Academic Press, pp. 337–360.

Filippova, G.N., Fagerlie, S., Klenova, E.M. *et al*. (1996) An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. *Mol. Cell Biol.*, **16**, 2802–2813.

Filippova, G.N., Qi, C.-F., Ulmer, J.E. *et al*. (2002) Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter its DNA-binding specificity. *Cancer Res.*, **62**, 48–52.

Filippova, G.N., Thienes, C.P., Penn, B.H. *et al*. (2001) CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nat. Genet.*, **28**, 335–343.

Fiorentino, F.P., Macaluso, M., Miranda, F. *et al*. (2011) CTCF and BORIS regulate Rb2/p130 gene transcription: a novel mechanism and a new paradigm for understanding the biology of lung cancer. *Mol. Cancer Res.*, **9**, 225–233.

Fitzpatrick, G.V., Pugacheva, E.M., Shin, J.-Y. *et al*. (2007) Allele-specific binding of CTCF to the multipartite imprinting control region KvDMR1. *Mol. Cell Biol.*, **27**, 2636–2647.

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

Gause, M., Morcillo, P. and Dorsett, D. (2001) Insulation of enhancer-promoter communication by a gypsy transposon insert in the Drosophila cut gene: cooperation between suppressor of hairy-wing and modifier of mdg4 proteins. *Mol. Cell Biol.*, **21**, 4807–4817.

Gerasimova, T.I., Byrd, K. and Corces, V.G. *et al*. (2000) A chromatin insulator determines the nuclear localization of DNA. *Mol. Cell*, **6**, 1025–1035.

Gerasimova, T.I., Gdula, D.A., Gerasimov, D.V. *et al*. (1995) A Drosophila protein that imparts directionality on a chromatin insulator is an enhancer of position-effect variegation. *Cell*, **82**, 587–597.

Gerasimova, T.I., Lei, E.P., Bushey, A.M. and Corces, V.G. (2007) Coordinated control of dCTCF and gypsy chromatin insulators in Drosophila. *Mol. Cell*, **28**, 761–772.

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

Ghosh, D., Gerasimova, T.I. and Corces, V.G. (2001) Interactions between the Su(Hw) and Mod(mdg4) proteins required for gypsy insulator function. *EMBO J.*, **20**, 2518–2527.

Gombert, W.M., Farris, S.D., Rubio, E.D. *et al*. (2003) The c-myc insulator element and matrix attachment regions define the c-myc chromosomal domain. *Mol. Cell Biol.*, **23**, 9338–9348.

Gombert, W.M. and Krumm, A. (2009) Targeted deletion of multiple CTCF-binding elements in the human C-MYC gene reveals a requirement for CTCF in C-MYC expression. *PLoS ONE*, **4**, e6109.

Guastafierro, T., Cecchinelli, B., Zampieri, M. *et al*. (2008) CCCTC-binding factor activates PARP-1 Affecting DNA methylation machinery. *J. Biol. Chem.*, **283**, 21873–21880.

Guelen, L., Pagie, L., Brasset, E. *et al*. (2008) Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*, **453**, 948–951.

Gyurkovics, H., Gausz, J., Kummer, J. and Karch, F. (1990) A new homeotic mutation in the Drosophila bithorax complex removes a boundary separating two domains of regulation. *EMBO J.*, **9**, 2579–2585.

Hadjur, S., Williams, L.M., Ryan, N.K. *et al*. (2009) Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. *Nature*, **460**, 410–413.

Handoko, L., Xu, H., Li, G. *et al*. (2011) CTCF-mediated functional chromatin interactome in pluripotent cells. *Nat. Genet.*, **43**, 630–638.

Hark, A.T., Schoenherr, C.J., Katz, D.J. *et al*. (2000) CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature*, **405**, 486–489.

Harrison, D.A., Gdula, D.A., Coyne, R.S. and Corces, V.G. (1993) A leucine zipper domain of the suppressor of Hairy-wing protein mediates its repressive effect on enhancer function. *Genes Dev.*, **7**, 1966–1978.

Hart, C.M., Cuvier, O. and Laemmli, U.K. (1999) Evidence for an antagonistic relationship between the boundary element- associated factor BEAF and the transcription factor DREF. *Chromosoma*, **108**, 375–383.

Heath, H., de Almeida, C.R., Sleutels, F. *et al*. (2008) CTCF regulates cell cycle progression of [alpha][beta] T cells in the thymus. *EMBO J.*, **27**, 2839–2850.

