

Roles of Chromatin insulators in gene regulation and diseases

Seyed Mohammad Javad Mousavi¹, Farshid Yeganeh^{1*}

¹Department of Immunology, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

Received: 5 Sep 2013

Revised : 15 Oct 2013

Accepted: 7 Nov 2013

Corresponding Author:

Farshid Yeganeh
Immunology Department, Shahid Beheshti University of Medical Sciences, Koodakyar st., Velenjak, Tehran, Iran.
Phone: 09127200243
Fax: 021-22439970
E-mail: fyeganeh@sbmu.ac.ir

Abstract

With advances in genetic science, the dynamic structure of eukaryotic genome is considered as basis of gene expression regulation. Long-distance communication between regulatory elements and target promoters is critical and the mechanisms responsible for this connection are just starting to emerge. Chromatin insulators are key determinants of proper gene regulation and precise organization of chromosomal structures. These elements are DNA-protein complexes which can mediate interactions within and between chromosomes. In doing so, they can bring close together regulatory sequences located at large distances from each other. In addition, insulators may be important in remodeling the architecture of chromosomes in the cell cycle. Here, we review recent data revealing the structure and properties of these DNA-protein complexes, their global role in regulation of transcription, and the mechanisms by which misregulation of their function leads to various human diseases such as cancers and autoimmune diseases.

Introduction

All eukaryotic genes have control elements which determine the tissue, the situations and the time of gene expression. The disruption of regulatory mechanisms of gene expression can have severe consequences, leading to developmental disorders and diseases such as cancer. So the proper regulation of gene activation and gene silencing is of utmost importance. To achieve this goal, several types of distinct regulatory elements play role to organize genome to distinct domains (1). The main classes of these elements are: enhancers, silencers, locus control regions (LCRs) and insulators. These elements are very far apart from each other sometimes, but still able to interact with each other. As a result, the pattern of transcription of a gene is determined by interaction between regulatory elements (2). Among them, chromatin insulators may be a key determinant for the proper regulation of gene expression and organization of chromosomes architecture (3, 4). These DNA sequences are introduced the first time in *Drosophila melanogaster* as regulatory factors that block the aberrant stimulatory effect of distant enhancer (5, 6). Further analysis showed that insulators additionally act as boundaries to limit heterochromatin and euchromatin domains (7, 8). In general Insulators are protein-DNA complexes; act as factors inhibiting inappropriate expansion of

heterochromatin or block stimulatory effect of enhancers (9).

Structure and organization of insulators

Insulators are classified into different families according to the insulator binding protein (IBP), which is essential for their activity (10). Genome wide studies (GWS) on *Drosophila*, have led to the identification of five insulator families (Table 1).

Table1. Table summarizes different *Drosophila melanogaster* insulators and their roles. IBP; Insulator Binding Protein.

Insulator	Abbreviation of IBP	Full name of IBP	Role
<i>gypsy</i>	Su(Hw)	Suppressor of hairywing	Female reproductivity
<i>scs'</i>	BEAF	Boundary element-associated factor	Chromosome structure and cell viability
<i>scs</i>	Zw5	Zeste-White 5	Wing, bristle and eye development
<i>Fab-7</i>	GAF	GAGA factor	cell viability
<i>Fab-8</i>	dCTCF	<i>Drosophila</i> CTCF homolog	Abdominal homeotic phenotypes

In vertebrates, the CCCTC-binding factor (CTCF) has been characterized as the only insulator protein (9). CTCF is a highly conserved protein containing 11 zinc-finger domains bind to DNA and has many homology in vertebrates (Ohlsson et al., 2001). This structure explains why there is only one insulator protein to regulate a wide variety of DNA sequences in vertebrate. In mice, CTCF homozygous gene knockout causes embryo loss before implantation. This finding shows the tremendous importance of this protein in various cellular processes (11).

