Chromatin Boundaries and Chromatin Domains

G. Felsenfeld, B. Burgess-Beusse, C. Farrell, M. Gaszner, R. Ghirlando, S. Huang, C. Jin, M. Litt, Frédérique Magdinier, V. Mutskov, et al.

Chromatin Boundaries and Chromatin Domains

G. Felsenfeld,[*] B. Burgess-Beusse,[*] C. Farrell,[*] M. Gaszner,[*] R. Ghirlando,[*] S. Huang,[*] C. Jin,[*] M. Litt,[*] F. Magdinier,[*] V. Mutskov,[*] Y. Nakatani,[†] H. Tagami,[†] A. West,[*‡] and T. Yusufzai[*]

[*]*Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892-0540;*
[†]*Dana Farber Cancer Institute, Boston, Massachusetts 02115*

Insulator elements were first described in *Drosophila*, but subsequent studies have shown that they are present in vertebrates as well (for review, see West et al. 2002). Over the past several years we have focused our attention on the properties of an insulator at the 5′ end of the chicken β-globin locus that has begun to provide an understanding of how such elements function. This work, as well as studies in other laboratories, has revealed that there are two distinct kinds of insulator activities, which are different in their function. The first of these is the *enhancer-blocking activity*, which can prevent interaction between a distal enhancer and a promoter when placed between them (Fig. 1A). This has the effect of preventing an incorrect interaction between regulatory elements in adjacent, but separately regulated, gene systems. The second insulator function is connected with *barrier activity*, which prevents condensed heterochromatin from extending into adjacent chromatin domains carrying transcriptionally active genes (Fig. 1B).

The chicken β-globin locus extends over 30 kb. It contains four members of the globin gene family, with different programs of expression during development (Fig. 2), which have been studied extensively (Felsenfeld 1993). Strong positive regulatory elements, components of the locus control region (LCR), are distributed both upstream of the gene cluster and within it. Further upstream is a DNase I "hypersensitive site," 5′HS4, which, unlike others in the locus, is not erythroid-specific, but is nuclease sensitive in all cells that have been tested (Reitman and Felsenfeld 1990). It seemed an attractive possibility that this marked the 5′ end of the open chromatin domain; in fact, it was shown not long afterward that immediately upstream of 5′HS4 there is an abrupt decrease of nuclease sensitivity and histone acetylation in globin-expressing cells, consistent with a transition from the open chromatin of the globin locus to a more inactive, condensed chromatin structure (Hebbes et al. 1994). We explored the possibility that this element might have the properties of an insulator (Chung et al. 1997).

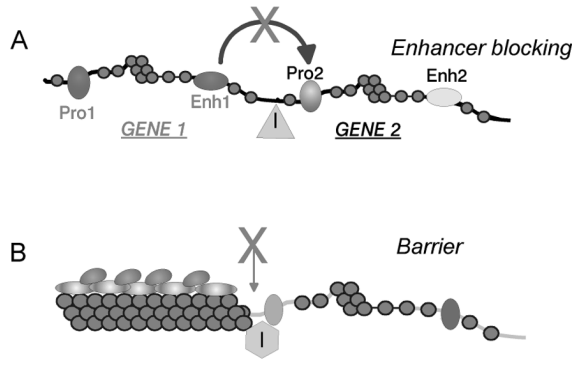

Figure 1. Two kinds of insulator function. (*A*) Enhancer blocking: The insulator (I) prevents Enhancer 1, belonging to gene system 1, from acting inappropriately on Promoter 2 in gene system 2. (*B*) Barrier function: The insulator (I) prevents the advance of heterochromatic regions on the *left* into transcriptionally active chromatin on the *right*.

