December 2002

The insulation of genes from external enhancers and silencing chromatin

Bonnie Burgess-Beusse
Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD

Catherine Farrell
Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD

Miklos Gaszner
Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD

Michael Litt
Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD

Vesco Mutskov
Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD

See next page for additional authors

The insulation of genes from external enhancers and silencing chromatin

Bonnie Burgess-Beusse, Catherine Farrell, Miklos Gaszner, Michael Litt, Vesco Mutskov, Felix Recillas-Targa, Melanie Simpson, Adam West, and Gary Felsenfeld*

Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892-0540

Insulators are DNA sequence elements that can serve in some cases as barriers to protect a gene against the encroachment of adjacent inactive condensed chromatin. Some insulators also can act as blocking elements to protect against the activating influence of distal enhancers associated with other genes. Although most of the insulators identified so far derive from *Drosophila*, they also are found in vertebrates. An insulator at the 5′ end of the chicken β-globin locus marks a boundary between an open chromatin domain and a region of constitutively condensed chromatin. Detailed analysis of this element shows that it possesses both enhancer blocking activity and the ability to screen reporter genes against position effects. Enhancer blocking is associated with binding of the protein CTCF; sites that bind CTCF are found at other critical points in the genome. Protection against position effects involves other properties that appear to be associated with control of histone acetylation and methylation. Insulators thus are complex elements that can help to preserve the independent function of genes embedded in a genome in which they are surrounded by regulatory signals they must ignore.

Although DNA methylation has been considered the primary chromosomal modification involved in transmission of epigenetic information, it is becoming clear that the chromatin proteins, especially the histones, are also involved in this process. It is understood that the histones bound to transcriptionally active and inactive regions of the genome are chemically modified in different ways, and mechanisms have been proposed by which these modified states can be propagated along the chromosome and perhaps transmitted during replication.

We have for some time been interested in the boundary elements called insulators. Detailed study of these elements has led us to mechanisms that are coupled to allele-specific expression at an imprinted locus, as well as to mechanisms by which propagation of inactive chromatin states may be modulated.

An Insulator as an Enhancer-Blocking Element

Within the vertebrate genome, transcriptionally active genes are embedded in an environment containing in some cases extensive regions of condensed chromatin. In other cases active genes may be located near other, silent, genes that have a different program of expression. The possibilities thus arise that the active gene will be inappropriately silenced by the condensed chromatin or will inappropriately activate the adjacent silent gene. It is equally possible that in other tissues or at other developmental stages, where this gene is inactive, signals from adjacent extraneous enhancers could cause incorrect patterns of expression.

During the past several years studies begun initially in *Drosophila* and now extended to vertebrates have identified DNA sequence elements called insulators that appear to function as blocks against both kinds of signals from the outside. Two kinds of assay have been developed to measure these properties (Fig. 1). The first assay measures "enhancer blocking," the ability to shield a promoter from the action of a distal enhancer without preventing the enhancer from working on a proximal promoter. The second assay measures the "barrier" activity that prevents the advance of adjacent condensed chromatin. We have applied both of these assays in the study of vertebrate insulators.

Our attention was first drawn to this problem through our interest in the role of chromatin structure in the regulation of gene expression in the chicken β-globin locus (Fig. 2). The gene cluster contains four members of the β-globin family, expressed at different developmental stages. The regulatory elements of these genes are marked by a series of erythroid-specific DNase hypersensitive sites (HSs), but at the 5′ end of the locus there is a "constitutive" HS (5′HS4) present in all tissues that have been examined (1–3). We speculated that this HS might mark the 5′ boundary of the "open" globin chromatin domain, and indeed subsequent work (4) has shown (Fig. 2) that there is an abrupt change from a chromatin structure characterized by general heightened nuclease sensitivity and a high level of histone acetylation (signs of an active globin chromatin locus), to a region of condensed chromatin further upstream, that is nuclease resistant and underacetylated.

