The role of CTCF in regulating nuclear organization

Adam Williams and Richard A. Flavell

The spatial organization of the genome is thought to play an important part in the coordination of gene regulation. New techniques have been used to identify specific long-range interactions between distal DNA sequences, revealing an ever-increasing complexity to nuclear organization. CCCTC-binding factor (CTCF) is a versatile zinc finger protein with diverse regulatory functions. New data now help define how CTCF mediates both long-range intrachromosomal and interchromosomal interactions, and highlight CTCF as an important factor in determining the three-dimensional structure of the genome.

A three-dimensional view of gene regulation

At the microscopic level, the nucleus is grossly partitioned into regions of heterochromatin and euchromatin, corresponding to inactive and active portions of the genome, respectively (1). In addition, each chromosome appears to be positioned nonrandomly, adopting preferred nuclear positions relative to each other in certain cell types (2). Interestingly, gene activation or silencing is frequently accompanied by locus relocalization within the nucleus (3–5), suggesting that the spatial organization of chromatin may be important for the regulation of gene expression.

Recently developed techniques, such as 3C (chromosome conformation capture) and RNA-TRAP (tagging and recovery of associated proteins) identify interacting regions of chromatin and reveal its overall three-dimensional organization within the nucleus (6, 7). In 3C, formaldehyde–cross-linked chromatin is digested with restriction enzymes and then ligated under conditions that favor the ligation of cross-linked fragments, which can then be detected by polymerase chain reaction (PCR) and sequenced. In RNA-TRAP, in situ hybridization is used to target horseradish peroxidase activity to primary transcripts associated with a transcribed gene. Localized peroxidase activity catalyses the covalent attachment of a biotin tag to nearby chromatin, which after purification on streptavidin agarose beads can be analyzed by PCR to determine the presence of specific interactions.

The use of these techniques has identified specific interactions between distal genomic sequences, revealing another level of nuclear organization (6–12). Although many of these interactions are likely to be coincidental and are driven by the need to share common resources, as with the colocalization of genes within transcription factories (13), several of these long-range interactions have been shown to possess important biological functions (8, 11, 12). These new findings have expanded our understanding of the mechanisms regulating gene expression, and have led to the replacement of the traditional linear model of gene regulation with a three-dimensional model in which long-range intrachromosomal and interchromosomal associations orchestrate and coordinate gene expression. In this model, distal regulatory elements physically interact both in cis and in trans with the genes that they control. Perhaps most surprisingly, regulatory elements on one chromosome can also directly regulate the expression of genes on other chromosomes (8, 11, 12). These new concepts might have potential clinical importance, as they provide an explanation for the preferential chromosomal translocation partners observed in certain cancers (14). On page 785 in this issue, Majumder et al. (15) help explain how CTCF might orchestrate these long-range interactions.

CTCF as a mediator of long-range interactions

A three-dimensional model of gene regulation raises the question of which factors control nuclear organization and long-range interactions. One potential candidate is CTCF, a ubiquitously expressed transcription factor that has multiple context-dependent functions. CTCF can act as an enhancer-blocking protein, can bind to boundary elements to prevent spreading of heterochromatin, and can function as both a transcriptional activator and a transcriptional silencer. How CTCF carries out these diverse functions is not completely clear, but recent work from several groups has revealed that CTCF may function through topological organization of the genome (Fig. 1). One study, for example, suggested that CTCF regulates gene expression by inducing the formation of long-range chromatin loops (16). Two years later, the first direct evidence for CTCF-mediated chromatin looping was revealed from analysis of the β-globin locus (17).

CTCF was later found to mediate interchromosomal interactions between the *Igf2/H19* imprinting control region (ICR) on chromosome 7 and the *Wsb1/Nf1* gene complex on chromosome 11 (12). Deletion of the *Igf2/H19* ICR or abrogation of CTCF expression disrupted this interchromosomal association and altered *Wsb1/Nf1* gene expression (12). More recently, CTCF has been implicated in driving X chromosome homologous pairing during the process of X chromosome

A.W and R.A.F. are at the Department of Immunobiology and Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, CT 06520

CORRESPONDENCE
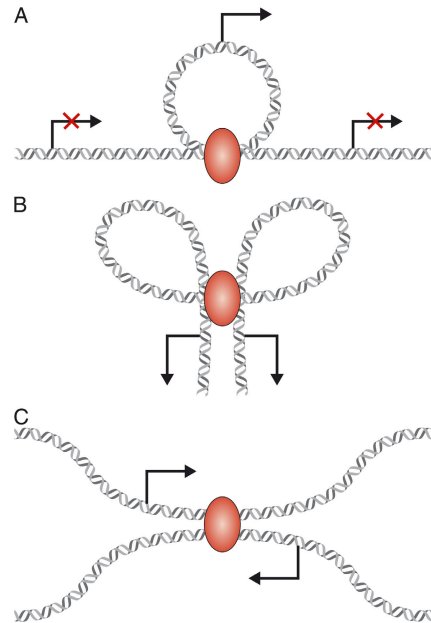R.A.F.: richard.flavell@yale.edu