ENHANCER-BLOCKING INSULATION AND CTCF

The analysis of the gypsy element in *Drosophila* had provided the first example of enhancer-blocking action, and we began our studies by testing whether a 1.2-kb fragment containing 5′HS4 could similarly prevent enhancer–promoter interaction. The assay we devised placed the element between a strong erythroid-specific enhancer and promoter that were driving expression of a drug resistance gene. The number of colonies able to grow under selective conditions was used as a measure of enhancer-blocking activity. We showed that the 1.2-kb element reduced colony number by an order of magnitude, but only when placed between the enhancer and promoter (Chung et al. 1997). A 250-bp "core" element derived from this fragment retained the enhancer-blocking function and allowed us to dissect the activity further. We report here on subsequent results that followed from these observations.

DNase footprinting experiments with nuclear extracts showed that there were five discrete protected regions within the 250-bp core, which could be tested individually for enhancer-blocking activity (Fig. 3) (Chung et al. 1997). Footprint II was necessary and sufficient for activity in our assay and provided a DNA sequence that could be used in gel retardation assays to follow purification of the protein that bound to it. This protein, CTCF, had been described earlier, but had not been associated with insulator activity (Bell et al. 1999).

[‡]Present address: Division of Cancer Sciences and Molecular Pathology, Gene Regulation and Mechanisms of Disease Section, Pathology Department, Glasgow University, Glasgow G11 6NT, Scotland, United Kingdom.

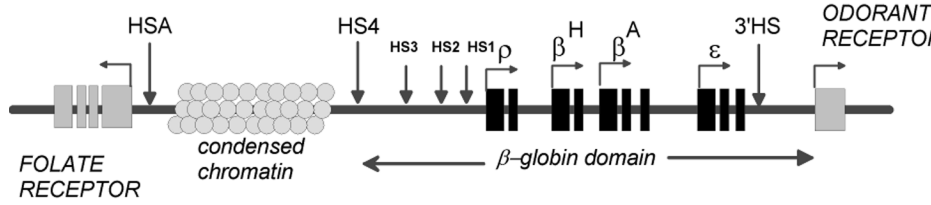

Figure 2. A map (not to scale) of the chicken β-globin domain and its neighbors. The distance between HS4, the globin 5′ insulator element, and 3′HS, the CTCF site at the 3′ end of the domain, is about 30 kb.

CTCF is a member of the zinc finger family, with the unusually large number of 11 fingers. Since its identification as the active agent in the enhancer-blocking activity of the 5′HS4 chicken β-globin element, it has been implicated in similar functions within a wide variety of vertebrate loci, including the 3′ end of the chicken β-globin locus (Saitoh et al. 2000). CTCF binding sites with enhancer-blocking action have been found within the mouse and human β-globin loci (Farrell et al. 2002; Tanimoto et al. 2003), and associated with the Tsix antisense gene in the mouse choice/imprinting center involved in control of X chromosome inactivation (Chao et al. 2002). Recent studies have identified a single CTCF site (Fig. 4) located between the mouse T cell receptor α (*TCR* α) gene and the downstream *Dad 1* gene, where it may play a role in preventing inappropriate interaction between enhancers committed to one gene and promoter controlling the other (Magdinier et al. 2004).

Perhaps the most convincing evidence for the biological role of such insulators is provided by studies of the regulation of the mouse and human *Igf2/H19* imprinted locus (Bell and Felsenfeld 2000; Hark et al. 2000; Kanduri et al. 2000). The *H19* gene is expressed only from the maternally transmitted allele, and *Igf2* only from the paternal (Fig. 5). *Igf2* expression is shut down on the maternal allele because of the presence of an insulator element with multiple binding sites for CTCF, which blocks the action of downstream endodermal enhancers that normally activate *Igf2*. These DNA sites are methylated on the imprinted paternal allele, which prevents CTCF binding, inactivates the insulator, and activates *Igf2* expression.