To test whether the DNA at 5′HS4 had properties of an insulator we devised a method to assay enhancer blocking activity (Fig. 3*A*), based on the ability of a 1.2-kb element that includes the HS to shield a reporter expressing a neomycin resistance gene from the action of a strong enhancer. As shown in Fig. 3*B*, the 1.2-kb element is quite effective in reducing the number of G418-resistant colonies, a measure of the strength of the blockade. Placing the 1.2-kb element outside the region between enhancer and promoter resulted in little blocking (2, 3), confirming that 5′HS4 has insulating properties. We dissected this region further and found that a 250-bp "core" sequence containing the HS was equally effective in enhancer blocking. Further subdivision, making use of the DNase I footprint patterns generated on the core by nuclear extracts, revealed that a single binding site (footprint II) was also active. This finding led to the identification of a known regulatory protein, CTCF, as the DNA binding factor responsible for enhancer blocking activity (5).

We asked where else CTCF sites with enhancer blocking activity could be found. It seemed plausible that if these were truly associated with boundaries there might also be one at the 3′ end of the β-globin locus. Indeed a constitutive HS with this activity, and binding CTCF, is found just upstream of the 3′ condensed chromatin region (ref. 6, Fig. 4). This region also harbors a gene for an odorant receptor (7). It seems possible that

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Self-Perpetuating Structural States in Biology, Disease, and Genetics," held March 22–24, 2002, at the National Academy of Sciences in Washington, DC.

Abbreviations: HS, hypersensitive site; FR, folate receptor; ICR, imprinted control region; IL-2R, IL-2 receptor.

*To whom reprint requests should be addressed. E-mail: gary.felsenfeld@nih.gov.

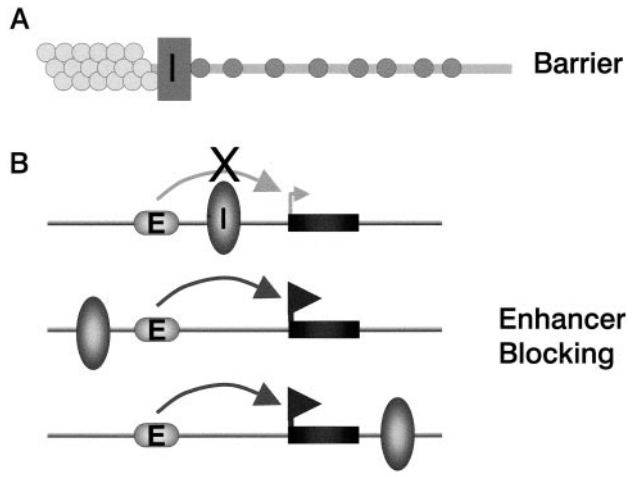

Fig. 1. Two kinds of insulator functions. (*A*) Some insulators may function as barriers against the encroachment of adjacent genomic condensed chromatin. (*B*) Some insulators may serve as positional enhancer-blocking elements that prevent enhancer action when placed between enhancer and promoter, but not otherwise.

the 3′ CTCF site could block cross-interaction between the regulatory elements of this receptor and those of the β-globin locus. CTCF sites are also present at conserved locations in the mouse and human loci that are embedded within clusters of odorant receptor genes (8). Further examination of the DNA upstream of the chicken 5′HS4 insulator reveals a somewhat different arrangement: Here, the upstream sequence is packaged as condensed chromatin containing CR1 repeat sequences. This sequence extends for about 16 kb and is followed by a gene (ref. 9, Fig. 4) for an erythroid-specific folate receptor (FR) that is expressed at a developmental stage preceding the point at which the globin genes are switched on. The FR gene and the globin genes are not expressed at the same stages, so that the presence of an enhancer blocking activity at 5′HS4 might once again be useful in avoiding cross-talk between the two gene systems.

A role for CTCF-mediated enhancer blocking activity has been demonstrated most clearly at the *Igf2/H19* locus in mouse and human (10–12). In this imprinted locus the maternally transmitted allele expresses H19 but not Igf2, whereas the paternally transmitted allele expresses Igf2 but not H19. Furthermore the paternal allele is methylated at a site (the ICR or imprinted control region) located between the two genes (Fig. 5). Earlier work had suggested that the ICR might contain an enhancer blocking activity that would prevent downstream endodermal enhancers from activating *Igf2.* Direct examination reveals that the ICR contains four CTCF binding sites in the mouse ICR and seven in human. Methylation of these sites abolishes CTCF binding. These results indicate that CTCF plays an important role as an insulator protein in allele-specific regulation at this imprinted locus, and that the insulator function can be modulated by DNA methylation, thus making the CTCF sites susceptible to epigenetic regulation. Quite recently a cluster of differentially methylated CTCF sites has been identified at the *Xist* gene promoter, and it has been suggested that these are enhancer-blocking elements important for X chromosome inactivation (13).