Figure 1. Potential mechanisms of CTCF–mediated genome organization. (A) CTCF (represented as a single ellipsoid for simplicity) could fold the genome into loop domains, thereby isolating genes from the influence of their neighbors. (B) Alternatively, CTCF could fold chromatin to bring distal regulatory elements or coregulated genes closer together to form a chromatin hub or holocomplex. (C) Similarly, CTCF could link genes or regulatory elements on different chromosomes to mediate interchromosomal regulation.

inactivation (18). Collectively, these results suggest that CTCF mediates at least some of its gene regulatory functions through the three-dimensional organization of the genome.

The growing interest in CTCF is highlighted in the study by Majumder et al. (15), which describes a new model of major histocompatibility complex (MHC) class II gene regulation that is mediated by CTCF-dependent, long-range intrachromosomal interactions. Expression of two divergently transcribed MHC class II genes, *HLA-DRB1* and *HLA-DQA1*, is driven by an intervening nuclear matrix–bound enhancer element called *XL9* (19). CTCF had been shown to bind to this enhancer element, but its effect on the expression of the two class II genes had not been investigated (19). In the new study, the authors find evidence for long-range chromatin loops between the promoters of the *HLA-DRB1* and *HLA-DQA1* genes and the intergenic *XL9* enhancer. These interactions depended on a complex consisting of CTCF, the transcription factor RFX, and the transactivator

CIITA. Knocking down CTCF using RNAi reduced the long-range interactions between the *XL9* enhancer element and the MHC class II genes and decreased expression of *HLA-DRB1* and *HLA-DQA1*. These findings provide a novel model for MHC class II expression, and also provide insight into several unanswered questions about the biology of CTCF, such as how these long-range interactions are mediated and how the tissue-specific functions of CTCF are regulated.

How does CTCF mediate long–range interactions?
Because CTCF can form dimers and maybe even oligomers, it is possible that CTCF molecules bound to distal elements could interact with each other, thereby driving loop formation or interchromosomal interactions (16). However, in a study by Ling et al. (12), the presence of CTCF on both interacting loci was not an absolute requirement for interchromosomal interaction, suggesting either that CTCF is part of a multiprotein bridging complex or that CTCF

recruits additional factors that then drive long-range interactions. Indeed, Majumder et al. (15) showed that CTCF requires at least two additional factors to function, at least in the control of the *HLA-DRB1* and *HLA–DQA1* genes. CTCF, RFX, and CIITA formed a complex even in the absence of DNA, and the loss of any one of these factors abolished long-range interactions with the intergenic enhancer element, precluding MHC class II expression.

Whether RFX and CIITA form a structural part of a bridging complex or whether they simply recruit CTCF and/or additional factors to the *XL9* enhancer remains to be resolved. It will be interesting to purify these and other bridging complexes and characterize their protein composition. Interestingly, recent data show that CTCF is able to recruit cohesins to specific genomic locations (20–22). Cohesins, better known for their role in mediating sister chromatid cohesion during mitosis, have proven chromatin bridging potential and are thus an obvious candidate for spatial organization of the genome.

How are these interactions controlled?
As CTCF is a ubiquitous factor, its function cannot be regulated by differential expression. Thus, long-range intra- and interchromosomal interactions must be regulated on other levels. This control might involve direct blocking of CTCF binding by developmentally regulated or allele-specific DNA methylation. The mono-allelic nature of the interaction between the *Igf2/H19* ICR and the *Wsb1/Nf1* locus, for example, appears to be conferred by imprinted DNA methylation of the noninteracting loci (12).

As CTCF may require the presence of other protein factors in some contexts, another level of control might involve regulation of the expression of these factors or their ability to interact with CTCF. Indeed, Majumder et al. (15) found that interactions between XL9 and the *HLA-DRB1* and *HLA-DQA1* genes could be induced in A431 epithelial cells by interferon γ treatment, which is essential for CIITA expression in these cells. In untreated A431 cells, CIITA is not expressed, precluding

long-range interactions between the *HLA-DRB1* and *HLA-DQA1* genes and the *XL9* enhancer and thus MHC class II expression.

