

Chromatin domain boundaries: insulators and beyond

Gong Hong WEI, De Pei LIU*, Chih Chuan LIANG

National Laboratory of Medical Molecular Biology, Institute of Basic Medical Sciences, Chinese Academy of Medical Sciences and Peking Union Medical College, 5 Dong Dan San Tiao, Beijing 100005, China

ABSTRACT

The eukaryotic genome is organized into functionally and structurally distinct domains, representing regulatory units for gene expression and chromosome behavior. DNA sequences that mark the border between adjacent domains are the insulators or boundary elements, which are required in maintenance of the function of different domains. Some insulators need others enable to play insulation activity. Chromatin domains are defined by distinct sets of post-translationally modified histones. Recent studies show that these histone modifications are also involved in establishment of sharp chromatin boundaries in order to prevent the spreading of distinct domains. Additionally, in some loci, the high-order chromatin structures for long-range looping interactions also have boundary activities, suggesting a correlation between insulators and chromatin loop domains. In this review, we will discuss recent progress in the field of chromatin domain boundaries.

Keywords: insulator, chromatin domain boundary, histone code, chromatin loop domain.

INTRODUCTION

Under high-resolution light or electron microscopy, the interior of the eukaryotic cell nucleus is structurally complex and non-homogeneous. Additionally, within the three-dimensional space of the cell nucleus chromosomes and genes are not randomly placed. Numerous studies have examined the organization of chromatin fibers and found that the eukaryotic genomes are partitioned into functionally autonomous domains in which gene expression is either repressed or facilitated. It is now apparent that chromatin structure plays an important role in regulating gene transcription by providing the proper local environment and binding sites for regulators to ensure proper spatial and temporal gene expression [1]. Proximal regulatory elements like promoters and certain distal enhancer element, such as enhancers and locus control regions (LCRs) are required to modulate individual genes within these domains [2, 3].

Mechanisms involved in delimiting independent structural and functional domains are poorly understood. It is unclear about the establishment and maintenance of gene

expression in response to specific chromatin domain antagonizing signals (either positive or negative) generated from surrounding chromatin regions. Additionally, it is not known about the mechanism of the distal enhancer-like elements in long-range chromatin regulatory interactions to ensure the proper expression of specific genes. A recent study by Gilbert *et al* through a global analysis of human chromatin fibre structure, found that the transitions between open and closed chromatin regions along a chromosome are sharp rather than gradual [4], which is consistent with the hypothesis that distinct boundaries may exist between these regions [4, 5]. The role of a specialized class of DNA element, known as insulators, is implicated in these processes.

Insulators or boundary elements are genetic elements near chromatin domain boundaries with distinct properties involving in gene expression alteration. First, they act as a barrier to shield the transgene from position effects to prevent the spread of repressive heterochromatin from one domain to the next (Fig. 1A) [5]. Secondly, they can prevent the communication between distant elements such as enhancers to influence gene expression, a function known as enhancer-blocking activity (Fig. 1B) [6]. Up to now, numerous insulators have been described in different organisms from yeast to mammals. Some possess both properties as an enhancer-blocker and a barrier [7]. Some in-

*Correspondence: De Pei LIU

Tel: +86-10-65296415 Fax: +86-10-65133086

Email: liudp@pumc.edu.cn

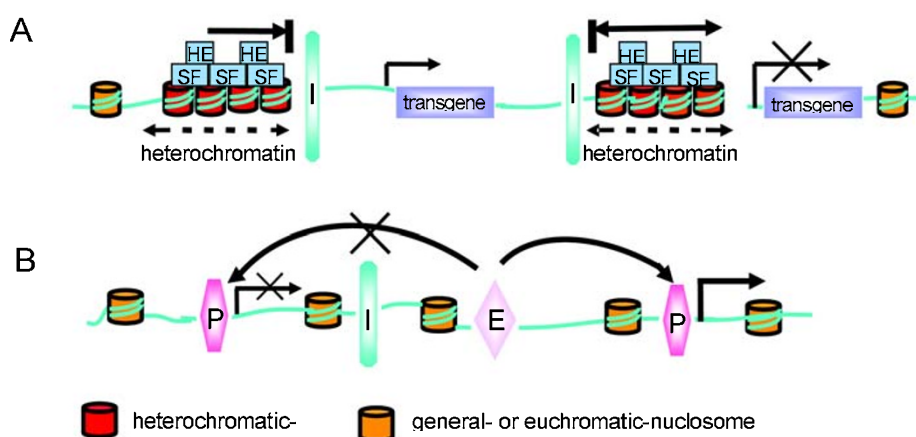


Fig. 1 Two basic properties of insulators. **(A)** Barrier activity. Heterochromatic domains are always established in some chromosomal regions through histone-modifying enzymes (HE) such as deacetylases and methyltransferases, which modify histone tails to create binding sites for silencing factors (SF). When a transgene stably integrates into these adjacent regions the spreading of the heterochromatin will influence the transgene expression. Flanking the transgene with two copies of an insulator (I) can protect it from such chromosomal position effect following stable integration. **(B)** Enhancer-blocking activity. An insulator efficiently blocks an enhancer's (E) action on a promoter (P) when placed between the two regulatory elements. The enhancer-mediated transcription of the promoter-directed downstream gene is repressed.

insulators contain separable activities [7]. Furthermore, histone modifications may play a pivotal role in barrier function as having been demonstrated that the 5'-HS4 insulator element of chicken β -globin can manipulate the histone code to form a barrier to prevent the spread of adjacent heterochromatin domain (Fig. 2) [8]. A recent significant progress shows that insulators can mediate the formation of independent chromatin loop domains through nuclear matrix or nuclear envelope tethering [9, 10]. These looped chromatin domains thus may have insulator function.