Insulators as Barriers

A second property that some insulators possess is the ability to protect against silencing caused by formation of condensed chromatin. This is the barrier function, which can be detected by assays designed to measure protection of stably integrated transgenes against position effects (Fig. 1). We established an assay to test barrier function by constructing a reporter expressing a fragment of the IL-2 receptor (IL-2R) (Fig. 6) driven by an erythroid-specific promoter and enhancer, and integrating it into a chicken erythroid cell line, 6C2 (14). Typically after 80–100 days in culture expression was extinguished in most lines. This extinction is a manifestation of position effects, i.e., the dependence of expression on the site of integration. We repeated the experiment with the same reporter, but flanked on each side by two copies of the 1.2-kb 5′HS4 element; now expression was maintained in nearly all lines even after 80–100 days of incubation. The 5′HS4 element thus protects against position effects, a second property possessed by some insulators.

Silencing of gene expression can involve a number of chromatin modifications, and insulators might interfere with some of these. We turned again to the chicken β-globin locus and examined the state of modification of the histones over the entire

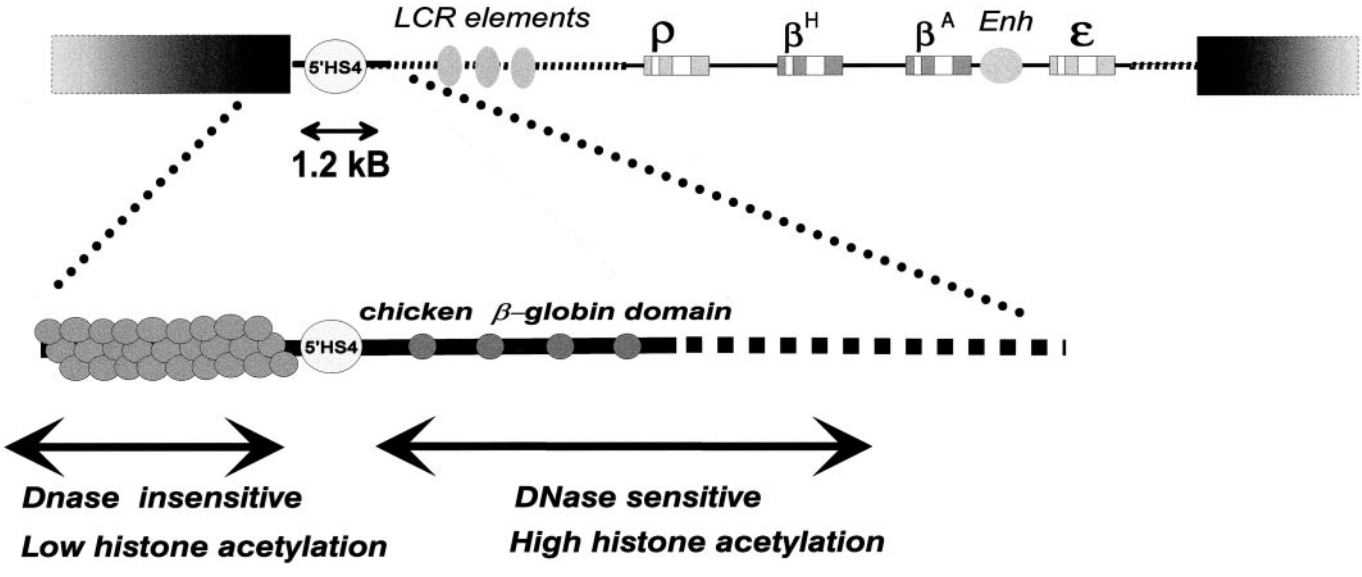

Fig. 2. The chicken β-globin locus (*Upper*) showing the four genes, the strong enhancer between the adult β gene (β^A) and the ε gene, and the constitutive HS, 5′HS4. The boundary between the open chromatin domain and the condensed chromatin domain further 5′, as determined by Hebbes *et al.* (4), is shown (*Lower*).

