Cell, Vol. 111, 151–154, October 18, 2002, Copyright ©2002 by Cell Press

Setting the Boundaries of Chromatin Domains and Nuclear Organization

Minireview

Mariano Labrador and Victor G. Corces[1]
Department of Biology
The Johns Hopkins University
3400 N. Charles Street
Baltimore, Maryland 21218

The nuclear architecture of the interphase nucleus is established by laying down an intricate three-dimensional framework of higher-order chromatin structure. This arrangement is essential for the integration of complex biological processes such as DNA replication, RNA processing, and transcription. Boundary or insulator elements are emerging as key players in the establishment and maintenance of this organization.

The extraordinary task of the eukaryotic cell nucleus in orchestrating gene expression while providing a mechanism to ensure the proper allocation of genetic and epigenetic information after cell division relies on the ability of the nuclear components to transition between two vastly different architectures: interphase and mitosis. The chromatin fiber in metaphase chromosomes is highly condensed and organized into loops attached at the base to a proteinaceous scaffold (Paulson and Laemmli, 1977). After mitosis, the chromatin fiber undergoes a process of decondensation necessary for the activation of transcription and DNA replication. During interphase, however, decondensation of chromatin is far from complete, as a large fraction of the chromatin fiber remains highly condensed. Evidence for highly condensed interphase chromatin interspersed with decondensed chromatin fibers is found in the band-interband pattern of *Drosophila* polytene chromosomes. Interspersed highly condensed chromatin may correspond to silent genes, repetitive DNA such as transposable elements, or large stretches of noncoding intergenic DNA; actively transcribed genes, on the other hand, must be present in regions of relatively decondensed chromatin. The eukaryotic nucleus must then have in place strategies to avoid the spreading of condensed chromatin into regions of actively transcribed genes and vice versa. DNA sequences found in organisms ranging from yeast to humans, known as boundary or insulator elements, display special properties suggestive of a role in the organization of the chromatin fiber.

Boundary and Insulator Elements Might Play a Role in Nuclear Organization

Insulators prevent interactions between enhancers and promoters when located between them, and they buffer transgenes from chromosomal position effects (Gerasimova and Corces, 2001; West et al., 2002). These two properties qualify boundaries or insulators as key candidates in the process of setting up independent domains of gene expression. Evidence for a role of insulators in the establishment of chromatin domains has come in part from the analysis of the gypsy insulator of *Drosophila* (Gerasimova et al., 2000). Protein components of this insulator are present at approximately 500 sites in the *Drosophila* genome. These sites are located at the boundaries between bands and interbands of polytene chromosomes, suggesting a role in separating condensed from decondensed chromatin. These 500 insulator sites coalesce into approximately 25 large structures, named insulator bodies, present mostly in the nuclear periphery of diploid cells. Each insulator body is composed of multiple individual insulator sites, which are brought together by interactions between proteins bound to the insulator DNA. Through this process, the insulators separate the chromatin fiber into loops or domains, forming rosette-like structures. These might be attached to a fixed perinuclear substrate, perhaps the nuclear lamina, which serves as a scaffold to maintain nuclear organization (Figure 1).

Additional support in favor of this idea comes from recent findings by Ishii et al. (2002) suggesting that boundary function in yeast is mediated by interactions with components of the nuclear pore. Ishii and collaborators used the E and I silencers of the yeast silent mating type locus to flank the *ADE2* and *URA3* reporter genes. In a normal situation, the silencers will recruit Sir protein complexes that will repress both genes. They then inserted Gal4 binding UASg sequences flanking the *ADE2* but not the *URA3* gene and screened a library expressing yeast proteins fused to Gal4. When proteins with boundary activity flank the *ADE2* gene, it becomes protected from the spreading of Sir proteins, whereas *URA3*, lying outside the protected domain, is silenced. Surprisingly, the screen and subsequent experiments reveal that proteins involved in nuclear-cytoplasmic transport, such as Cse1p, Mex67p, and Los1p, or the Nup2p receptor of the nuclear pore complex, exhibit strong boundary activity. Although a physiological role for nuclear pore components in the establishment of endogenous boundaries remains to be demonstrated, the results convincingly suggest that tethering DNA to the Nup2p receptor of the nuclear pore, or Nup2p itself, translates into the formation of a domain that is refractory to the spreading of silencing proteins. These results, and those described above for the gypsy insulator of *Drosophila*, suggest that boundary/insulator elements may play a role in the regulation of transcription by organizing the chromatin fiber through the attachment of the DNA to a more or less fixed perinuclear substrate. The yeast results are especially appealing because they suggest a strategy by which the nucleus might arrange actively transcribed chromatin domains in the proximity of nuclear pores, thus ensuring timely transport of newly synthesized RNAs to the cytoplasm.