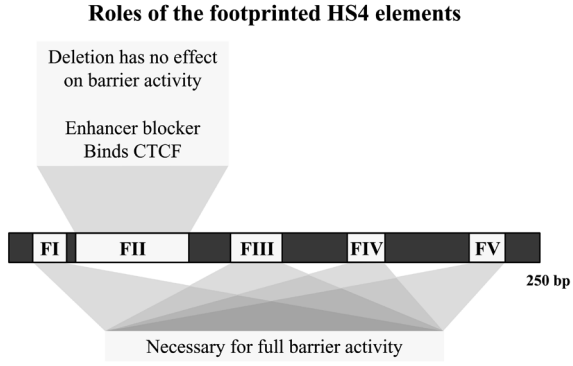

Figure 3. The five footprinted regions of the β-globin HS4 "core" insulator element, showing the separate contributions of each element to enhancer-blocking or barrier activity.

Since it seemed unlikely that CTCF could carry out this function without the intervention of other proteins, the next step was to search for factors with which it interacted. We expressed a double epitope-tagged version of the protein and immunopurified complexes containing candidate cofactors (Yusufzai et al. 2004). Among the proteins purified in this way were poly ADP-ribose polymerase, Set1, H2A, H2A.Z, and, most notably, the nucleolar protein, nucleophosmin, which was present in the greatest abundance. When a singly purified complex was centrifuged on a glycerol gradient, some of the CTCF cosedimented with nucleophosmin in two fractions, indicating that a well-defined complex had formed between them. Chromatin immunoprecipitation (ChIP) using antibodies against CTCF and nucleophosmin showed that these proteins also colocalized in vivo over the two known CTCF sites at either end of the chicken β-globin locus (Fig. 6A). Fluorescence in situ hybridization analysis of cell lines carrying multiple copies of the insulator showed localization of the insulator elements at the surface of the nucleolus, where nucleophosmin is also concentrated. Mutation of the CTCF site abolished localization (Yusufzai et al. 2004). In other experiments, it was shown also that CTCF can form homodimers and probably higher-order oligomers as well.

The class of models suggested by these results is quite similar to one proposed earlier by Corces and his collaborators to explain the action of the gypsy element in *Drosophila* (Gerasimova et al. 2000). These models invoke the generation of separate loop domains, mediated by the protein bound to the insulators, to generate structures in which enhancer and promoter occupy separate topologically independent loops. The β-globin insulator is tethered to the nucleolar surface by its interaction with nucleophosmin, and this would be sufficient to create such domains (Fig. 6B). However, the ability of CTCF molecules to interact with one another would also result in the formation of loops. In the case of the gypsy element, both clustering and tethering to the nuclear envelope have been observed.

There are two kinds of enhancer mechanisms arising from these models that could serve to explain insulation. Either the enhancer can no longer make a necessary physical contact with the promoter, or some processive signal that normally passes from enhancer to promoter is blocked at the point where the insulator helps to form the