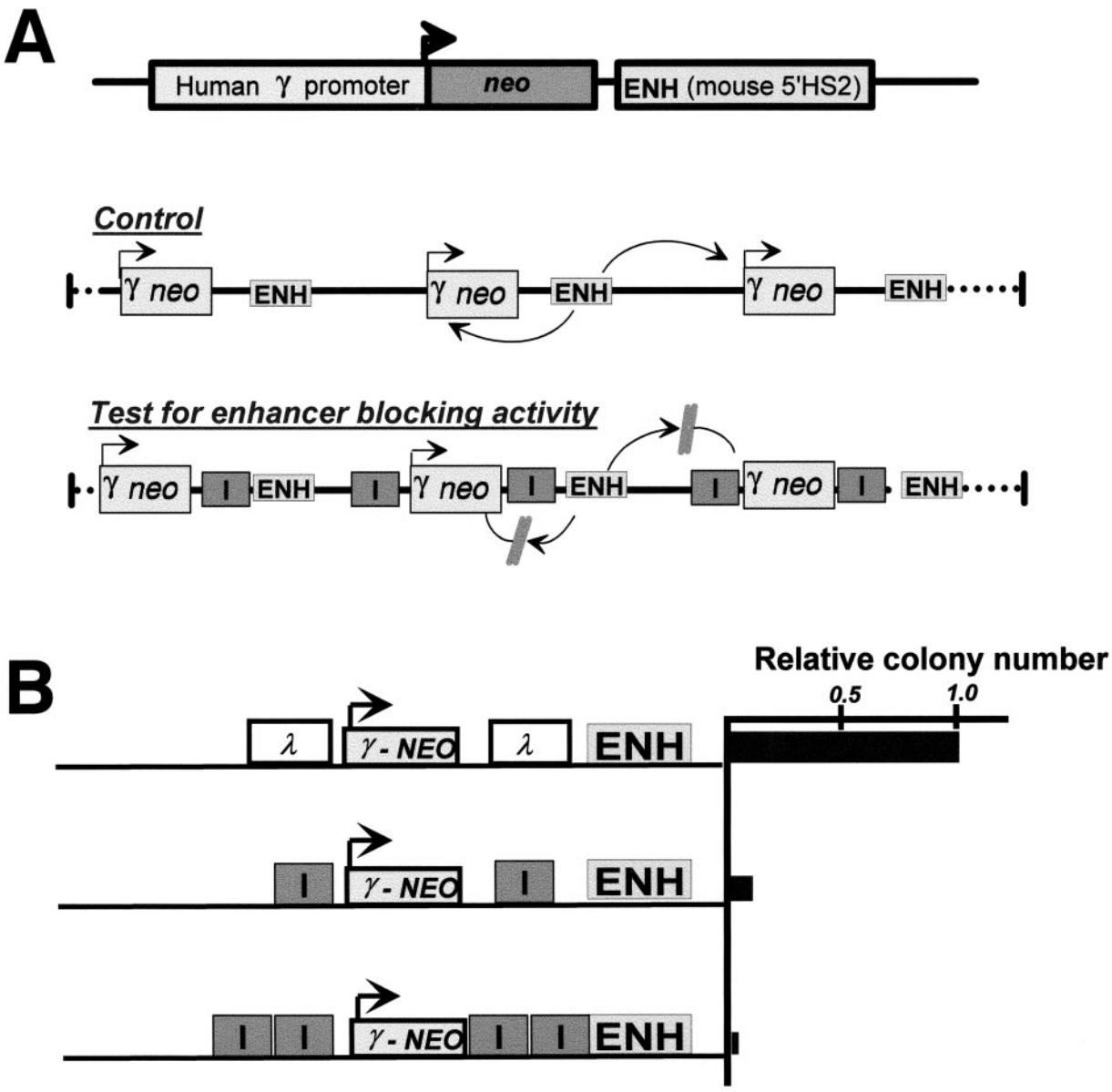

Fig. 3. (*A*) Construction used in an assay for enhancer (ENH) blocking activity (2, 3). Expression of a gene conferring G418 resistance (*neo*) is driven by an erythroid-specific enhancer and promoter. This plasmid is stably transfected into K562 human erythroleukemia cells, and G418-resistant colonies are counted. Typically transfection produces tandem integrants (Control). The test for the putative insulator (I) is to insert it so that it can block enhancer action, and again to count colonies. (*B*) Results of inserting a 1.2-kb fragment (see Fig. 2) containing 5′HS4 on colony number in the above assay. The control has an approximately equal length of λ phage DNA on either side of the reporter to keep distances constant. The 5′HS4 element strongly reduces enhancer (ENH) activity. Other experiments show that it has a much smaller effect when placed on the other side of the enhancer, confirming the positional enhancer blocking activity (2, 3).