Boundary/Insulator Elements Delimit Regions of Altered Chromatin Structure

Covalent modifications of histone N-terminal tails have emerged in recent years as critical determinants of the transcriptional status of chromatin (Jenuwein and Allis, 2001). Some of these modifications, such as acetylation, occur not only in promoter regions of genes but also ubiquitously throughout large sections of the genome

[1]Correspondence: corces@jhu.edu

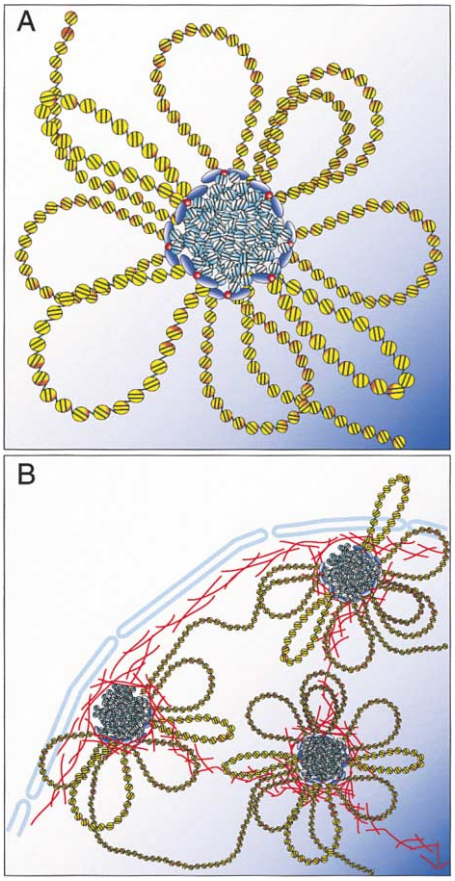

Figure 1. Boundary/Insulator Elements Organize the Chromatin Fiber in the Nucleus by Establishing Separate Compartments of Higher-Order Chromatin Structure

(A) Domains of open chromatin (yellow nucleosomes) are flanked by insulators (purple ovals and red spheres) that interact together to form a loop. Highly condensed chromatin (blue nucleosomes) is restricted to a distinct compartment. Chromatin remodeling and histone-modifying enzymes that contribute to the condensation of the chromatin are abundant in the inner compartment, whereas proteins involved in opening chromatin may be recruited by insulators and are enriched in the outer compartment.

(B) Diagram showing part of a nucleus with compartmentalized chromatin, anchored to the nuclear periphery by interactions of the insulators with the nuclear lamina or nuclear pore complexes.

(Vogelauer et al., 2000). As a consequence, *cis*-regulatory sequences in genes might not be the first level of transcriptional regulation. Rather, regulation of global chromatin structure, specified by factors other than those present at the promoter and regulatory sequences, might be an important prerequisite to specify the final transcriptional estate of a gene. Boundary/insulator elements might play this role by establishing domains of open chromatin whose subsequent maintenance involves global changes in histone modification within the domain.

In yeast, heterochromatin-like structures occur in telomeric regions, rDNA genes, and mating-type loci. We might then expect to find chromatin boundaries flanking domains of silenced chromatin at these regions. To gain insights into the mechanisms by which these different chromatin domains are established, Grewal and collaborators used chromatin immunoprecipitation (ChIP) analysis to create a high-resolution map of chromatin-associated proteins within a 47 kb region containing and surrounding the yeast mating-type locus (Noma et al., 2001). The proteins analyzed included the fission yeast Swi6 protein (the homolog of the *Drosophila* heterochromatin-associated protein HP1), histone H3 methylated at lysine 9 (a histone N-terminal modification that correlates with heterochromatin), and histone H3 methylated at lysine 4 (which correlates with actively transcribed chromatin). Within this 47 kb region, identical inverted repeats containing chromatin boundaries flank the silent region of the mating-type locus, which spans 20 kb and contains the repressed *mat2* and *mat3* genes. The Swi6 protein is present within the silenced region, which also contains histone H3 methylated at Lys9. In contrast, regions of active chromatin outside of the silenced domain lack Swi6 and contain histone H3 methylated at Lys4 instead. These results suggest that boundary elements can separate regions of biochemically different chromatin structure, with different protein composition and histone tail modifications.