In this article, we will emphasize the recent progress on dynamic insulator elements and their associated proteins, especially the role of histone code in chromatin domain boundary activity and the correlations between chromatin insulating and looping. Based on the recent discoveries, we propose two models for the formation and function of higher order chromatin loop domains in mammalian β -globin loci.

CHROMATIN INSULATORS AND ASSOCIATED PROTEINS

Insulator elements were first discovered in *Drosophila* and subsequently found in various organisms including yeast, sea urchins, *Xenopus*, chickens, mice, and humans [6, 7, 11-13]. Insulators exert their functions through interacting with specific DNA-binding proteins and associated factors. Genetic and biochemical analyses have identified a large number of protein factors required for insu-

lator functions.

One of the best characterized *Drosophila* insulators is the Su(Hw) chromatin insulator, which was first identified within the genome of the gypsy retrotransposon. It contains a 340 bp sequence with 12 binding sites for DNA binding protein Su(Hw). Su(Hw) interacts with Mod(mdg) 2.2, a BTB/POZ domain protein, which is essential for Su(Hw) insulator function. When two Su(Hw) insulators were inserted between enhancer and promoter, the enhancer-blocking activity was neutralized [14, 15], suggesting that the two insulators interact, probably via the bound protein complexes to form a mini-loop to bring the enhancer and promoter into close proximity. There are accumulating evidence to support this hypothesis.

Using fluorescence *in situ* hybridization (FISH) technique, Byrd *et al* showed that DNA sequences located between two endogenous Su(Hw) insulators form loops attaching to the nuclear matrix. Insertion of a gypsy insulator into one of these loops results in the formation of two smaller loops. Functional insulator proteins are required for the formation of the loops [16]. Recently, Melnikova *et al* found that the insertion of GAGA factor (GAF) binding sites between the enhancer and the Su(Hw) insulator allows bypass of the insulator, suggesting that GAF and Mod(mdg4) proteins may interact with each other to promote loop formation, thus restore the communication between the upstream enhancer and the downstream reporter gene [17]. Another interesting study identified CP190 as a third

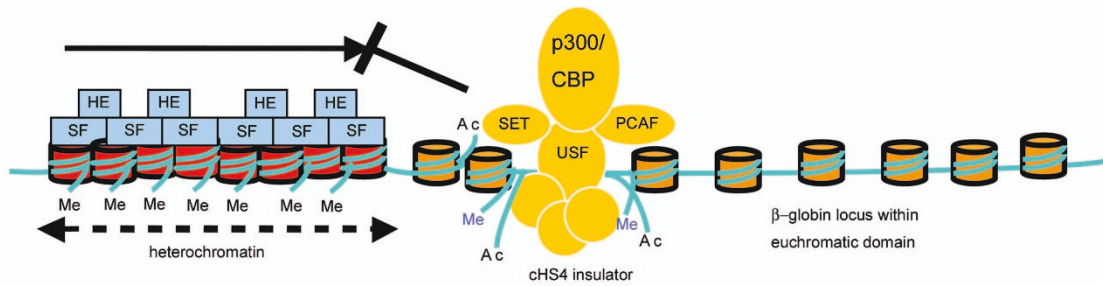


Fig. 2 A model for chromatin barrier function of insulator and the histone code. The cHS4 insulator binding proteins USF1/USF2 (USF) recruit the histone acetyltransferases (HATs) PCAF and P300/CBP (H3 and H4 acetylation) and methyltransferase SET7/9 (H3 lysine 4 methylation), which form a histone hyperacetylation- and H3 lysine 4 methylation-enriched chromatin boundary to block the spread of adjacent heterochromatin structure [8, 42]. The insulator perhaps recruit other putative protein factors to form a barrier complex to prevent heterochromatin spreading. Ac and Me are abbreviations of acetyl and methyl groups, respectively. Especially, Me (blue) mean methylation of H3 lysine 4, which are active chromatin marks. Other Me mean repressive chromatin marks such as methylation of H3 lysine 9.

component of the Su(Hw) insulator. CP190 contains both a BTB/POZ protein-protein interaction domain and three C2H2 zinc finger DNA binding domains [18]. The authors showed that CP190 can bind to both Su(Hw) and Mod (mdg)2.2. Thus a model was proposed that these insulator components form higher order chromatin structures through protein-DNA and protein-protein interactions to regulate gene expression by compartmentalizing genes into these transcriptionally independent domains [18]. Similar to Mod(mdg) and CP190, all GAFs contain BTB-domains at their N-termini. Some of the binding sites for GAF proteins are involved in insulator function [19]. These results suggest that interactions between different BTB-domain-containing proteins can promote the formation of high-order chromatin loop structure, and thus regulate insulator function.