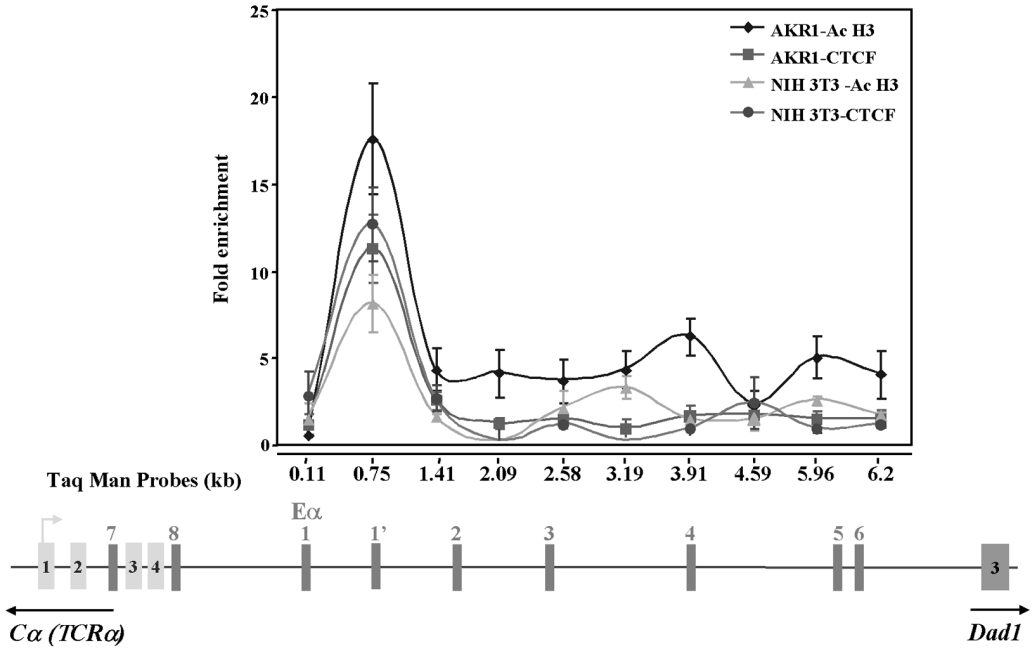

Figure 4. The distribution of histone H3 acetylation and CTCF binding in vivo, determined by chromatin immunoprecipitation, for the region between the mouse *T cell receptor* α gene and the downstream *Dad1* gene. Experiments were carried out both in the fibroblast line NIH-3T3 and in a lymphocyte line, AKR1. Hypersensitive sites 1–6 previously described (Zhong and Krangel 1999; Ortiz et al. 2001) are shown below as dark vertical bars. The peak at 0.75 kb corresponds to hypersensitive site 1′. (Reprinted, with permission, from Magdinier et al. 2004.)

base of the loop. Evidence for such processive mechanisms has been reported (Hatzis and Talianidis 2002). It should be recalled, however, that there are other proteins that interact with CTCF (Yusufzai et al. 2004), and that their possible role in enhancer blocking remains to be explored.

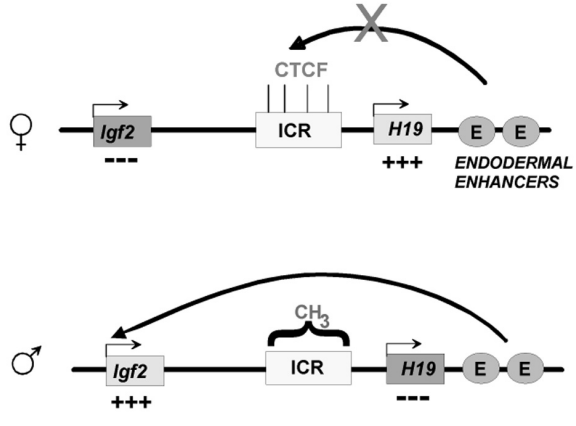

Figure 5. The imprinted *Igf2/H19* locus in mouse. The imprinted control region (ICR) of the maternally transmitted allele binds CTCF and acts as an enhancer-blocking insulator to prevent the downstream endodermal enhancers (E) from activating *Igf2* expression. In the paternal allele, the DNA of the ICR is methylated, CTCF does not bind, and the enhancers can activate *Igf2* expression (Bell and Felsenfeld 2000; see also Hark et al. 2000; Kanduri et al. 2000). (Reprinted, with permission, from Bell and Felsenfeld 2000. http://www.nature.com)

BOUNDARY FUNCTIONS AND THE SEPARATION OF INSULATOR ACTIVITIES

The second kind of insulator serves to provide a barrier against invasion of an open chromatin region by heterochromatin. We devised an assay in which we measured the expression in erythroid cell lines of an integrated transgene surrounded by the element to be tested for barrier activity. In most lines carrying unprotected constructs, expression is extinguished by 20–80 days in culture. This probably reflects integration into endogenous sites that are silent; some lines do continue to express the reporter gene, in many cases presumably by trapping a strong enhancer (Pikaart et al. 1998). We tested the ability of both the full 1.2-kb globin HS4 insulator element and the 250-bp "core" to protect against extinction of expression. When the reporter is flanked on each side by two copies of either of these elements, silencing does not occur. Thus the globin insulator also possesses the second, "barrier" activity that protects against position effect.