Heterochromatin regions flanking transcriptionally active chromatin are more abundant in higher eukaryotes than in yeast, offering an even more complex scenario of the interplay between active and inactive chromatin regions in the genome. A particularly well-studied case is the chicken β-globin locus, which is flanked by a ~16 kb block of condensed chromatin on one side and an odorant receptor gene on the other. Fensenfeld and collaborators have mapped two DNase I hypersensitive sites (5′HS4 and 3′HS) that correspond to insulator sequences containing binding sites for the CTCF protein and flank the β-globin domain (Litt et al., 2001a, 2001b). The 5′HS4 insulator delimits the block of condensed chromatin, and the 3′HS insulator is interposed between the β-globin domain and the odorant receptor gene. Using ChIP analysis, Litt and colleagues have determined the distribution of acetylated histones H3 and H4 (multiacetylated H3, tetra acetylated H4, and acetylated H4 at lysines 5, 8, or 12) and methylated H3 at lysines 9 and 4 along 55 kb spanning the β-globin domain and flanking chromatin regions, before and after activation of transcription in the locus. Results show that the global levels of H3 and H4 acetylation and H3 Lys4 methylation increase in the region located between the two CTCF insulator sites during the developmental stage at which activation of β-globin transcription takes place. The heterochromatic block, however, is always enriched in Lys9-methylated H3, whereas acetylation of H3 and H4 is low in this region. A recent report by Mutskov and collaborators (Mutskov et al., 2002) further reinforces the idea that boundary/insulator elements have a role in establishing domains of open chromatin characterized by global changes in histone modification. Using transgenes flanked by the β-globin insulators, the authors show that insulators are required to maintain high levels of histone acetylation independent of the transcriptional state of the gene or the presence of active enhancers in the domain. Interestingly, the mouse and human β-globin loci lack the heterochromatic block present in the chicken β-globin locus. Instead, several

olfactory receptor genes are present on both sides of the domain, also flanked by CTCF insulators (Farrell et al., 2002). The lack of heterochromatin adjacent to the mammalian locus suggests that, whereas the role of the insulator in chicken is to maintain an open chromatin estate to allow transcription of β-globin genes, the role of the insulators in human and mouse may be, in addition, to prevent regulatory sequences from one domain to activate transcription in an unrelated domain.

Boundary/Insulator Elements Might Establish Functional Domains of Gene Expression

If insulators define blocks of transcriptionally active chromatin, one may anticipate that genes residing within these domains would be transcriptionally coregulated. Work in recent years has addressed this issue using a genomic approach to integrate near whole-genome expression patterns with the location of genes in the chromosome. The results suggest that organisms as divergent as yeast, *Drosophila*, and humans exhibit local clustering of genes that are coexpressed at the same time and in the same tissues. Using correlation maps of gene expression for all ORFs from each chromosome, data obtained in yeast show that adjacent genes and nonadjacent genes located nearby tend to display correlated expression patterns (Cohen et al., 2000). Although coexpression was higher between pairs of genes that share the same promoter regions, coregulation was still significant between pairs of genes in tandem, in a convergent orientation, or with unrelated UAS regulatory regions. Roughly, 19% of 2018 coexpressed genes turned out to be related in function, suggesting that there is a certain trend for genes with similar functions to be present in the same region of a chromosome. An interpretation of the data may be that, in the absence of other barriers, global chromatin structure may coordinate transcription of neighboring genes, probably favoring a trend that could bring together, within a chromatin domain, genes with related functions over evolutionary time.

One could argue that this hypothesis would only be appropriate for compact genomes such as that of yeast. This argument has been tested by analyzing coexpression of genes in the genome of flies and humans, where intergenic regions are significantly longer than in yeast. Surprisingly, the same functional organization is observed in genomes from these two organisms. Coexpression profile analyses using high-density oligonucleotide microarrays covering 80 different experimental conditions show that 20% of *Drosophila* genes are found as part of one of the 200 groups of adjacent genes that are similarly expressed (Spellman and Rubin, 2002). Each one of these groups contains between 10 and 30 members. Although the clustering is similar to that found in yeast, the authors did not find any functional relationship between the genes present in each cluster, further favoring the hypothesis that open chromatin structure may trigger the activation of all genes in a domain. Interestingly, the human genome also shows significant clustering of coexpressed genes, although in this case, after removing duplicated genes from the domains, the only significant coregulated clusters correspond to housekeeping genes (Caron et al., 2001; Lercher et al., 2002).