To infer the distribution and organization of the insulators in genome, specific staining of insulator binding protein is need (1). The presence of proteins associated with chromatin insulators in genome have been mapped by chromatin immunoprecipitation and Microarray (ChIP-chip) method (12, 13). Throughout the DNA, Insulators attach to the thousands motifs. Differences in the motifs of DNA, can affect the protein binding capacity and functional properties of the insulators (Fig.1). Comparison of the exact positions of the CTCF binding site in 38 human cell lines revealed that while the majority of insulators are fixed, but there are thousands of CTCF specific binding site (14). So, CTCF epigenetically divided the genome to distinct transcriptional units which have unique expression characteristics. According to recent findings, it is expected that some interactions between insulators are cell type-specific and have a vital role in dynamism of gene expression during development (15)

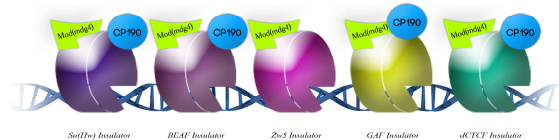


Fig 1. Organization of *Drosophila* insulators. Each insulator is composed of a DNA binding protein that bind to a specific sequence of DNA, plus accessory proteins that mediate interactions between insulator sites. In *Drosophila*, accessory proteins include Mod (mdg4) and CP190. Homologues of Mod (mdg4) and CP190 have not been found in vertebrates.

There has been a great interest in identifying where potential insulators are located in the eukaryotic genome, because knowledge of these elements can help understand how cis-regulatory elements coordinate expression of the target genes. Kim et al described 13804 CTCF-binding sites in potential insulators of the human genome, discovered experimentally in primary human fibroblasts. Most of these sequences are located far from the transcriptional start sites (are generally located an average of 48 kb away from promoters). They showed that nearly half of the CTCF binding sites are located in the

intergenic regions (46%). Only about 20% CTCF sites are near transcription start sites. Unexpectedly, a significant number of CTCF binding sites fall within genes, with 22% in the introns and 12% in the exons. Based on the results, they concluded that the distribution of CTCF binding sites correlates strongly with the number of genes on each chromosome (16).

Insulator roles in gene regulation

As described above, insulator proteins are present in intergenic regions as well as at the 5', or 3' of genes. The genomic distribution of insulators suggests that they can play broad roll in gene regulation. The most of information about the role of insulators is achieved from experiments on transgenic constructs. For example, in the study of "*blocking of enhancer*" transgene insulator placed between the enhancer and promoter; then the degrees of inhibition (insulation degree) was evaluated (17). Based on functional characteristics, insulators are divided into two main types (2). The first group is "*enhancer blockers*" which have regulatory role and prevent aberrant activation of genes. When this sequence placed between an enhancer and promoter, stops the enhancer stimulatory effect on promoter (1). This finding explains how, the stimulatory activity of enhancer is limited to a specific promoter, despite their ability to stimulate distant sequences (fig.2A).

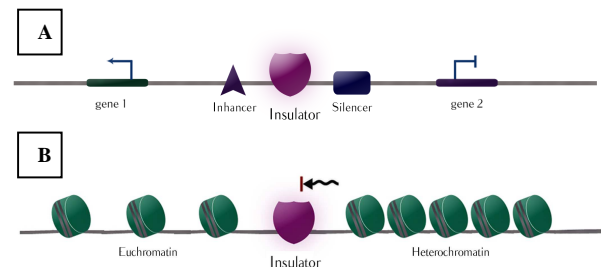


Fig 2. Hypothetical model explains how the insulators regulate gene expression. (A); *enhancer-blocker insulators* regulate gene expression by limiting enhancer effect to gene 1 and silencer effect to gene 2 (B) *Barrier insulators* physically limit heterochromatin regions, to avoid their spreading to other parts of the genome that have to be expressed.

The second group is "*Barrier sequences*". This group is the factors that prevent the extension of heterochromatin into adjacent euchromatin structures (fig. 2 B). It should be noted that all insulators do not use the same mechanism (1). Different factors attached to insulators; promote different mechanisms their use to insulate (18).

Insulator roles in nuclear structure

Recent studies showed that insulators facilitate intra- and inter-chromosomal interactions to affect transcription, imprinting and recombination (i.e VDJ

recombination in T cell or B cell receptors production). In addition, CTCF binds chromatin boundaries leads to anchor the DNA to cellular structures such as the nuclear lamina (19). Insulators operate primarily through creation and maintaining the three-dimensional structure of DNA by mediation of CTCF protein (Fig. 3).