54-kb region containing the globin genes and the upstream FR (15, 16). Chromatin immunoprecipitation (ChIP) experiments with antibodies to acetylated histones H3 and H4 revealed elevated levels of acetylation associated with activation of the individual genes. Thus there was strong acetylation over the FR gene in 6C2 cells (Fig. 7), which express this gene, but in 10-day embryonic erythrocytes, corresponding to a later developmental stage, high levels of acetylation are shifted to the globin gene cluster (15). At all stages, the approximately 16-kb condensed chromatin region upstream of the globin genes remains unacetylated. Furthermore, there is a peak of acetylation over the 5′HS4 insulator element (as well as over the HSA enhancer of the FR gene) in all cells we examined, including a DT40 lymphocyte line in which neither globin nor FR genes are active (15). Recent evidence (16–20) implicates histone methylation as well in the regulation of expression: Methyl group modifications at histone H3 lysines 4 or 9 are associated with active or inactive chromatin, respectively. We again used ChIP methods to measure patterns of histone methylation over the same region (Fig. 7). We found a striking correlation between previously observed patterns of histone acetylation and lysine 4 methylation and anticorrelation between acetylation and methylation of lysine 9. The clearest example of such distinct regions of modification is found in the mating type locus of *Schizosaccharomyces pombe*, where the heterochromatic and euchromatic domains of the mating type locus are distinguished by similar patterns of methylation and acetylation (18).

Work in a number of laboratories has led to models for propagation of the condensed chromatin state (15, 17, 18). These