The model proposed by Majumder et al. (15) predicts that several chromatin organization states might exist for the MHC class II locus. For example, the MHC class II locus in B cells, which express CTCF, RFX, and CIITA, is in an active state with long-range interactions between the *XL9* enhancer and the *HLA-DRB1* and *HLA-DQA1* genes driving constitutive class II expression. In contrast, most nonimmune cells do not express CIITA or MHC class II. In these cells, however, regulatory elements in the MHC class II locus are still bound by transcription factors such as RFX, CREB, and NF-Y (23, 24). And these silent MHC class II promoters display histone modifications normally associated with accessible chromatin (25), suggesting that in cells lacking CIITA, the MHC class II genes are poised and ready for class II expression. According to the data from Majumder et al. (15), the induction of CTIIA expression in these cells would lead to recruitment of CTCF to the *XL9* enhancer, the formation of long-range interactions between the *XL9* enhancer and the *HLA-DRB1* and *HLA-DQA1* genes, and initiation of MHC class II expression. However, it must be noted that in the absence of genetic evidence, it is not possible to rule out an alternative hypothesis in which the formation of long-range interactions is a consequence (rather than the cause) of MHC class II transcription.

Is CTCF a universal regulator?

The number of CTCF binding sites in the human genome was recently estimated to be between 14,000 and 15,000, and it is likely that a comparable number are present within the mouse genome (26, 27). Interestingly, most of these sites appear to be occupied in a cell type–independent manner (27). It is therefore tempting to speculate that CTCF is important for the maintenance of nuclear architecture and, by default, the regulation of many genes. This does not appear to be the case, however, according

to Majumder et al. (15), as knocking down CTCF did not result in global changes in gene expression. Thus, it is highly likely that other factors also serve to regulate long-range interactions and nuclear structure. The protein SATB1, for example, has been implicated in the regulation of thymocyte nuclear architecture and the initiation of cytokine expression by directing long-range interactions within the T helper 2 cytokine locus on mouse chromosome 11 (28, 29). Other factors, such as cohesins, MeCP2, and MENT, which have been shown to have chromatin bridging potential, may also help organize the three-dimensional structure of the nucleus (20–22, 30, 31).

Future perspectives

Further characterization of MHC class II gene expression may reveal additional CTCF-mediated long-range interactions within the MHC class II gene complex. Indeed, the MHC class II gene *HLA-DRA* has previously been shown to form DNA loops that bring other regulatory elements into close proximity with its promoter sequences (25). Additional CTCF binding sites have also been identified within the MHC class II locus, highlighting the potential for a general role for CTCF in regulating MHC class II expression (27, 32).

From a more global perspective, long-range interactions and interchromosomal associations are likely to be widespread within the nucleus, meriting genome-wide analyses to define the nuclear "interactome" (11, 33–35). The task will then be to sort through these data to identify which interactions are functional and which are coincidental. In addition, isolation and characterization of chromatin bridging complexes, such as those containing CTCF, may reveal additional players in long-range regulation. One of the most interesting questions that needs to be addressed is whether these long-range interactions are stable, static structures or ephemeral complexes that form from transient interactions.

REFERENCES

1. Dillon, N. 2004. Heterochromatin structure and function. *Biol. Cell.* 96:631–637.

2. Parada, L., and T. Misteli. 2002. Chromosome positioning in the interphase nucleus. *Trends Cell Biol.* 12:425–432.

3. Hewitt, S.L., F.A. High, S.L. Reiner, A.G. Fisher, and M. Merkenschlager. 2004. Nuclear repositioning marks the selective exclusion of lineage-inappropriate transcription factor loci during T helper cell differentiation. *Eur. J. Immunol.* 34:3604–3613.

4. Kosak, S.T., J.A. Skok, K.L. Medina, R. Riblet, M.M. Le Beau, A.G. Fisher, and H. Singh. 2002. Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science.* 296:158–162.

5. Delaire, S., Y.H. Huang, S.W. Chan, and E.A. Robey. 2004. Dynamic repositioning of CD4 and CD8 genes during T cell development. *J. Exp. Med.* 200:1427–1435.

6. Carter, D., L. Chakalova, C.S. Osborne, Y.F. Dai, and P. Fraser. 2002. Long-range chromatin regulatory interactions in vivo. *Nat. Genet.* 32:623–626.

7. Dekker, J., K. Rippe, M. Dekker, and N. Kleckner. 2002. Capturing chromosome conformation. *Science.* 295:1306–1311.

8. Spilianakis, C.G., M.D. Lalioti, T. Town, G.R. Lee, and R.A. Flavell. 2005. Interchromosomal associations between alternatively expressed loci. *Nature.* 435:637–645.

9. Spilianakis, C.G., and R.A. Flavell. 2004. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat. Immunol.* 5:1017–1027.

10. Tolhuis, B., R.J. Palstra, E. Splinter, F. Grosveld, and W. de Laat. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell.* 10:1453–1465.