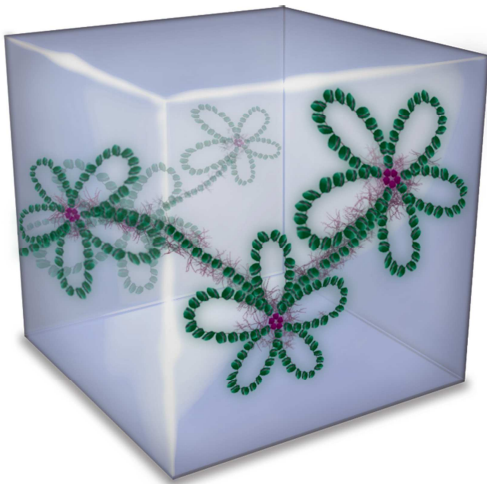


Fig 3. In the nucleus, Insulators play main role to construct the three-dimensional organization of the chromatin fiber. Green beads represent nucleosomes. The nuclear lamina/matrix is shown in red and Insulator proteins are shown as shield of pink.

Two CTCF molecules form a homodimer structure, causing a loop to DNA may serve as an anchor to attach insulator sites to the nuclear lamina or nuclear matrix (20, 21). Therefore, insulators can contribute in interphase chromatin organization via interaction with specific nuclear structures such as the nuclear pore complexes (19, 22-23).

How Insulator protein binds to DNA

In vertebrates, methylation of CpG sequences around CTCF binding site is the main mechanism that controls the binding of insulator proteins to DNA. This issue has been studied in detail in H19 and Insulin-like growth factor 2 (Igf2) loci, and found that CTCF binds only to non-methylated CpG sequences (24). The Igf2 and H19 genes are imprinted, resulting in silencing of the maternal and paternal alleles, respectively. On the maternal allele, CTCF binds the imprinting control region (ICR) and interacts with the IGF2 promoter, blocking activation by the upstream enhancers, which are then able to interact with the H19 gene. On the paternal allele, the CTCF-binding sites in the ICR are methylated and CTCF is unable to bind the ICR. The enhancers can then activate the IGF2 gene while repressors bind H19. This mechanism has been observed in

several other loci. So, CTCF binding regulation through DNA methylation is a common mechanism to regulate the function of insulators. The second mechanism controls the insulators function is shared by mammals and invertebrates such as drosophila. In this mechanism, competitor proteins compete with CTCF for binding to DNA. For example CTCF like protein (CTCF-L or BORIS) binds to similar sequences in DNA and inhibit CTCF binding capacity (25).

CTCF is linked to diseases

CTCF involves in multiple features of genetic and epigenetic, including regulation of proliferation, differentiation, apoptosis, genomic imprinting and X-chromosome inactivation (26, 27). Recent studies have been shown that abnormal methylation pattern of CTCF targets are associated with several cancers (28). For example abnormal methylation pattern of CTCF targets in IGF2/H19 imprint center result in overexpression of IGF2, which is a prominent characteristic of many ovarian malignancies (29). Additionally, recent in-vivo and in-vitro studies have been shown that retinoblastoma (RB) has CTCF dependent gene expression pattern (30). Moreover, in cancers of lung, breast, colorectal and multiple myeloma, a tumor suppressor gene P16 (INK4a) is a frequent target of epigenetic silencing in the absence of CTCF (31- 33).

CTCF gene is located on the 16q22. Thus far, no deletions or truncating frame-shift mutations were detected in CTCF gene. This finding reveals that complete loss of CTCF function is incompatible with cell survival. Consequently, CTCF could be considered as an unusual type of tumor suppressor gene, showing “*change of function*” instead of complete “*loss of function*” in cancer (34).

Emerging evidences also point to the role of CTCF deregulation in diseases other than cancers. For example, CTCF malfunction due to inappropriate DNA methylation pattern was observed in Systemic lupus erythematosus patients (35). Additional, CTCF controls non-coding transcription repeat; and has many roles in preventing of *trinucleotide repeat-associated* diseases. Therefore, loss of CTCF-binding sequences which is associated with expansion of repeats, lead to development of diseases such as myotonic dystrophy, spinocerebellar ataxia, spinocerebellar ataxia, dentatorubral-pallidoluysian atrophy, and Huntington's disease (36, 37).