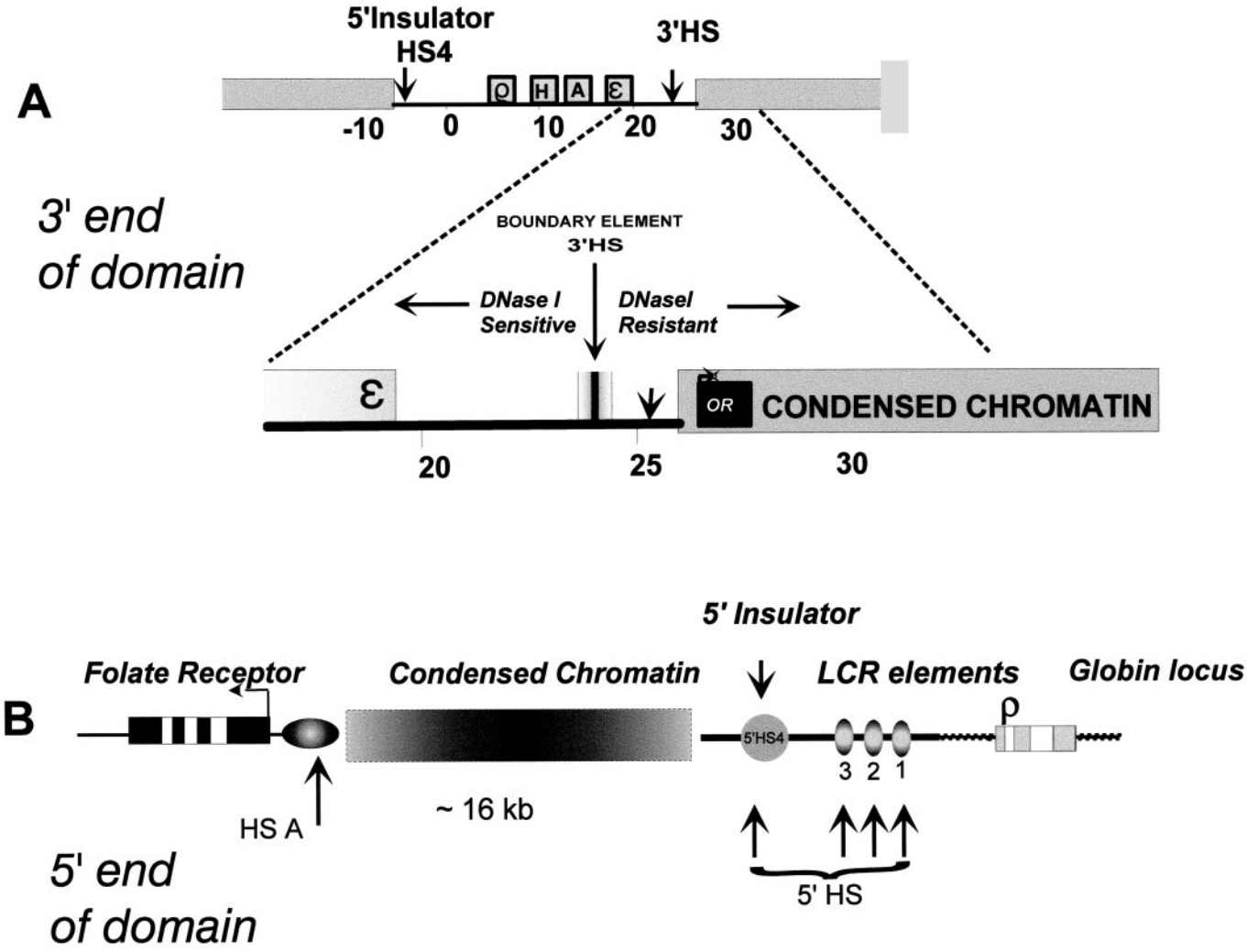

Fig. 4. The ends of the open β-globin domain. (*A*) 3′end. A second constitutive HS that binds CTCF and has enhancer blocking activity is found upstream of the beginning of a condensed chromatin region containing a gene for an odorant receptor (8). (*B*) 5′ end. A condensed chromatin region extends for about 16 kb upstream of 5′HS4, and beyond that is an erythroid-specific FR gene (9).

are based on the following observations: (*i*) the heterochromatin protein HP1 [or in *S. pombe* its homolog Swi6 (18)] binds selectively to histone H3 that is methylated at lysine 9 in the amino terminal tail, and (*ii*) the enzyme responsible for methylating lysine 9, Suv39H1, interacts with HP1. This finding leads to the proposal (15, 17, 18) that a nucleosome methylated at that site will indirectly recruit Suv39H1 and promote methylation of sites on the adjacent nucleosome. Our observation that the 5′HS4 insulator is a major center of histone acetylation suggests that one of its roles at the 5′ end of the β-globin locus is to continually acetylate the adjacent upstream nucleosome, in particular lysine 9 of H3. By doing so, it prevents methylation of that residue and thus terminates the propagation of the condensation signal. A related mechanism has been proposed for barrier activity in *Saccharomyces cerevisiae*, where a site that binds histone acetylases is sufficient to prevent extension of silencing from HMR-E (21).

These results suggest that *in vivo* the 5′HS4 insulator has two functions: it serves as an enhancer blocking element to screen out upstream signals, and it also acts as a barrier against the advance of the condensed chromatin structure immediately upstream. In fact, recent results indicate that these two functions are separable (22). The enhancer blocking function, as discussed

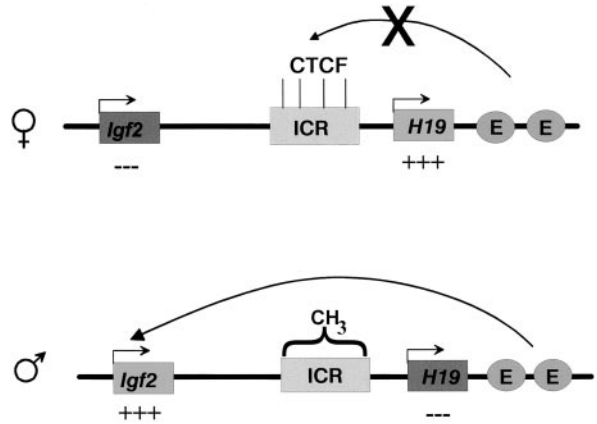

Fig. 5. The mouse *Igf2*/*H19* imprinted locus. In the maternally transmitted allele, *Igf2* is silent, but in the paternal allele it is expressed, and the ICR is methylated. The ICR has been shown to contain four CTCF binding sites, which have strong enhancer blocking properties; enhancer blocking is abolished by methylation of cytosines at CpG sites within the ICR. These and other results lead to a model in which the maternal ICR blocks the action of downstream endodermal enhancers (E) on the *Igf2* promoter. Methylation of the paternal ICR abolishes enhancer blocking and permits *Igf2* activation (10–12).