Another kind of well-characterized *Drosophila* insulators is the scs (special chromatin structures) and scs' insulators. They are located ~15 kb apart and flanking one of the *hsp70* loci [20, 21]. The proteins involved in their functions have also been identified [20, 21]. The two elements contain binding sites for proteins Zw-5 and boundary element-associated factor (BEAF), respectively. Recent results show that the scs and scs' insulators can interact with each other *in vivo* via Zw-5 and BEAF32 proteins bound to each element [20]. By using chromosome conformation capture (3C) technique, loop formation can be visualized [20]. Insulators could subdivide chromosome into autonomous higher order looped domains. However, this study only detected the localization of the Zw5 and BEAF proteins to scs and scs', respectively. It

is still a question whether or not these interactions will occur in polytene chromosome. More recently, another possible role of the two insulator elements was demonstrated. Kuhn *et al* suggest that the two elements do not limit the extent of heat shock-induced chromatin decondensation at region 87 of *Drosophila* polytene chromosomes and that scs and scs' may play a role in maintaining the transcriptional fidelity in this gene-dense region [21].

In vertebrates, the first best-characterized insulator is the CHS4, which is located near the 5' end of the chicken β -globin locus. Similar to the *Drosophila* Su(Hw) and scs insulators, the CHS4 element has barrier activity to protect transgenes from position effects and functions as an enhancer blocker. However, its barrier and enhancer-blocking function are separable [8, 22]. There are different proteins mediating the two kinds of functions, i.e. USF for barrier and the CCCTC-binding factor (CTCF) for enhancer blocking activity (Fig. 2). USF is a ubiquitous mammalian bHLH-ZIP transcription factor, while CTCF contains 11-zinc finger DNA-binding domains to bind to a wide range of cis-elements by varied contribution of individual zinc fingers [23]. CHS4 (fingerprint II) fragment containing CTCF binding sites is both necessary and sufficient for the enhancer-blocking activity [22]. Furthermore, CHS4 insulator could specifically protect the promoter from DNA methylation [24]. A recent study found that, in addition to histone deacetylation and the loss of histone H3-K4 methylation, specific promoter methylation also contributes to transgene silencing [25]. Therefore, the property of CHS4 could provide a potential strategy to sustain nor-

mal activity of foreign regulatory elements in transgenic assay [26].

Recent evidence suggests that human and mouse CTCF are also involved in insulator function. Conserved CTCF sites were found at 5' HS5 and 3' HS1 of both human and mouse *β-globin* loci [27]. This suggests that 3' HS1 and 5' HS5 may function as chromatin domain boundaries to prevent inappropriate interactions between *β-globin* regulatory elements and those of neighboring chromatin domains or subdomains.

As a chromatin insulator protein, CTCF is also involved in the imprinting of the *Igf2/H19* locus through binding to a maternal differentially methylated domain (DMD) (Fig. 3) [28]. Recent data propose that CTCF and the *H19* DMD together represent a chromatin insulator to regulate the mutually exclusive access of *H19* and *Igf2* to a set of shared enhancers in the downstream of *H19*. It is the first mammalian insulator shown to have a function at the endogenous locus, and its activity is DNA methylation-dependent.

CTCF binding to DMD is methylation sensitive and methylation itself is dependent on CTCF binding. When methylation prevents CTCF to bind to DMD, *H19* is bypassed and *Igf2* is activated by the enhancers. Mutation of CTCF binding site within DMD leads to loss of *H19* DMD chromatin insulator activity, *Igf2* is activated and maternal *H19* expression is also interestingly reduced. These results suggest that CTCF may also be a direct transcriptional activator of *H19* [28, 29]. However, more recently, Murrell *et al.* found that *H19* DMD mediated the binding of *Igf2* and *H19* into parent-specific chromatin loops [30]. This higher-order chromatin domain moving between an active and a silent chromatin state might cause decreased *H19* expression and the loss of the chromatin insulator function of *H19* DMD (Fig. 3). Another study found that the binding of CTCF to *H19* DMD could protect the *de novo* methylation of *H19* DMD and the increased methylation in the *H19* DMD is associated with substantial loss of CTCF protein [31]. More recently, it was shown that poly(ADP-ribosyl)ation on CTCF could regulate its insulator activity without influencing CTCF-DMD interactions [32]. This mechanism is not restricted to the imprinted *Igf2/H19* locus. ChIP (chromatin immunopurification) - on-chip assay showed that the link between CTCF, poly(ADP-ribosyl)ation and chromatin insulator function extends to more than 140 mouse CTCF target sites. This observation provides an alternative means to activate and deactivate insulator function without removing CTCF from the corresponding DNA binding regions [33].

Very recently, analysis of potential CTCF binding sites on a genome-wide basis was performed. More than 200

new CTCF target regions were identified by generating DNA microarrays using clones derived from chromatin immunopurified DNA sequences followed by ChIP-on-chip hybridization analysis [34]. The insulator trapping assays showed that the majority of these targets manifest insulator functions and the interested are potential heterochromatic sequences. The main properties of these CTCF-target sites are shared between heterochromatic and euchromatic domains, implicating that CTCF might have function in organizing active expression domains within the context of heterochromatin [34]. Like CTCF, the Su(Hw) insulator protein and GAFs that we mentioned above also bind hundreds of their respective sites within corresponding genomes [19]. One can envisage that there is still a large number of undiscovered chromatin insulators or boundary elements existing in invertebrate and vertebrate genomes.