Conclusion

In Comparison of DNA with a language, insulators are similar to a “comma”; they define the boundaries to affect enhancer-promoter interactions and interfere with the spreading of heterochromatin. Although they identified first in drosophila, but, today, from

yeast to humans, Insulators are considered as one of the key regulators of gene expression. Stability of transcriptionally active genes during cell division, reflect the accurate function of regulatory elements in the cell. Even though they are defined based on their boundary ability, their role appears to be more general. Intra- and inter-chromosomal interactions mediated by insulator proteins may form a web of contacts between insulators that give rise to precise patterns of nuclear organization. So it is expected that any defect in insulators, causing abnormal gene expression, resulting in diseases and cancers.

References

1. Bruce Alberts AJ, Julian Lewis, Martin Raff, Keith Roberts, Peter Walter. *Molecular Biology of the Cell*. 5th ed. USA: Garland Science; 2012 June 7.
2. Bulger M, Groudine M. Enhancers: the abundance and function of regulatory sequences beyond promoters. *Developmental biology*. 2010; 339 (2): 250-7. PMID: 20025863
3. Capelson M, Corces VG. Boundary elements and nuclear organization. *Biology of the cell / under the auspices of the European Cell Biology Organization*. 2004; 96 (8): 617-29. PMID: 15519696
4. Vogelmann J, Valeri A, Guillou E, Cuvier O, Nollmann M. Roles of chromatin insulator proteins in higher-order chromatin organization and transcription regulation. *Nucleus*. 2011; 2 (5): 358-69. PMID: 21983085
5. Geyer PK, Corces VG. DNA position-specific repression of transcription by a *Drosophila* zinc finger protein. *Genes & development*. 1992; 6 (10): 1865-73. PMID: 1327958
6. Reitman M, Lee E, Westphal H, Felsenfeld G. Site-independent expression of the chicken beta A-globin gene in transgenic mice. *Nature*. 1990; 348 (6303): 749-52. PMID: 2175398
7. Gyurkovics H, Gausz J, Kummer J, Karch F. A new homeotic mutation in the *Drosophila* bithorax complex removes a boundary separating two domains of regulation. *The EMBO journal*. 1990; 9 (8): 2579-85. PMID: 1973385
8. Udvardy A, Maine E, Schedl P. The 87A7 chromomere. Identification of novel chromatin structures flanking the heat shock locus that may define the boundaries of higher order domains. *Journal of molecular biology*. 1985; 185 (2): 341-58. PMID: 2997449
9. Wallace JA, Felsenfeld G. We gather together: insulators and genome organization. *Current opinion in genetics & development*. 2007; 17 (5): 400-7. PMID: 17913488
10. Maeda RK, Karch F. Making connections: boundaries and insulators in *Drosophila*. *Current opinion in genetics & development*. 2007; 17 (5): 394-9. PMID: 17904351
11. Splinter E, Heath H, Kooren J, Palstra RJ, Klous P, Grosveld F, et al. CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes & development*. 2006; 20 (17): 2349-54. PMID: 16951251
12. Bushey AM, Ramos E, Corces VG. Three subclasses of a *Drosophila* insulator show distinct and cell type-specific genomic distributions. *Genes & development*. 2009; 23 (11): 1338-50. PMID: 19443682
13. Negre N, Brown CD, Shah PK, Kheradpour P, Morrison CA, Henikoff JG, et al. A comprehensive map of insulator elements for the *Drosophila* genome. *PLoS genetics*. 2010; 6 (1): e1000814. PMID: 20084099
14. Chen H, Tian Y, Shu W, Bo X, Wang S. Comprehensive identification and annotation of cell type-specific and ubiquitous CTCF-binding sites in the human genome. *PloS one*. 2012; 7 (7): e41374. PMID: 22829947
15. Wood AM, Van Bortle K, Ramos E, Takenaka N, Rohrbach M, Jones BC, et al. Regulation of chromatin organization and inducible gene expression by a *Drosophila* insulator. *Molecular cell*. 2011; 44 (1): 29-38. PMID: 21981916
16. Kim TH, Abdullaev ZK, Smith AD, Ching KA, Loukinov DI, Green Roland D, et al. Analysis of the Vertebrate Insulator Protein CTCF-Binding Sites in the Human Genome. *Cell*. 2007; 128 (6): 1231-45. PMID: 17382889
17. Phillips JE, Corces VG. CTCF: master weaver of the genome. *Cell*. 2009; 137 (7): 1194-211. PMID: 19563753
18. Benjamin Lewin JEK, Stephen T Kilpatrick, Elliott S Goldstein. *Lewin's genes X*. 10th ed: Sudbury, Mass. : Jones and Bartlett 2011.
19. Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, et al. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature*. 2008; 453 (7197): 948-51. PMID: 18463634
20. Hou C, Zhao H, Tanimoto K, Dean A. CTCF-dependent enhancer-blocking by alternative chromatin loop formation. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105 (51): 20398-403. PMID: 19074263
21. Yusufzai TM, Tagami H, Nakatani Y, Felsenfeld G. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Molecular cell*. 2004; 13 (2): 291-8. PMID: 14759373
22. Ishii K, Laemmli UK. Structural and dynamic functions establish chromatin domains. *Molecular cell*. 2003; 11 (1): 237-48. PMID: 12535536
23. Kalverda B, Fornerod M. Characterization of genome-nucleoporin interactions in *Drosophila* links chromatin insulators to the nuclear pore complex. *Cell Cycle*. 2010; 9 (24): 4812-7. PMID: 21150273
24. Bell AC, Felsenfeld G. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature*. 2000; 405 (6785): 482-5. PMID: 10839546
25. Loukinov DI, Pugacheva E, Vatolin S, Pack SD, Moon H, Chernukhin I, et al. 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 USA*. 2002; 99 (10): 6806-11. PMID: 12011441
26. Filippova GN. Genetics and epigenetics of the multifunctional protein CTCF. *Current topics in developmental biology*. 2008; 80: 337-60. PMID: 17950379