The obvious next step was to dissect the core element to isolate the components responsible for this activity. We undertook successive deletion of each of the five footprinted regions of the core described above. To our surprise, deletion of the CTCF binding site had no effect on the barrier activity. In contrast, deletion of any one of the other four footprint regions severely inhibited barrier function. The HS4 insulator enhancer-blocking function is therefore separable from the barrier function; HS4 is a compound element with interspersed regulatory sequences (Recillas-Targa et al. 2002).

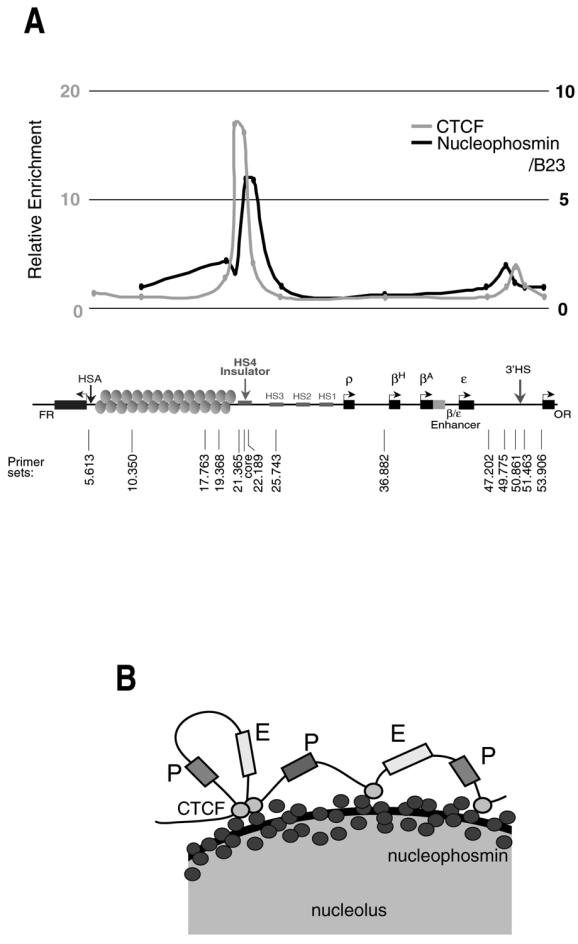

Figure 6. (*A*) Chromatin immunoprecipitation over the chicken β-globin domain with antibodies to CTCF and nucleophosmin show that nucleophosmin colocalizes with CTCF over the insulators at each end of the domain. (*B*) A possible model for enhancer-blocking action, in which CTCF is tethered to the nucleolar surface to form topologically isolated loop domains. Equivalent structures can be generated simply by pair-wise or clustered interactions between CTCF molecules (see text). (Reprinted, with permission, from Yusufzai et al. 2004.)

HISTONE MODIFICATIONS AND BARRIER FUNCTION

A clue as to how barriers might be established and maintained comes from studies of the distribution of histone modifications over the β-globin locus. Chromatin immunoprecipitation studies show that there are peaks of histone acetylation over 5′HS4 that are present in every cell type that has been studied (Fig. 7) (Litt et al. 2001b). Immediately upstream of 5′HS4 and the β-globin locus is a region of condensed chromatin (see below), marked by dimethylation of histone H3 lysine 9 (Litt et al. 2001a). This methylation mark is characteristic of heterochromatin. Schemes have been proposed in which the presence of this modification on one nucleosome can recruit the methylating enzyme necessary to modify an adjacent nucleosome, leading to propagation of the inactivation signal. We have suggested that the presence of a high level of H3 lysine 9 acetylation at HS4 could prevent the adjacent histones from being methylated at lysine 9, preventing the advance of the propagating silencing signal. When the HS4 insulator flanks a transgene, the targeted histone acetylation by the insulators may prevent the influence of surrounding endogenous condensed chromatin in the same way.