In addition, CTCF has also been proposed a function of maintaining the expression of specific genes within domains of inactive X chromosome. The escape of such genes from X inactivation suggests the existence of specific chromatin boundary elements between domains. Filippova *et al* have recently reported, for the first time, the characterization of boundary elements between adjacent domains of inactive and active chromatin along the inactive X chromosome [35]. Their results show that CTCF binding sites are present at the 5' ends of mouse *Jarid1c* and *Eif2s3x* and of human *EIF2s3* genes. Each of them is able to escape X inactivation and is adjacent to a gene subject to X inactivation. These boundary regions are also associated with a high level of histone H3 acetylation. They propose that the binding of CTCF to both ends of an escaped domain might isolate such domains within a separate chromatin loop [35], in support of the hypothesis proposed by Mukhopadhyay *et al* as discussed earlier in this review [34]. It is very likely that CTCF has the potential function to establish autonomous domains within the heterochromatic regions.

Taken together, a number of insulators and boundary elements and their associated proteins have now been identified from invertebrate species to vertebrates. These insulator proteins are usually interact with each other or other factors to exhibit the insulator functions. Although an impressive variety of effects have been reported, many potential mechanisms of insulator action still remain unclear. Recent developments have revealed that some insulator proteins can recruit the histone-modifying enzymes to establish sharp chromatin domain boundary environment and that others can tether to specific nuclear sub-structure to subdivide chromatin domains into various separate environment.

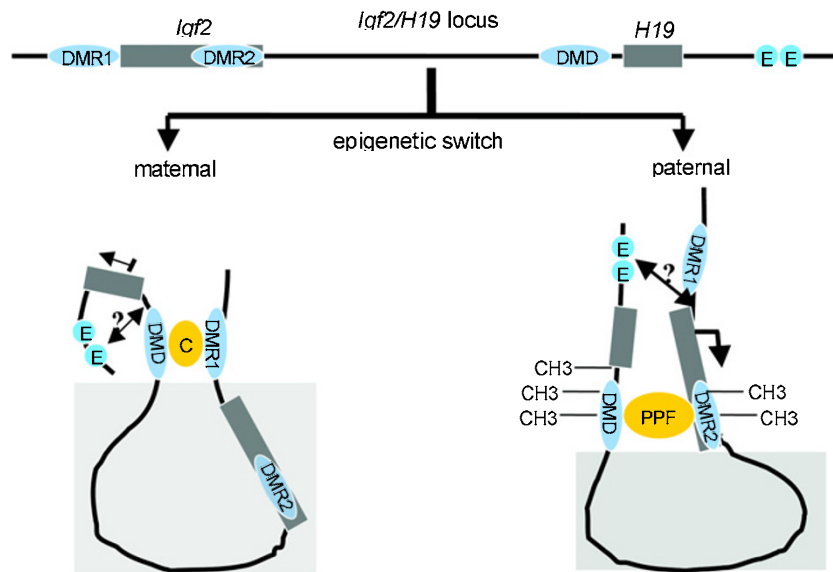


Fig. 3 One insulator involved chromatin-loop model at the *Igf2/H19* locus. On the maternal allele, the hypomethylated *H19* DMD boundary element binding with CTCF and related protein complex (C) is associated with the unmethylated silencer element on *Igf2* DMR1. The interaction between them forms two chromatin domains, with *H19* in an active domain (enhancer present) and *Igf2* in an inactive domain away from the enhancers. It results in *H19* expression. On the paternal chromosome, the methylated inactive *H19* DMD interacts with the methylated DMR2 through putative protein factors (PPF), placing both *H19* and *Igf2* in an active chromatin domain. The proximity of *Igf2* promoter and distant enhancers makes *Igf2* express, and *H19* is silent because of its promoter methylation although it is also in the active domain [30]. The hypothetical inactive chromatin domains are shown as shaded area. CH3 is referred to as methyl group. Physical interactions between the genes and the shared enhancers still need further investigation.

INSULATOR FUNCTION AND THE HISTONE CODE

Previous studies have characterized the flexible insulator functions as enhancer blockers and chromatin domain barrier (Fig. 1). Several recent observations suggest that particular patterns of histone modifications form histone codes that can regulate the structure of higher-order chromatin [24, 36]. The boundary elements are able to attenuate the propagation of neighboring repressive chromatin domains (Fig. 2) [24, 36]. These lead to illustrate the direct mechanisms of insulator barrier function.

Recent progress in the chromatin research field has fundamentally reshaped our views of chromatin states. Different histone modifications may differentially mark local chromatin domains. Acetylation of histone H3 and H4 and the methylation of histone H3 on lys4 are correlated with active chromatin domains, while lys9 of histone H3 is usually methylated over heterochromatin [37, 38]. Analysis of the histone acetylation and methylation patterns of the chicken β -globin locus has shown that histones within the domain are hyperacetylated and have a relatively high level of methylation on lys4 of histone H3 [39, 40]. In contrast, histones in the flanking domain of the locus, only a short distance upstream of the CHS4 insulator are

hypoacetylated and are enriched in histone H3 with methylated lys9. The region immediately surrounding CHS4 displays a constitutively higher level of histone acetylation, even in non-erythroid cells. These modifications are not found at the 3' HS enhancer-blocking element, which does not have a barrier activity [39, 40]. These results suggest that the barrier element present in CHS4 may be an entry site for histone acetyltransferase (HAT) that aids the formation of a barrier to protect the globin gene cluster from the influence of the condensed heterochromatin immediately upstream of CHS4.