27. Heath H, Ribeiro de Almeida C, Sleutels F, Dingjan G, van de Nobelen S, Jonkers I, et al. CTCF regulates cell cycle progression of alphabeta T cells in the thymus. *The EMBO journal*. 2008; 27 (21): 2839-50. PMID: 18923423
28. Klenova EM, Morse HC, 3rd, Ohlsson R, Lobanenko VV. The novel BORIS + CTCF gene family is uniquely involved in the epigenetics of normal biology and cancer. *Seminars in cancer biology*. 2002; 12 (5): 399-414. PMID: 12191639
29. Huang Z, Murphy SK. Increased Intragenic IGF2 Methylation is Associated with Repression of Insulator Activity and Elevated Expression in Serous Ovarian Carcinoma. *Frontiers in oncology*. 2013; 3: 131. PMID: 23745176
30. De La Rosa-Velazquez IA, Rincon-Arango H, Benitez-Bribiesca L, Recillas-Targa F. Epigenetic regulation of the human retinoblastoma tumor suppressor gene promoter by CTCF. *Cancer research*. 2007; 67 (6): 2577-85. PMID: 17363576
31. Rodriguez C, Borgel J, Court F, Cathala G, Forne T, Piette J. CTCF is a DNA methylation-sensitive positive regulator of the INK/ARF locus. *Biochemical and biophysical research communications*. 2010; 392 (2): 129-34. PMID: 20051228
32. Witcher M, Emerson BM. Epigenetic silencing of the p16(INK4a) tumor suppressor is associated with loss of CTCF binding and a chromatin boundary. *Molecular cell*. 2009; 34 (3): 271-84. PMID: 19450526
33. Valadan, R., Rafiei, A., Tehrani, M., & Nejatollahi, F. Resistance to HER2-targeted therapy. *Research in Molecular Medicine*. 2013; 1 (1): 1-9.
34. Filippova GN, Qi CF, Ulmer JE, Moore JM, Ward MD, Hu YJ, et al. Tumor-associated zinc finger mutations in the CTCF transcription factor selectively alter its DNA-binding specificity. *Cancer research*. 2002; 62 (1): 48-52. PMID: 11782357
35. Robertson KD. DNA methylation and human disease. *Nature reviews Genetics*. 2005; 6 (8): 597-610. PMID: 16136652
36. Cho DH, Thienes CP, Mahoney SE, Analau E, Filippova GN, Tapscott SJ. Antisense transcription and heterochromatin at the DM1 CTG repeats are constrained by CTCF. *Molecular cell*. 2005; 20 (3): 483-9. PMID: 16285929
37. Filippova GN, Thienes CP, Penn BH, Cho DH, Hu YJ, Moore JM, et al. CTCF-binding sites flank CTG/CAG repeats and form a methylation-sensitive insulator at the DM1 locus. *Nature genetics*. 2001; 28 (4): 335-43. PMID: 11479593