The genome-wide expression data suggest that domains of gene expression are persistent across the phylogenetic scale. Clustering of genes may obey different nonexclusive evolutionary strategies. For example, genes with similar function might tend to cluster together during evolution, as it seems to occur in yeast. In other cases, genes with similar patterns of expression, such as housekeeping genes or genes for which a low level of expression is not deleterious for the cell, might tend to share the same domain, as occurs in humans and flies. In addition, the data show that tissue-specific genes in humans do not cluster together, leaving still open the question of how these genes preserve their transcriptional specificity and that of their neighbors. A possible explanation is that all regulatory elements with a strong effect on the neighboring chromatin, such as enhancers, silencers, or heterochromatin, are compart-

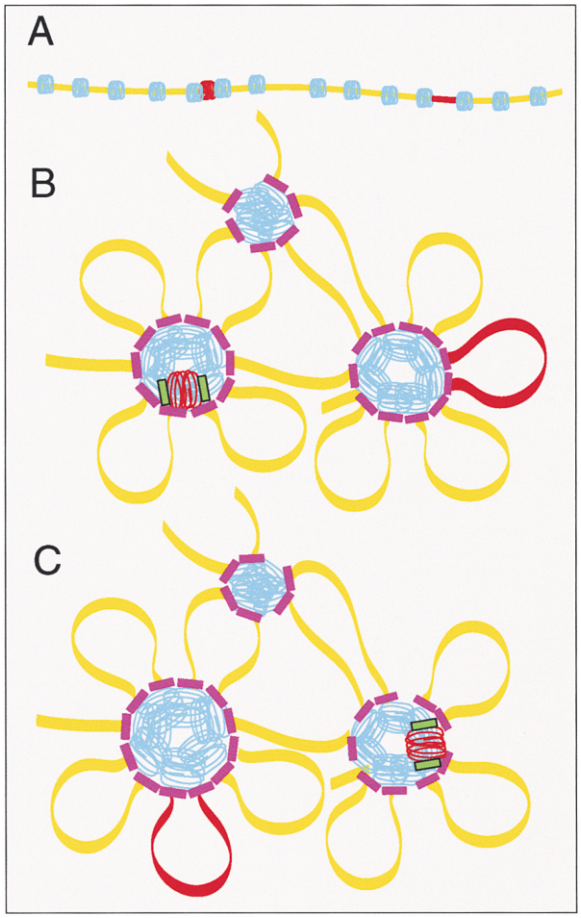

Figure 2. Regulation of Boundary/Insulator Function Gives Rise to Different Patterns of Chromatin Organization

(A) Linear layout of interphase chromatin. Highly condensed chromatin is shown in blue and open chromatin domains are shown in yellow. Domains with regulatable insulators, whose activity can be changed during cell differentiation, are shown in red.

(B) During development, domains of higher-order chromatin structure are organized by active insulators (purple squares). Inactive insulators and the domain they flank (green squares) remain in the heterochromatin compartment.

(C) In a particular tissue, a chromatin domain becomes open after activation of the flanking insulators, and a different domain becomes heterochromatic after insulator inactivation.

mentalized in the nucleus in domains defined by insulators that do not interfere with each other. In either case, the arrangement of the interphase chromatin into loop domains, predicted by the distribution of *Drosophila* gypsy insulator sites and the attachment of yeast boundaries to the nuclear pore complex, suggest a mechanism by which functional groups of transcriptionally coregulated genes could be established.

Can Boundary/Insulators Be Regulated?

Specific histone modifications inside the β-globin domain, flanked by CTCF insulators, are induced during development of erythroid cells as a function of the transcriptional state of the β-globin genes. Such developmental changes in a chromatin domain suggest the possibility that the function of boundary elements could also be regulatable, therefore providing a new regulatory level for the control of gene expression (Figure 2). A well-documented instance of regulatable activity of insulators is found in the epigenetic control of transcription of the *H19* and *Igf2* genes in mammals (Bell and Felsenfeld, 2000; Hark et al., 2000). The imprinting control region (ICR), present between *Igf2* and shared enhancers located 3′ to the *H19* gene, contains insulator sequences that bind the CTCF protein. In the paternal chromosome, ICR sequences are methylated, inhibiting the binding of CTCF to the insulator. The paternal *Igf2* gene is then activated by the enhancers located downstream of *H19*. In the same paternal chromosome, the *H19* gene is not transcribed due to methylation of the promoter. This scenario is reversed in the maternal chromosome, where the ICR is not methylated, allowing binding of CTCF. This prevents the enhancers from activating *Igf2* and allows the activation of *H19*. In this case, methylation of insulator sequences controls binding of the CTCF protein and, consequently, insulator activity.