More recently, West *et al* provide convincing evidence to support the hypothesis (Fig. 2) [8]. The authors previously have identified the *CHS4* core region, which is a sequence of 250 bp with five distinct protein binding regions or "footprints" [22]. Among them, footprintII (FII) shows enhancer-blocking property via its binding protein CTCF as described in previous sections. Other regions have potential barrier activity. The USF proteins are footprint IV (FIV) binding factors with barrier activity. In their recent study, the authors found that the USF proteins could recruit histone modifiers such as PCAF, P300/CBP and SET719 to the *CHS4* region *in vivo* [8]. The establishment of the USF-associated barrier complex helps to block

the spread of heterochromatin domain immediately upstream of globin locus. This is the first time a demonstration of a mechanism for how vertebrate barrier develop (Fig. 2).

Another study by Zhao *et al* suggests that loss of *CHS4* enhancer blocking activity impairs the remodeling of promoter chromatin by the enhancer [41]. *CHS4* provides a positional block to enhancer-mediated histone acetylation domain and reduces recruitment of the HATs CBP and P300 to the corresponding enhancer. In contrast, the region of the CTCF motif in the *CHS4* interposed between enhancer and promoter is the target of histone acetylation [41]. These results show that CTCF partially plays the role of barrier protein as USF. Further evidence comes from recent research of several chromatin domain boundaries on the mammalian X chromosome. These boundary elements contain CTCF binding sites and have CTCF-dependent barrier activity. The boundary regions are enriched in a high level of histone H3 acetylation [35]. These studies puzzle us about the role of USF as a barrier and CTCF as an enhancer blocker [42]. We need more detailed investigation and stronger evidence in order to understand the complexity of insulator functions in the future.

Collectively, it is evident that a histone code underlies barrier activity generated by a class of insulator-associated proteins. Histone modification such as acetylation is probably accompanied by parallel changes of other covalent modifications, such as phosphorylation, methylation, ubiquitination and /or ATP-dependent remodeling events [37, 38, 43]. Thus it will be interesting to see in the future whether or not such parallel histone modifications is existing to function as an insulator barrier to block the spreading of neighboring chromatin domains. Finally deciphering the histone code of insulator mechanism will be fundamental to understand its barrier function.

INSULATING AND LOOPING

Insulator has many functions with various mechanisms involved [6, 7, 11, 13, 44]. Several different models have been described to explain the boundary functions of insulators. These models can be grouped into two types of mutually nonexclusive catalogues: the dynamic modification and the passive physical models [45]. The former model, as we discussed above, suggests that boundary proteins mediate an active process by recruiting a multi-protein complex, bearing acetyltransferase and/or nucleosome remodeling activities, which hyperacetylates the histones at their binding sites and/or precludes nucleosome assembly. Thus, it can counteract silencing by interfering the linear spreading of neighboring heterochromatin domain and can prevent crosstalk of enhancers [8, 41]. The latter suggests that boundary elements congregate to form

a subnuclear compartment and/or are tethered to a nuclear structure, which then establish a physically autonomous domain to block the negative regulatory influences of surrounding heterochromatin [9, 10]. Accumulating evidence is emerging to support the later model. For example, recent data suggest that insulators can mediate the establishment of independent chromatin loop domains through their interactions with each other and /or tethering to the nuclear envelope or nuclear matrix [9, 20].

In yeast, a genetic screen was used to isolate nuclear-cytoplasmic transportin proteins that have strong boundary activities and are components of nuclear pore complexes (NPC). Identified factors can mediate the formation of an independent transcriptionally active loop minidomain within a preexisting heterochromatin region by tethering the HML locus to the NPC, linking boundary function to nuclear envelope [9]. Others also showed that when tethered to DNA, the transportins can reposition chromatin to the nuclear envelope, and the binding to the NPCs is required for the barrier activity of these transportins [10]. It is very likely that nuclear peripheral heterochromatin-associated active genes are tethered to the NPC and looped out for ready access by transcriptional regulators in the nucleoplasm. Alternatively, some active and inactive regions containing repressors and activators, may exist in the nuclear periphery, respectively. Thus, differentially regulated loci will be tethered to different regions. In either case, NPCs may have an additional function in defining transcriptionally active and inactive regions by insulating the loci from the influences of functionally distinct chromatin domains [46]. Taken together, these data illustrate that attachment to the nuclear periphery represents one important mechanism of insulator function. It also suggests that nuclear periphery is not merely a silencing milieu but also plays a more complex and subtle role in gene regulation [46].

Some studies recently suggest that CTCF protein, which binds insulators in the *β -globin* locus in chickens, mice, and humans also can target insulators to the nucleolar surface [47, 48]. It was shown that CTCF in nuclear extracts forms a stable and well-defined complex with a nucleolar protein nucleophosmin [47]. This protein is present in vivo at *CHS4* insulator sites as well as at exogenous CTCF binding sites. Furthermore, exogenous insulator sites are localized to the nuclear periphery in a CTCF-dependent manner. Further study indicates that CTCF and the *CHS4* insulator sequence are highly concentrated in the nuclear matrix fraction [48]. Collectively, the results show that the *CHS4* insulator is a CTCF-dependent nuclear matrix-associated element. The authors propose that CTCF-associated insulator complex through tethering to nuclear matrix subdivide the chromosome into looped domains that

are functionally isolated from one another, such that an enhancer in one loop could not efficiently interact with a promoter in another [47]. Additionally, this insulator CTCF-dependent action may share some molecular mechanisms with the example in yeast as discussed above.