Recent evidence from our laboratories suggests that other modifications induced by HS4 may similarly block the advance of other silencing signals. A similar mechanism has been proposed for the barrier function at the end of the mating type locus in yeast (Donze and Kamakaka 2001).

In this connection it is interesting to note that silencing of the integrated but uninsulated reporter used in our barrier assays can be prevented by growing the cells in the constant presence of Trichostatin A, which inhibits histone deacetylation (Mutskov et al. 2002). In the absence of such inhibition, the kinetics of loss of histone H3 and H4 acetylation closely parallel loss of transcript; promoter DNA methylation, usually associated with silencing, occurs only afterward. The kinetics of silencing exclude DNA methylation as the primary causative event in this transgene system (Mutskov and Felsenfeld 2004).

We are presently dissecting the HS4 region further to identify the proteins responsible for the barrier behavior. It is already clear that multiple proteins are involved, and that these include histone-modifying enzymes that correspond to the histone modifications observed in the neighborhood of the insulator.

CONDENSED CHROMATIN

The condensed chromatin region that begins immediately upstream of 5′HS4 extends for about 16 kb and is followed at its 5′ end by a gene for an erythroid-specific folate receptor (Prioleau et al. 1999). It is important to understand the structure of such regions, which in vivo may well affect the expression of nearby genes. Our early studies of this region showed that it was composed of regularly arrayed nucleosomes, and that the DNA within it was highly methylated at CpG sites. We therefore took advantage of accessible and unmethylated HpaII restriction sites at the borders of the region to excise it from the nucleus. Sedimentation in a sucrose gradient revealed an essentially monodisperse fragment as revealed either by Southern blotting or polymerase chain reaction (PCR) analysis (Fig. 8). The individual gradient fractions could be studied in the analytical ultracentrifuge, yielding a precise value for the sedimentation coefficient of the heterochromatin particle. We were also able to measure the buoyant density of the particle. The combined density and sedimentation information gave its frictional coefficient, a measure of particle shape. The measured value corresponds to an extended rod-like particle, and is consistent with results obtained from electron micrograph and diffraction studies of chromatin fibers that detect a rod-like structure about 30 nm in diameter (Ghirlando et al. 2004).

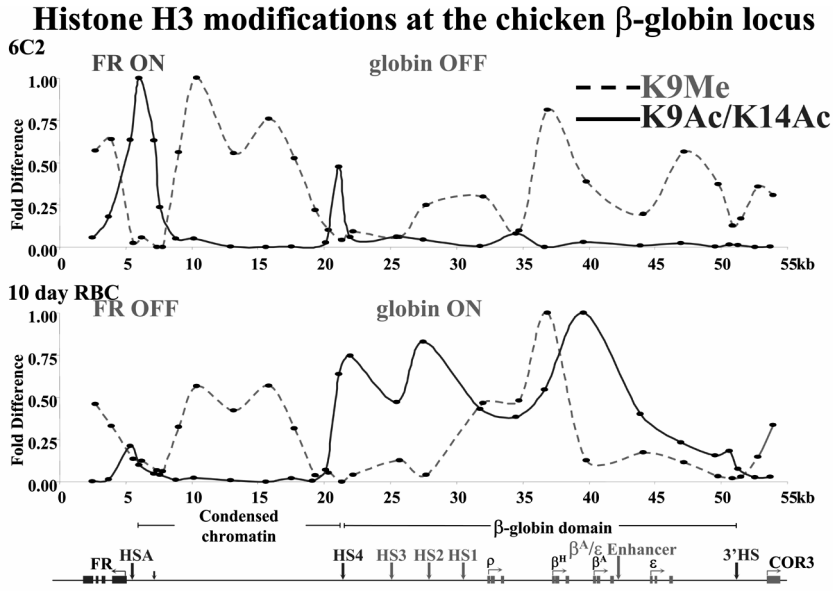

Figure 7. Distribution of histone H3 lysine 9 methylation and lysine 4 methylation over the β-globin region in the chicken erythroleukemia cell line 6C2 (*top*) and 10-day chick embryonic red blood cells (*bottom*). (*Top*, Reprinted, with permission, from Litt et al. 2001a; *bottom*, reprinted, with permission, from Litt et al. 2001b.)