Hikichi, T., Kohda, T., Kaneko-Ishino, T., *et al*. (2003) Imprinting regulation of the murine Meg1/Grb10 and human GRB10 genes; roles of brain-specific promoters and mouse-specific CTCF-binding sites. *Nucleic Acids Res.*, **31**, 1398–1406.

Holdorf, M.M., Cooper, S.B., Yamamoto, K.R. and Miranda, J.L. *et al*. (2011) Occupancy of chromatin organizers in the epstein-barr virus genome. *Virology*, **415**, 1–5.

Holohan, E.E., Kwong, C., Adryan, B. *et al*. (2007) CTCF genomic binding sites in *Drosophila* and the organisation of the bithorax complex. *PLoS Genet.*, **3**, e112.

Hong, J.A., Kang, Y., Abdullaev, Z. *et al*. (2005) Reciprocal binding of CTCF and BORIS to the NY-ESO-1 promoter coincides with derepression of this cancer-testis gene in lung cancer cells. *Cancer Res.*, **65**, 7763–7774.

Hoover, K.K., Chien, A.J. and Corces, V.G. (1993) Effects of transposable elements on the expression of the forked gene of Drosophila melanogaster. *Genetics*, **135**, 507–526.

Hou, C., Dale, R. and DeanA. (2010) Cell type specificity of chromatin organization mediated by CTCF and cohesin. *Proc. Natl. Acad. Sci.*, **107**, 3651–3656.

Ishihara, K., Oshimura, M. and Nakao, M. (2006) CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Mol. Cell*, **23**, 733–742.

Jelinic, P., Stehle, J.-C. and Shaw, P. (2006) The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. *PLoS Biol.*, **4**, e355.

Jelinic, P. and Shaw, P. (2007) Loss of imprinting and cancer. *J. Pathol.*, **211**, 261–268.

Jiang, N., Emberly, E., Cuvier, O. and Hart, C.M. (2009) Genome-wide mapping of boundary element-associated factor (BEAF) binding sites in Drosophila melanogaster Links BEAF to transcription. *Mol. Cell Biol.*, **29**, 3556–3568.

Johnatty, S.E., Beesley, J., Chen, X. *et al*. (2010) Evaluation of candidate stromal epithelial cross-talk genes identifies association between risk of serous ovarian cancer and *TERT*, a cancer susceptibility 'hot-spot'. *PLoS Genet.*, **6**, e1001016.

Jothi, R., Cuddapah, S., Barski, A. *et al*. (2008) Genome-wide identification of in vivo protein–DNA binding sites from ChIP-Seq data. *Nucleic Acids Res.*, **36**, 5221–5231.

Kagey, M.H., Newman, J.J., Bilodeau, S. *et al*. (2011) Mediator and cohesin connect gene expression and chromatin architecture. *Nature*, **467**, 430–435.

Kang, Y., Hong, J.A., Chen, G.A. *et al*. (2007) Dynamic transcriptional regulatory complexes including BORIS, CTCF and Sp1 modulate NY-ESO-1 expression in lung cancer cells. *Oncogene*, **26**, 4394–4403.

Karch, F., Galloni, M., Sipos, L. *et al*. (1994) Mcp and Fab-7: molecular analysis of putative boundaries of *cis*-regulatory domains in the bithorax complex of *Drosophila* melanogaster. *Nucleic Acids Res.*, **22**, 3138–3146.

Kellum, R. and Schedl, P. (1992) A group of scs elements function as domain boundaries in an enhancer-blocking assay. *Mol. Cell Biol.*, **12**, 2424–2431.

Kim, T.H., Abdullaev, Z.K., Smith, A.D. *et al*. (2007) Analysis of the vertebrate insulator protein CTCF-binding sites in the human genome. *Cell*, **128**, 1231–1245.

Kim, Y.J., Cecchini, K.R. and KimT.H. (2011) Conserved, developmentally regulated mechanism couples chromosomal looping and heterochromatin barrier activity at the homeobox gene A locus. *Proc. Natl. Acad. Sci.*, **108**, 7391–7396.

Kizuka, Y., Kitazume, S., Yoshida, M. and Taniguchi, N. (2011) Brain-specific expression of N-acetylglucosaminyltransferase IX (GnT-IX) is regulated by epigenetic histone modifications. *J. Biol. Chem.*, **286**, 31875–31884.

Klenova, E.M., Chernukhin, I.V., El-Kady, A. *et al*. (2001) Functional phosphorylation sites in the C-terminal region of the multivalent multifunctional transcriptional factor CTCF. *Mol. Cell Biol.*, **21**, 2221–2234.