CTCF boundary elements functioning in chromatin looping has also been found in the imprinted *Igf2/H19* domain (Fig. 3). The mouse *Igf2* gene is separated by ~100 kb from non-coding gene *H19* on mouse distal chromosome 7. In humans and mice, the *Igf2* and *H19* genes are reciprocally imprinted, giving paternal-specific expression of *Igf2* and maternal-specific expression of *H19* [49]. A set of elements is involved in transcription of these genes. Differentially methylated region1 (DMR1), located upstream of *Igf2* promoter1, and DMR2, located within exon 6 of *Igf2*, have a silencer function and an activator function, respectively. *H19* DMD contains a CTCF-associated chromatin boundary as described. These chromatin boundaries, silencers and activators collaboratively regulate the reciprocal expression of the two genes in a methylation-sensitive manner. Additional sets of enhancers locate downstream of *H19* to promote higher expression (Fig. 3) [28]. It is generally agreed that the potential molecular mechanism for the reciprocal expression of *Igf2* and *H19* genes involves higher-order chromatin structures, which enable CTCF boundary elements to interact so as to form chromatin loop domains. Murrell *et al* recently detected preferential interactions between the *H19* DMD and the *Igf2* DMRs and indeed demonstrated that such gene expression pattern is involved in epigenetic-controlled formation of different chromatin-loop domains [30]. The different interactions of *H19* DMD and *Igf2* DMRs generate a simple epigenetic switch for *Igf2* through which it is moved into either an inactive domain or an active domain in close to the enhancers of *H19* downstream (Fig. 3). This model provides a novel mechanism for the actions of chromatin boundaries and other regulatory elements [28, 30]. Insulators may have an ability to relocate specific domains to new nuclear environments as part of their mechanism of action.

Taken together with several studies on *Drosophila* Su (Hw) insulator discussed previously, it is evident that insulator functions in the establishment of chromatin loop domains. However, with the development of 3C technology, recently identified looped chromatin domain at the mammalian β -globin locus may therefore function as insulators, suggesting the function of chromatin loops in insulating.

Enhancers, LCRs and insulators always regulate gene expression from a long distance. Insulators have features to form chromatin loop. Recent studies demonstrated a direct physical contact between the LCR and expressed genes in the active mammalian β -globin locus to form

specific chromatin loop domains [50]. Further study in transgenic mice shows that these loop domains are developmentally conserved between mice and humans during erythroid differentiation [51]. The investigators propose the loop formation to be a general phenomenon that is a key element in setting up autonomous expression of genes irrespective of their position in the genome [52]. Its activity may be involved in insulating the chromatin inside the loop from the influence of what is outside by acting as a physical barrier to prevent the spreading of heterochromatin or by concentrating chromatin modifiers that antagonize silent chromatin structure [52, 53]. This stable loop formation may be contributed to position independent expression of human β -globin gene cluster in transgenic mice [54].

Our lab has recently detected the spatial clustering of regulatory elements and genes in murine α -globin locus (Xin L, Zhou GL, Liu DP *et al*, unpublished data). Our results suggest that a similar chromatin loop conformation exists in the α -globin locus compared to β -globin locus. However, the expression of human α -globin genes was demonstrated to be position-dependent [55]. Studies in human, mouse and rabbit indicate that the α -globin locus displays characteristics of ubiquitously expressed housekeeping gene, in contrast to β -globin and other inactive genes studied to date [56]. This demonstrates that α - and β -globin loci are divergent in nuclear location with different chromatin status, and their expression occurs within different nuclear domains in haemopoietic cells. A precise regulation of these genes ensures a balanced protein amount from both loci [56]. Considering the differences between both loci in vivo, in combination of recent studies on mammalian β -globin loci, we propose that the formation of the higher order chromatin conformation on β -globin locus has additional molecular mechanisms. Furthermore, the clustering and looping of β -globin gene cluster perhaps have additional functions. Two models for the formation and function of chromatin loop domain at the mammalian β -globin loci are suggested.

The first model is that CTCF boundary elements in the 5' and 3' regions of mammalian β -globin locus interact with each other and /or tether to nuclear structure to form the chromatin loop domain. Such conformation conversely functions as insulator buffering against surrounding repressive chromatin domains as discussed previously. Recent studies suggest that, like chicken β -globin locus, CTCF insulator elements flanking the mouse and human β -globin loci are conserved [27]. Another model is that CTCF-associated chromatin boundaries may help to relocate the β -globin locus to an active nuclear compartment. For instance, the β -globin promoter localized in pericentromeric regions can not be accessed by NF-E2, a

heterodimer of p18 and p45. The β -globin gene activation requires transcription factor NF-E2. Therefore, in order to be fully activated, the β -globin locus first has to be released from heterochromatin and relocated to a specific euchromatic region to which NF-E2 can be targeted [57]. Additionally, distanced localization of the human β -globin locus from centromeric heterochromatin is required to obtain a DNA nuclease-sensitive, open chromatin state, which is marked by acetylation of histones H3 and H4 [58].