CONCLUSIONS

The original purpose of this study was to investigate the structure of chromatin at the boundaries of the chicken β-globin locus and to understand how longer-range chromatin structure might mediate globin gene expression. During the ensuing studies we identified a complex insulator element at the 5′ end of the locus that separated it from an extended adjacent heterochromatic region and, beyond that, a folate receptor gene with a different program of expression from that of the globin genes. Dissection of the insulator showed that it had separate elements capable both of serving as a barrier to heterochromatinization and of blocking distal enhancers. The latter activity was attributable to binding of the protein CTCF, and we were able to show that CTCF insulator sites are present elsewhere in the genome, notably at the Igf2/H19 imprinted locus. Recent results show that CTCF can interact both with itself and with a nucleolar protein, nucleophosmin; this suggests possible mechanisms of enhancer-blocking action. Barrier activity is also mediated by the globin insulator, but recent results in our laboratory show that it involves a different set of binding sites and proteins, appropriate to what we believe is its task of preventing the immediately adjacent, compact, and inactive 16-kb heterochromatin domain from advancing.

This series of investigations also has revealed a lot about the role of chromatin structure in regulating gene expression. The analysis of the developmentally regulated distribution of histone modifications over the β-globin locus provides important correlations between these modifications and the state of expression over the globin and folate receptor genes, as well as the inactive 16-kb heterochromatic region. Recent data from our laboratory (G. Felsenfeld et al., unpubl.) extends this work to a variety of other modifications, with the ultimate goal of determining the time course of activation and inactivation events. The studies of enhancer-blocking activity take us in yet another direction, calling attention to the likely importance of higher-order structures and interaction with components of the nuclear architecture in regulating expression. The study of insulators has thus provided a powerful way of studying problems of chromatin structure and gene expression that might otherwise be difficult to address.

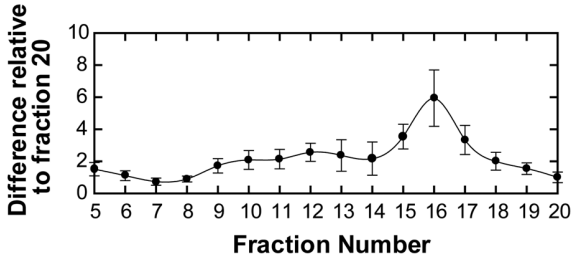

Figure 8. Sucrose gradient sedimentation of a *Hpa*II digest of 6C2 cell nuclei. The digestion releases an ~16-kb condensed chromatin fragment that lies between the 5′ end of the chicken β-globin domain and the folate receptor gene further upstream (see Fig. 2). The fragment was detected by PCR, and the chromatin in the peak fraction was subjected to analytical ultracentrifugation. (Reprinted, with permission, from Ghirlando et al. 2004.)

REFERENCES

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

Bell A.C., West A.G., and Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98:** 387.

Chao W., Huynh K.D., Spencer R.J., Davidow L.S., and Lee J.T. 2002. CTCF, a candidate trans-acting factor for X-inactivation choice. *Science* **295:** 345.

Chung J.H., Bell A.C., and Felsenfeld G. 1997. Characterization of the chicken beta-globin insulator. *Proc. Natl. Acad. Sci.* **94:** 575.

Donze D. and Kamakaka R.T. 2001. RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in *Saccharomyces cerevisiae*. *EMBO J.* **20:** 520.

Farrell C.M., West A.G., and Felsenfeld G. 2002. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol.* **22:** 3820.

Felsenfeld G. 1993. Chromatin structure and the expression of globin-encoding genes. *Gene* **135:** 119.

Gerasimova T.I., Byrd K., and Corces V.G. 2000. A chromatin insulator determines the nuclear localization of DNA. *Mol. Cell* **6:** 1025.