Kurukuti, S., Tiwari, V.K., Tavoosidana, G. *et al*. (2006) CTCF binding at the H19 imprinting control region mediates maternally inherited higher-order chromatin conformation to restrict enhancer access to Igf2. *Proc. Natl. Acad. Sci.*, **103**, 10684–10689.

Kyrchanova, O., Toshchakov, S., Podstreshnaya, Y. *et al*. (2008) Functional interaction between the Fab-7 and Fab-8 Boundaries and the upstream promoter region in the *Drosophila* Abd-B gene. *Mol. Cell Biol.*, **28**, 4188–4195.

Labrador, M. and Corces, V.G. (2003) Extensive exon reshuffling over evolutionary time coupled to trans-splicing in Drosophila. *Genome Res.*, **13**, 2220–2228.

Labrador, M., Mongelard, F., Baxter, E.M. *et al*. (2001) Protein encoding by both DNA strands. *Nature*, **409**, 1000.

Lanzuolo, C., Roure, V., Dekker, J. *et al*. (2007) Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. *Nat. Cell Biol.*, **9**, 1167–1174.

Lei, E.P. and Corces, V.G. (2006) RNA interference machinery influences the nuclear organization of a chromatin insulator. *Nat. Genet.*, **38**, 936–941.

Li, H.B., Müller, M., Bahechar, I.A. *et al*. (2011) Insulators, not polycomb response elements, are required for long-range interactions between polycomb targets in Drosophila melanogaster. *Mol. Cell Biol.*, **31**, 616–625.

Libby, R.T., Hagerman, K.A., Pineda, V.V. *et al*. (2008) CTCF *cis*-regulates trinucleotide repeat instability in an epigenetic manner: a novel basis for mutational hot spot determination. *PLoS Genet.*, **4**, e1000257.

Ling, J.Q., Li, T., Hu, J.F. *et al*. (2006) CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science*, **312**, 269–272.

Loukinov, D.I., Pugacheva, E., Vatolin, S., *et al*. (2002) BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. *Proc. Natl. Acad. Sci. U.S.A.*, **99**, 6806–6811.

Lutz, M., Burke, L.J., Barreto, G. *et al*. (2000) Transcriptional repression by the insulator protein CTCF involves histone deacetylases. *Nucleic Acids Res.*, **28**, 1707–1713.

Lutz, M., Burke, L.J., LeFevre, P. *et al*. (2003) Thyroid hormone-regulated enhancer blocking: cooperation of CTCF and thyroid hormone receptor. *EMBO J.*, **22**, 1579–1587.

Majumder, P. and Boss, J.M. (2011) DNA methylation dysregulates and silences the HLA-DQ locus by altering chromatin architecture. *Genes Immunol.*, **12**, 291–299.

Majumder, P., Gomez, 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.

Martowicz, M.L., Meyer, M.B. and Pike, J.W. (2011) The mouse RANKL gene locus is defined by a broad pattern of histone H4 acetylation and regulated through distinct distal enhancers. *J. Cell. Biochem.*, **112**, 2030–2045.

Meguro-Horike, M., Yasui, D.H., Powell, W. *et al*. (2011) Neuron-specific impairment of inter-chromosomal pairing and transcription in a novel model of human 15q-duplication syndrome. *Hum. Mol. Genet.*, **20**, 3798–3810.

Melnikova, L., Juge, F., Gruzdeva, N. *et al*. (2004) Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 14806–14811.

Mishiro, T., Ishihara, K., Hino, S. *et al*. (2009) Architectural roles of multiple chromatin insulators at the human apolipoprotein gene cluster. *EMBO J.*, **28**, 1234–1245.

Mohan, M., Bartkuhn, M., Herold, M. *et al*. (2007) The Drosophila insulator proteins CTCF and CP190 link enhancer blocking to body patterning. *EMBO J.*, **26**, 4203–4214.

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.

Nativio, R., Heeson, S. and Reik, W. *et al*. (2011) Disruption of genomic neighbourhood at the imprinted IGF2-H19 locus in Beckwith–Wiedemann syndrome and Silver–Russell syndrome. *Hum. Mol. Genet.*, **20**, 1363–1374.

Nativio, R., Wendt, K.S., Ito, Y. *et al*. (2009) Cohesin is required for higher-order chromatin conformation at the imprinted *IGF2-H19* locus. *PLoS Genet.*, **5**, e1000739.