11. Lomvardas, S., G. Barnea, D.J. Pisapia, M. Mendelsohn, J. Kirkland, and R. Axel. 2006. Interchromosomal interactions and olfactory receptor choice. *Cell.* 126:403–413.

12. Ling, J.Q., T. Li, J.F. Hu, T.H. Vu, H.L. Chen, X.W. Qiu, A.M. Cherry, and A.R. Hoffman. 2006. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science.* 312:269–272.

13. Osborne, C.S., L. Chakalova, K.E. Brown, D. Carter, A. Horton, E. Debrand, B. Goyenechea, J.A. Mitchell, S. Lopes, W. Reik, and P. Fraser. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* 36:1065–1071.

14. Osborne, C.S., L. Chakalova, J.A. Mitchell, A. Horton, A.L. Wood, D.J. Bolland, A.E. Corcoran, and P. Fraser. 2007. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. *PLoS Biol.* 5:e192.

15. Majumder, P., J.A. Gomez, B.P. Chadwick, and J.M. Boss. 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.

16. Yusufzai, T.M., H. Tagami, Y. Nakatani, and G. Felsenfeld. 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared

insulator mechanisms across species. *Mol. Cell.* 13:291–298.

17. Splinter, E., H. Heath, J. Kooren, R.J. Palstra, P. Klous, F. Grosveld, N. Galjart, and W. de Laat. 2006. CTCF mediates long-range chromatin looping and local histone modification in the beta–globin locus. *Genes Dev.* 20:2349–2354.

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

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

20. Parelho, V., S. Hadjur, M. Spivakov, M. Leleu, S. Sauer, H.C. Gregson, A. Jarmuz, C. Canzonetta, Z. Webster, T. Nesterova, et al. 2008. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell.* 132:422–433.

21. Wendt, K.S., K. Yoshida, T. Itoh, M. Bando, B. Koch, E. Schirghuber, S. Tsutsumi, G. Nagae, K. Ishihara, T. Mishiro, et al. 2008. Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature.* 451:796–801.

22. Stedman, W., H. Kang, S. Lin, J.L. Kissil, M.S. Bartolomei, and P.M. Lieberman. 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.

23. Kara, C.J., and L.H. Glimcher. 1991. In vivo footprinting of MHC class II genes: bare promoters in the bare lymphocyte syndrome. *Science.* 252:709–712.

24. Beresford, G.W., and J.M. Boss. 2001. CIITA coordinates multiple histone acetylation modifications at the HLA-DRA promoter. *Nat. Immunol.* 2:652–657.

25. Gomez, J.A., P. Majumder, U.M. Nagarajan, and J.M. Boss. 2005. X box–like sequences in the MHC class II region maintain regulatory function. *J. Immunol.* 175:1030–1040.

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

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

28. Cai, S., H.J. Han, and T. Kohwi-Shigematsu. 2003. Tissue-specific nuclear architecture and gene expression regulated by SATB1. *Nat. Genet.* 34:42–51.

29. Cai, S., C.C. Lee, and T. Kohwi-Shigematsu. 2006. SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat. Genet.* 38:1278–1288.

30. Springhetti, E.M., N.E. Istomina, J.C. Whisstock, T. Nikitina, C.L. Woodcock, and S.A. Grigoryev. 2003. Role of the M-loop and reactive center loop domains in the folding and bridging of nucleosome arrays by MENT. *J. Biol. Chem.* 278:43384–43393.

31. Georgel, P.T., R.A. Horowitz-Scherer, N. Adkins, C.L. Woodcock, P.A. Wade, and J.C. Hansen. 2003. Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J. Biol. Chem.* 278:32181–32188.

32. Arnold, R., W. Maueler, G. Bassili, M. Lutz, L. Burke, T.J. Epplen, and R. Renkawitz. 2000. The insulator protein CTCF represses transcription on binding to the (gt)(22)(ga)(15) microsatellite in intron 2 of the HLA-DRB1(*)0401 gene. *Gene.* 253:209–214.

33. Simonis, M., P. Klous, E. Splinter, Y. Moshkin, R. Willemsen, E. de Wit, B. van Steensel, and W. de Laat. 2006. Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* 38:1348–1354.

34. Zhao, Z., G. Tavoosidana, M. Sjolinder, A. Gondor, P. Mariano, S. Wang, C. Kanduri, M. Lezcano, K.S. Sandhu, U. Singh, et al. 2006. Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and inter-chromosomal interactions. *Nat. Genet.* 38:1341–1347.

35. Wurtele, H., and P. Chartrand. 2006. Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended Chromosome Conformation Capture methodology. *Chromosome Res.* 14:477–495.