The potential ability of insulators to determine the transcriptional state and local chromatin structure in the interphase nucleus is highlighted by recent results suggesting the possibility that insulators may also help to discriminate between the two female X chromosomes as part of the mechanism that leads to X chromosome inactivation in mammals (Chao et al., 2002). Recognition sites for the CTCF protein are found in the mouse Xist locus, suggesting the possibility that binding of CTCF to these sites determines the choice between active and inactive X chromosomes. Hypothetically, binding of CTCF to insulator sequences will prevent downstream enhancers from activating transcription of Xist and will allow the transcription of Tsix, an RNA transcript complementary to Xist. In the model, activation of Tsix also contributes to transcriptional repression of Xist, whereas DNA methylation of the CTCF binding sites in the inactive X chromosome prevents Tsix from being expressed, allowing transcription of Xist and initiation of X inactivation. The potential to regulate CTCF insulators by DNA methylation suggests the possibility that X inactivation may be triggered by the inactivation of all the CTCF boundaries in the X chromosome and the subsequent spreading of heterochromatin, no longer contained in the repressed domains, to the complete chromosome. In this case, as in the imprinting of the *H19* and *Igf2* genes, insulator activity can be regulated by DNA methylation.

Concluding Remarks

Boundary or insulator elements might establish domains of gene expression by translating linear information contained in the chromatin fiber into a three-dimensional structure capable of compartmentalizing the genome (Figure 1). This compartmentalization is also suggested by the presence of insulators in the band-interband boundaries of *Drosophila* polytene chromosomes, indicating that insulators may help organize chromatin into discontinuous local compartments of condensed and decondensed chromatin (Figure 2). Analysis of histone modifications within insulated domains suggests that these sequences might function by promoting or allowing the activity of histone-modifying enzymes that contribute to the maintenance of open chromatin. The enzymatic activity of these proteins might prevent the spreading of heterochromatin proteins from one compartment into the other. This chromatin organization reflects the apparent clustering of actively expressed genes in eukaryotic genomes and may explain the enhancer-blocking properties of insulators. Insulators might interfere with enhancer function by inhibiting changes in DNA topology or histone modification. The directionality of this interference with enhancer-promoter communication can be explained by their separation into topologically independent domains. Attachment of insulators to the nuclear lamina or nuclear pore complexes may provide the scaffold necessary to facilitate the barrier function.

Selected Reading

Bell, A.C., and Felsenfeld, G. (2000). Nature *405*, 482–485.

Caron, H., van Schaik, B., van der Mee, M., Baas, F., Riggins, G., van Sluis, P., Hermus, M.C., van Asperen, R., Boon, K., Voute, P.A., et al. (2001). Science *291*, 1289–1292.

Chao, W., Huynh, K.D., Spencer, R.J., Davidow, L.S., and Lee, J.T. (2002). Science *295*, 345–347.

Cohen, B.A., Mitra, R.D., Hughes, J.D., and Church, G.M. (2000). Nat. Genet. *26*, 183–186.

Farrell, C.M., West, A.G., and Felsenfeld, G. (2002). Mol. Cell. Biol. *22*, 3820–3831.

Gerasimova, T.I., and Corces, V.G. (2001). Annu. Rev. Genet. *35*, 193–208.

Gerasimova, T.I., Byrd, K., and Corces, V.G. (2000). Mol. Cell *6*, 1025–1035.

Hark, A.T., Schoenherr, C.J., Katz, D.J., Ingram, R.S., Levorse, J.M., and Tilghman, S.M. (2000). Nature *405*, 486–489.

Ishii, K., Arib, G., Lin, C., Van Houwe, G., and Laemmli, U.K. (2002). Cell *109*, 551–562.

Jenuwein, T., and Allis, C.D. (2001). Science *293*, 1074–1080.

Lercher, M.J., Urrutia, A.O., and Hurst, L.D. (2002). Nat. Genet. *31*, 180–183.

Litt, M.D., Simpson, M., Gaszner, M., Allis, C.D., and Felsenfeld, G. (2001a). Science *293*, 2453–2455.

Litt, M.D., Simpson, M., Recillas-Targa, F., Prioleau, M.N., and Felsenfeld, G. (2001b). EMBO J. *20*, 2224–2235.

Mutskov, V.J., Farrell, C.M., Wade, P.A., Wolffe, A.P., and Felsenfeld, G. (2002). Genes Dev. *16*, 1540–1554.

Noma, K., Allis, C.D., and Grewal, S.I. (2001). Science *293*, 1150–1155.

Paulson, J.R., and Laemmli, U.K. (1977). Cell *12*, 817–828.

Spellman, P.T., and Rubin, G.M. (2002). J. Biol. *1*, 5.

Vogelauer, M., Wu, J., Suka, N., and Grunstein, M. (2000). Nature *408*, 495–498.

West, A.G., Gaszner, M., and Felsenfeld, G. (2002). Genes Dev. *16*, 271–288.