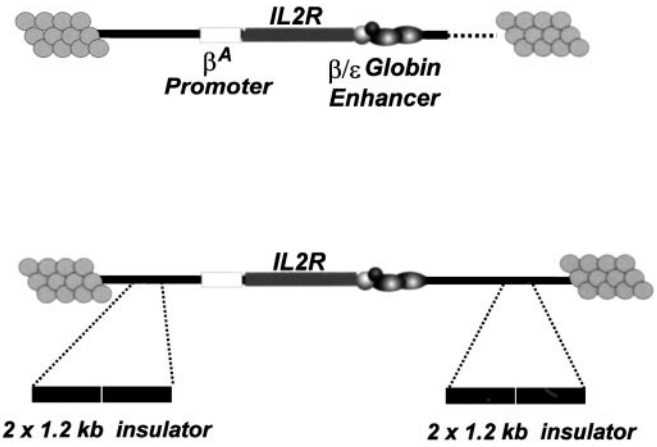

Fig. 6. Constructions to measure protection against position effects (barrier function). A reporter containing a fragment of the IL-2R driven by the chicken adult β-globin promoter and the downstream β/ε enhancer (see Fig. 2) is stably transformed into the avian erythroid line 6C2. Expression of IL-2R on the cell surface is monitored by FACS analysis. (*Upper*) The control construction. (*Lower*) The control reporter is surrounded by two copies of the 1.2-kb chicken 5′HS4 insulator on each side. In most lines transformed with the control construction, IL-2R expression was extinguished after 80–100 days in culture. Almost all lines carrying the insulated construction still expressed IL-2R after 80–100 days (14).

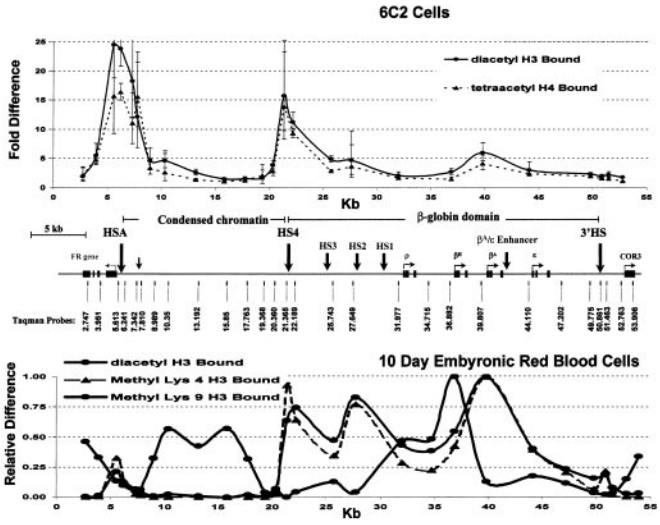

Fig. 7. Chromatin immunoprecipitation of modified histones across the β-globin locus. (*Top*) Diacetylated histone H3 and tetraacetylated histone H4 in 6C2 cells, arrested at the CFU-E stage of chicken erythroid development. (*Middle*) Map of the locus showing positions of HSs above the line and of probes used in PCR detection below the line. (*Bottom*) Patterns of histone H3 methylation at lysines 4 and 9 and diacetylated (lysines 9 and 14) histone H3 across the locus in 10-day embryonic chicken erythrocytes. (Adapted from refs. 15 and 16.)

above, depends on the CTCF site, whereas the barrier function depends on the other four subregions of the 250-bp insulator core and does not require CTCF. We note that the CTCF site at the 3′ end of the globin locus is not accompanied by the other subregions present in 5′HS4 and is also not a site of strong acetylation. The condensed chromatin at the 3′ end of the locus is facultative, i.e., it must open for expression of the odorant receptor, and it may therefore be distinct in properties from the constitutive condensed chromatin at the 5′ end and may not require an acetylated barrier.