Nègre, N., Brown, C.D., Shah, P.K. *et al*. (2010) A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS Genet.*, **6**, e1000814.

Nguyen, P., Cui, H., Bisht, K.S. *et al*. (2008) CTCFL/BORIS is a methylation-independent DNA-binding protein that preferentially binds to the paternal H19 differentially methylated region. *Cancer Res.*, **68**, 5546–5551.

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.

Oliver, D., Sheehan, B., South, H. *et al*. (2011) The chromosomal association/dissociation of the chromatin insulator protein Cp190 of Drosophila melanogaster is mediated by the BTB/POZ domain and two acidic regions. *BMC Cell Biol.*, **11**, 101.

Ottaviani, A., Rival-Gervier, S., Boussouar, A. *et al*. (2009a) The *D4Z4* Macrosatellite repeat acts as a CTCF and A-type lamins-dependent insulator in facio-scapulo-humeral dystrophy. *PLoS Genet.*, **5**, e1000394.

Ottaviani, A., Schluth-Bolard, C., Rival-Gervier, S. *et al*. (2009b) Identification of a perinuclear positioning element in human subtelomeres that requires A-type lamins and CTCF. *EMBO J.*, **28**, 2428–2436.

Pai, C.Y., Lei, E.P., Ghosh, D. and Corces, V.G. (2004) The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol. Cell*, **16**, 737–748.

Pant, V., Mariano, P., Kanduri, C. *et al*. (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.

Parelho, V., Hadjur, S., Spivakov, M. *et al*. (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell*, **132**, 422–433.

Pathak, R.U., Rangaraj, N., Kallappagoudar, S. *et al*. (2007) Boundary element-associated factor 32B connects chromatin domains to the nuclear matrix. *Mol. Cell Biol.*, **27**, 4796–4806.

Phillips, J.E. and Corces, V.G. (2009) CTCF: master weaver of the genome. *Cell*, **137**, 1194–1211.

Qi, C.F., Martensson, A., Mattioli, M. *et al*. (2003) CTCF functions as a critical regulator of cell-cycle arrest and death after ligation of the B cell receptor on immature B cells. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 633–638.

Renaud, S., Loukinov, D., Alberti, L. *et al*. (2010) BORIS/CTCFL-mediated transcriptional regulation of the hTERT telomerase gene in testicular and ovarian tumor cells. *Nucleic Acids Res.*, **39**, 862–873.

Renaud, S., Loukinov, D., Bosman, F.T. *et al*. (2005) CTCF binds the proximal exonic region of hTERT and inhibits its transcription. *Nucleic Acids Res.*, **33**, 6850–6860.

Renaud, S., Pugacheva, E.M., Delgado, M.D. *et al*. (2007) Expression of the CTCF-paralogous cancer-testis gene, brother of the regulator of imprinted sites (BORIS), is regulated by three alternative promoters modulated by CpG methylation and by CTCF and p53 transcription factors. *Nucleic Acids Res.*, **35**, 7372–7388.

Rodriguez, C., Borgel, J., Court, F. *et al*. (2010) CTCF is a DNA methylation-sensitive positive regulator of the INK/ARF locus. *Biochem. Biophys. Res. Commun.*, **392**, 129–134.

Roy, S., Gilbert, M.K. and Hart, C.M. (2007) Characterization of BEAF mutations isolated by homologous recombination in Drosophila. *Genetics*, **176**, 801–813.

Rubio, E.D., Reiss, D.J., Welcsh, P.L. *et al*. (2008) CTCF physically links cohesin to chromatin. *Proc. Natl. Acad. Sci.*, **105**, 8309–8314.

Salamon, D., Banati, F., Koroknai, A. *et al*. (2009) Binding of CCCTC-binding factor in vivo to the region located between Rep* and the C promoter of Epstein-Barr virus is unaffected by CpG methylation and does not correlate with Cp activity. *J. Gen. Virol.*, **90**, 1183–1189.

Sandhu, K.S., Shi, C., Sjolinder, M. *et al*. (2009) Nonallelic transvection of multiple imprinted loci is organized by the H19 imprinting control region during germline development. *Genes Dev.*, **23**, 2598–2603.

Schweinsberg, S., Hagstrom, K., Gohl, D. *et al*. (2004) The enhancer-blocking activity of the Fab-7 boundary from the Drosophila bithorax complex requires GAGA-factor-binding sites. *Genetics*, **168**, 1371–1384.