In summary, the studies discussed above have demonstrated a significant role of insulators acting on chromatin loop formation. Conversely, chromatin loops always maintain insulating activities to some extent. One can envisage that insulating and looping may correlate with each other. This information in conjunction with the detailed studies on the mammalian β -globin loci make it possible to predict the establishment of higher order clustering and looping conformation according to chromatin boundary mechanisms.

CONCLUDING REMARKS

The concept of insulator elements arose as an explanation for the existence of independently regulatory domains along the chromosomes. Although many studies over the past several years have established the basic properties of insulators, it's apparent that we are just at the beginning of understanding the global functions and mechanisms of insulators. There are probably many surprises ahead.

Practically, insulators have attractively potential applications in gene therapy and transgenesis. For a long time, gene therapists are trying to exogenously insert a wild-type gene to restore the appropriate expression and function of a damaged gene. Various viral vectors have been designed to deliver the correct copy to particular types of cells or tissues [26]. However, transgenes delivered by these vectors are usually susceptible to chromosomal position effects that will influence transgene expression and frequently lead to transcriptional silencing, thus poses a significant obstacle in gene therapy (Fig. 1). Additionally, the stochastic insertion of transgene will probably cause inappropriate activation of adjacent genes after integration into the genome. This mechanism could explain recent observation of a leukemia-like disorder in a gene therapy trial against an immune disease [59]. The incorporation of insulators into gene therapy vectors might offer a solution to this problem. Therefore, elucidating the whole bandwidth of insulator mechanisms is an exciting challenge in clinical trials and lead therapeutic concepts into novel gene therapy directions.

ACKNOWLEDGEMENTS

We apologize to colleagues whose work is not or only

indirectly cited due to space limitations and the scope of the review. This work was supported by the grant from the National Natural Science Foundation of China (No. 30393110).

REFERENCES

- Xin L, Liu DP, Ling CC. A hypothesis for chromatin domain opening. *Bioessays* 2003; **25**:507-14
- Felsenfeld G, Groudine M. Controlling the double helix. *Nature* 2003; **421**:448-53
- Wei GH, Liu DP, Liang CC. Charting gene regulatory networks: strategies, challenges and perspectives. *Biochem J* 2004; **381**:1-12
- Gilbert N, Boyle S, Fiegler H, et al. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* 2004; **118**:555-66
- Sun FL, Elgin SC. Putting boundaries on silence. *Cell* 1999; **99**:459-62
- Capelson M, Corces VG. Boundary elements and nuclear organization. *Biol Cell*. 2004; **96**:617-29
- West AG, Gaszner M, Felsenfeld G. Insulators: many functions, many mechanisms. *Genes Dev* 2002; **16**:271-88
- West AG, Huang S, Gaszner M, Litt MD, Felsenfeld G. Recruitment of histone modifications by USF proteins at a vertebrate barrier element. *Mol Cell* 2004; **16**:453-63
- Ishii K, Arib G, Lin C, Van Houwe G, Laemmli UK. Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* 2002; **109**:551-62
- Shinkura N, Forrester WC. Pushing the envelope: chromatin boundaries at the nuclear pore. *Mol Cell* 2002; **9**:1156-8
- Bell AC, West AG, Felsenfeld G. Insulators and boundaries: versatile regulatory elements in the eukaryotic genome. *Science* 2001; **291**:447-50
- Zhan HC, Liu DP, Liang CC. Insulator: from chromatin domain boundary to gene regulation. *Hum Genet* 2001; **109**:471-8
- Kuhn EJ, Geyer PK. Genomic insulators: connecting properties to mechanism. *Curr Opin Cell Biol* 2003; **15**:259-65.
- Muravyova E, Golovnin A, Gracheva E, et al. Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science* 2001; **291**:495-8
- Cai HN, Shen P. Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science* 2001; **291**:493-5
- Byrd K, Corces VG. Visualization of chromatin domains created by the gypsy insulator of *Drosophila*. *J Cell Biol* 2003; **162**:565-74
- Melnikova L, Juge F, Gruzdeva N, et al. Interaction between the GAGA factor and Mod(mdg4) proteins promotes insulator bypass in *Drosophila*. *Proc Natl Acad Sci U S A*. 2004; **101**:14806-11
- Pai CY, Lei EP, Ghosh D, Corces VG. The centrosomal protein CP190 is a component of the gypsy chromatin insulator. *Mol Cell* 2004; **16**:737-48
- Lehmann M. Anything else but GAGA: a nonhistone protein complex reshapes chromatin structure. *Trends Genet* 2004; **20**:15-22
- Blanton J, Gaszner M, Schedl P. Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev* 2003; **17**:664-75
- Kuhn EJ, Hart CM, Geyer PK. Studies of the role of the *Droso-*