How widely are insulators distributed in the genome? A considerable number of elements have been identified in *Drosophila*, each with its own characteristic site and associated proteins (for a recent general review of insulators see ref. 23). Barrier functions have been identified in both *S. pombe* and *S. cerevisiae* (20, 21). In vertebrates, the CTCF site appears to be widely distributed and to function in many cases as an enhancer blocking agent. However, no other barrier (position effect) elements have yet been identified similar to that found at the β-globin locus. Evidently many genes will not require insulators as a protection against inappropriate interaction of neighboring signals, but it seems reasonable to look for them in cases (such as the Ig and T cell receptor gene loci) where multiple regulatory elements are clustered. As the results in *Drosophila* and yeast have shown, diverse proteins and sites may be involved. There is no reason to think that vertebrates will be less complex, so it is likely that many novel insulators remain to be identified.

Conclusion

Insulator elements are clearly an important family of regulatory sequences, likely to be distributed widely in the genome. The connection of enhancer blocking activity with imprinted loci, and tentatively with the X-inactivation locus (13), suggests a role in the establishment of epigenetic imprinting marks. The barrier function is connected with the maintenance of boundaries that may also be established through epigenetic signals such as histone methylation. We are still not completely certain of detailed mechanisms of insulator action, but as we learn more, we will also understand more about how the cell controls and exploits epigenetic signals.

1. Reitman, M. & Felsenfeld, G. (1990) *Mol. Cell. Biol.* **10,** 2774–2786.
2. Chung, J. H., Whiteley, M. & Felsenfeld, G. (1993) *Cell* **74,** 505–514.
3. Chung, J. H., Bell, A. C. & Felsenfeld, G. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 575–580.
4. Hebbes, T. R., Clayton, A. L., Thorne, A. W. & Crane-Robinson, C. (1994) *EMBO J.* **13,** 1823–1830.
5. Bell, A. C., West, A. G. & Felsenfeld, G. (1999) *Cell* **98,** 387–396.
6. Saitoh, N., Bell, A. C., Recillas-Targa, F., West, A. G., Simpson, M., Pikaart, M. & Felsenfeld, G. (2000) *EMBO J.* **19,** 2315–2322.
7. Bulger, M., van Doorninck, J. H., Saitoh, N., Telling, A., Farrell, C., Bender, M. A., Felsenfeld, G., Axel, R., Groudine, M. & von Doorninck, J. H. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 5129–5134.
8. Farrell, C. M., West, A. G. & Felsenfeld, G. (2002) *Mol. Cell. Biol.* **22,** 3820–3831.
9. Prioleau, M. N., Nony, P., Simpson, M. & Felsenfeld, G. (1999) *EMBO J.* **18,** 4035–4048.
10. Bell, A. C. & Felsenfeld, G. (2000) *Nature (London)* **405,** 482–485.
11. Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M. & Tilghman, S. M. (2000) *Nature (London)* **405,** 486–489.
12. Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C. F., Wolffe, A., Ohlsson, R. & Lobanenkov, V. V. (2000) *Curr. Biol.* **10,** 853–856.
13. Chao, W., Huynh, K. D., Spencer, R. J., Davidow, L. S. & Lee, J. T. (2002) *Science* **295,** 345–347.
14. Pikaart, M. J., Recillas-Targa, F. & Felsenfeld, G. (1998) *Genes Dev.* **12,** 2852–2862.
15. Litt, M. D., Simpson, M., Recillas-Targa, F., Prioleau, M. N. & Felsenfeld, G. (2001) *EMBO J.* **20,** 2224–2235.
16. Litt, M. D., Simpson, M., Gaszner, M., Allis, C. D. & Felsenfeld, G. (2001) *Science* **293,** 2453–2455.
17. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. (2001) *Nature (London)* **410,** 116–120.
18. Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B. D., Sun, Z. W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C. P., Allis, C. D., *et al*. (2000) *Nature (London)* **406,** 593–599.
19. Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C. & Kouzarides, T. (2001) *Nature (London)* **410,** 120–124.
20. Noma, K., Allis, C. D. & Grewal, S. I. (2001) *Science* **293,** 1150–1155.
21. Donze, D. & Kamakaka, R. T. (2001) *EMBO J.* **20,** 520–531.
22. Recillas-Targa, F., Pikaart, M. J., Burgess-Beusse, B., Bell. A. C., Litt, M., Gaszner, M. & Felsenfeld, G. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 6883–6888.
23. West, A. G., Gaszner, M. & Felsenfeld, G. (2002) *Genes Dev.* **16,** 271–288.