Ghirlando R., Litt M.D., Prioleau M.N., Recillas-Targa F., and Felsenfeld G. 2004. Physical properties of a genomic condensed chromatin fragment. *J. Mol. Biol.* **336:** 597.

Hark A.T., Schoenherr C.J., Katz D.J., Ingram R.S., Levorse J.M., and Tilghman S.M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* **405:** 486.

Hatzis P. and Talianidis I. 2002. Dynamics of enhancer-promoter communication during differentiation-induced gene activation. *Mol. Cell* **10:** 1467.

Hebbes T.R., Clayton A.L., Thorne A.W., and Crane-Robinson C. 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J.* **13:** 1823.

Kanduri C., Pant V., Loukinov D., Pugacheva E., Qi C.F., Wolffe A., Ohlsson R., and Lobanenkov V.V. 2000. Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. *Curr. Biol* **10:** 853.

Litt M.D., Simpson M., Gaszner M., Allis C.D., and Felsenfeld G. 2001a. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. *Science* **293:** 2453.

Litt M.D., Simpson M., Recillas-Targa F., Prioleau M.N., and Felsenfeld G. 2001b. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *EMBO J.* **20:** 2224.

Magdinier F., Yusufzai T.M., and Felsenfeld G. 2004. Both CTCF-dependent and -independent insulators are found between the mouse T cell receptor alpha and Dad1 genes. *J. Biol. Chem*. **279:** 25381.

Mutskov V. and Felsenfeld G. 2004. Silencing of transgene transcription precedes methylation of promoter DNA and histone H3 lysine 9. *EMBO J.* **23:** 138.

Mutskov V.J., Farrell C.M., Wade P.A., Wolffe A.P., and Felsenfeld G. 2002. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev.* **16:** 1540.

Ortiz B.D., Harrow F., Cado D., Santoso B., and Winoto A. 2001. Function and factor interactions of a locus control region element in the mouse T Cell receptor-α/Dad1 gene locus. *J. Immunol.* **167:** 3836.

Pikaart M.J., Recillas-Targa F., and Felsenfeld G. 1998. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* **12:** 2852.

Prioleau M.N., Nony P., Simpson M., and Felsenfeld G. 1999. An insulator element and condensed chromatin region separate the chicken beta-globin locus from an independently regulated erythroid-specific folate receptor gene. *EMBO J.* **18:** 4035.

Recillas-Targa F., Pikaart M.J., Burgess-Beusse B., Bell A.C., Litt M.D., West A.G., Gaszner M., and Felsenfeld G. 2002. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc. Natl. Acad. Sci.* **99:** 6883.

Reitman M. and Felsenfeld G. 1990. Developmental regulation of topoisomerase II sites and DNase I-hypersensitive sites in the chicken beta-globin locus. *Mol. Cell. Biol.* **10:** 2774.

Saitoh N., Bell A.C., Recillas-Targa F., West A.G., Simpson M., Pikaart M., and Felsenfeld G. 2000. Structural and functional conservation at the boundaries of the chicken beta-globin domain. *EMBO J.* **19:** 2315.

Tanimoto K., Sugiura A., Omori A., Felsenfeld G., Engel J.D., and Fukamizu A. 2003. Human beta-globin locus control region HS5 contains CTCF- and developmental stage-dependent enhancer-blocking activity in erythroid cells. *Mol. Cell. Biol.* **23:** 8946.

West A.G., Gaszner M., and Felsenfeld G. 2002. Insulators: Many functions, many mechanisms. *Genes Dev.* **16:** 271.

Yusufzai T.M., Tagami H., Nakatani Y., and Felsenfeld G. 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol. Cell* **13:** 291.

Zhong X. and Krangel M.S. 1999. Enhancer-blocking activity within the DNase I hypersensitive site 2 to 6 region between the TCR α and Dad1 genes. *J. Immunol.* **163:** 295.