Sekimata, M., Pérez-Melgosa, M., Miller, S.A. *et al*. (2009) CCCTC-binding factor and the transcription factor T-bet orchestrate T helper 1 cell-specific structure and function at the interferon-[gamma] locus. *Immunity*, **31**, 551–564.

Sopher, B.L., Ladd, P.D., Pineda, V.V. *et al*. (2011) CTCF regulates ataxin-7 expression through promotion of a convergently transcribed, antisense noncoding RNA. *Neuron*, **70**, 1071–1084.

Splinter, E., Heath, H., Kooren, J. *et al*. (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes Dev.*, **20**, 2349–2354.

Stedman, W., Kang, H., Lin, S. *et al*. (2008) Cohesins localize with CTCF at the KSHV latency control region and at cellular c-myc and H19/Igf2 insulators. *EMBO J.*, **27**, 654–666.

Tam, A.S., Devereux, T.R., Patel, A.C. *et al*. (2003) Perturbations of the Ink4a/Arf gene locus in aflatoxin B1-induced mouse lung tumors. *Carcinogenesis*, **24**, 121–132.

Tempera, I., Klichinsky, M. and Lieberman, P.M. (2011) EBV Latency types adopt alternative chromatin conformations. *PLoS Pathog.*, **7**, e1002180.

Tempera, I., Wiedmer, A., Dheekollu, J. and Lieberman, P.M. (2010) CTCF Prevents the epigenetic drift of EBV latency promoter Qp. *PLoS Pathog.*, **6**, e1001048.

Tolhuis, B., Palstra, R.-J., Splinter, E. *et al*. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell*, **10**, 1453–1465.

Torrano, V., Chernukhin, I., Docquier, F. *et al*. (2005) CTCF Regulates growth and erythroid differentiation of human myeloid leukemia cells. *J. Biol. Chem.*, **280**, 28152–28161.

Vazquez, J., Muller, M., Pirrotta, V. and Sedat, J.W. *et al*. (2006) The Mcp element mediates stable long-range chromosome-chromosome interactions in Drosophila. *Mol. Biol. Cell*, **17**, 2158–2165.

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

Wendt, K.S., Yoshida, K., Itoh, T. *et al*. (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature*, **451**, 796–801.

Weth, O. *et al*. (2010) Modular insulators: genome wide search for composite CTCF/Thyroid hormone receptor binding sites. *PLoS ONE*, **5**, e10119.

Witcher, M. and Emerson, B.M. (2009) Epigenetic silencing of the p16INK4a tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Mol. Cell*, **34**, 271–284.

Wood, A.M., Bortle, K.V., Ramos, E. *et al*. (2011) Regulation of chromatin organization and inducible gene expression by a Drosophila insulator. *Mol. Cell* **44**, 29-38.

Xiao, T., Wallace, J. and Felsenfeld, G. *et al*. (2011) Specific sites in the C terminus of CTCF interact with the SA2 subunit of the cohesin complex and are required for cohesin-dependent insulation activity. *Mol. Cell Biol.*, **31**, 2174–2183.

Xu, J., Huo, D., Chen, Y. *et al*. (2010) CpG island methylation affects accessibility of the proximal BRCA1 promoter to transcription factors. *Breast Cancer Res. Treat.*, **120**, 593–601.

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

Xu, Z., Wei, G., Chepelev, I. *et al*. (2011) Mapping of INS promoter interactions reveals its role in long-range regulation of SYT8 transcription. *Nat. Struct. Mol. Biol.*, **18**, 372–378.

Yao, H., Brick, K., Evrard, Y. *et al*. (2010) Mediation of CTCF transcriptional insulation by DEAD-box RNA-binding protein p68 and steroid receptor RNA activator SRA. *Genes Dev.*, **24**, 2543–2555.

Yoon, B., Herman, H., Hu, B. *et al*. (2005) Rasgrf1 imprinting is regulated by a CTCF-dependent methylation-sensitive enhancer blocker. *Mol. Cell Biol.*, **25**, 11184–11190.

Yu, W., Ginjala, V., Pant, V. *et al*. (2004) Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat. Genet.*, **36**, 1105–1110.

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.

Zhang, Y. Liang, J., Li, Y. *et al*. (2010) CCCTC-binding factor acts upstream of FOXA1 and demarcates the genomic response to estrogen. *J. Biol. Chem.*, **285**, 28604–28613.

Zhao, K., Hart, C.M. and Laemmli, U.K. (1995) Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell*, **81**, 879–889.

Zhao, Z., Tavoosidana, G., Sjölinder, M. *et al*. (2006) Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat. Genet.*, **38**, 1341–1347.