- phila* scs and scs' insulators in defining boundaries of a chromosome puff. *Mol Cell Biol* 2004; **24**:1470-80
- 22 Bell AC, West AG, Felsenfeld G. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 1999; **98**:387-96
 - 23 Ohlsson R, Renkawitz R, Lobanenkov V. CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. *Trends Genet* 2001; **17**:520-7
 - 24 Mutskov VJ, Farrell CM, Wade PA, Wolffe AP, Felsenfeld G. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev* 2002; **16**:1540-54
 - 25 Mutskov V, Felsenfeld G. Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *EMBO J* 2004; **23**:138-49
 - 26 Recillas-Targa F, Valadez-Graham V, Farrell CM. Prospects and implications of using chromatin insulators in gene therapy and transgenesis. *Bioessays* 2004; **26**:796-807
 - 27 Farrell CM, West AG, Felsenfeld G. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol Cell Biol* 2002; **22**:3820-31
 - 28 Kato Y, Sasaki H. Imprinting and looping: epigenetic marks control interactions between regulatory elements. *Bioessays* 2005; **27**:1-4
 - 29 Lee JT. Molecular links between X-inactivation and autosomal imprinting: X-inactivation as a driving force for the evolution of imprinting? *Curr Biol* 2003; **13**:R242-54
 - 30 Murrell A, Heeson S, Reik W. Interaction between differentially methylated regions partitions the imprinted genes *Igf2* and *H19* into parent-specific chromatin loops. *Nat Genet* 2004; **36**:889-93
 - 31 Fedoriw AM, Stein P, Svoboda P, Schultz RM, Bartolomei MS. Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. *Science* 2004; **303**:238-40
 - 32 Yu W, Ginjala V, Pant V, et al. Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat Genet* 2004; **36**:1105-10
 - 33 Jeong S, Pfeifer K. Shifting insulator boundaries. *Nat Genet* 2004; **36**:1036-7
 - 34 Mukhopadhyay R, Yu W, Whitehead J, et al. The binding sites for the chromatin insulator protein CTCF map to DNA methylation-free domains genome-wide. *Genome Res* 2004; **14**:1594-602
 - 35 Filippova GN, Mimi K, Cheng, James M. Moore, et al. Boundaries between chromosomal domains of X inactivation and sscpe bind CTCF and lack CpG methylation during early development. *Developmental Cell* 2005; **8**:31-42
 - 36 Donze D, Kamakaka RT. Braking the silence: how heterochromatic gene repression is stopped in its tracks. *Bioessays* 2002; **24**:344-9
 - 37 Jenuwein T, Allis CD. Translating the histone code. *Science* 2001; **293**:1074-80
 - 38 Fischle W, Wang Y, Allis CD. Histone and chromatin cross-talk. *Curr Opin Cell Biol* 2003; **15**:172-83
 - 39 Litt MD, Simpson M, Recillas-Targa F, Prioleau MN, Felsenfeld G. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J* 2001; **20**:2224-35
 - 40 Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* 2001; **293**:2453-5
 - 41 Zhao H, Dean A. An insulator blocks spreading of histone acetylation and interferes with RNA polymerase II transfer between an enhancer and gene. *Nucleic Acids Res* 2004; **32**:4903-19
 - 42 Zhou J, Berger SL. Good fences make good neighbors: barrier elements and genomic regulation. *Mol Cell* 2004; **16**:500-2
 - 43 Fischle W, Wang Y, Allis CD. Binary switches and modification cassettes in histone biology and beyond. *Nature* 2003; **425**:475-9
 - 44 Oki M, Kamakaka RT. Blockers and barriers to transcription: competing activities? *Curr Opin Cell Biol* 2002; **14**:299-304
 - 45 Ishii K, Laemmli UK. Structural and dynamic functions establish chromatin domains. *Mol Cell* 2003; **11**:237-48
 - 46 Misteli T. Spatial positioning: a new dimension in genome function. *Cell* 2004; **119**:153-6
 - 47 Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* 2004; **13**:291-8
 - 48 Yusufzai TM, Felsenfeld G. The 5'-HS4 chicken β -globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci U S A* 2004; **101**:8620-4
 - 49 Lewis A, Murrell A. Genomic imprinting: CTCF protects the boundaries. *Curr Biol* 2004; **14**:R284-6
 - 50 Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W. Looping and interaction between hypersensitive sites in the active β -globin locus. *Mol Cell* 2002; **10**:1453-65
 - 51 Palstra RJ, Tolhuis B, Splinter E, et al. The β -globin nuclear compartment in development and erythroid differentiation. *Nat Genet* 2003; **35**:190-4
 - 52 de Laat W, Grosveld F. Spatial organization of gene expression: the active chromatin hub. *Chromosome Res* 2003; **11**:447-59
 - 53 Chambeyron S, Bickmore WA. Does looping and clustering in the nucleus regulate gene expression? *Curr Opin Cell Biol* 2004; **16**:256-62
 - 54 Grosveld F, van Assendelft GB, Greaves DR, Kollias G. Position-independent, high-level expression of the human β -globin gene in transgenic mice. *Cell* 1987; **51**:975-85
 - 55 Feng DX, Liu DP, Huang Y, et al. The expression of human α -like globin genes in transgenic mice mediated by bacterial artificial chromosome. *Proc Natl Acad Sci U S A* 2001; **98**:15073-7
 - 56 Brown KE, Amoils S, Horn JM, et al. Expression of α - and β -globin genes occurs within different nuclear domains in haemopoietic cells. *Nat Cell Biol* 2001; **3**:602-6
 - 57 Isogai Y, Tjian R. Targeting genes and transcription factors to segregated nuclear compartments. *Curr Opin Cell Biol* 2003; **15**:296-303
 - 58 Schubeler D, Francastel C, Cimbora DM, et al. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev* 2000; **14**:940-50
 - 59 Hacein-Bey-Abina S, von Kalle C, Schmidt M, et al. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med* 2003; **348**:255-6