

# Looping versus linking: toward a model for long-distance gene activation

Michael Bulger<sup>1</sup> and Mark Groudine<sup>1-3</sup>

<sup>1</sup>Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109 USA; <sup>2</sup>Department of Radiation Oncology, University of Washington Medical School, Seattle, Washington 98195 USA

Locus control regions (LCRs) are defined by their ability, in transgenic assays, to direct high-level, tissue-specific expression of linked genes at all sites of integration examined and at moderately constant levels per gene copy. The most extensively examined LCR is that associated with the  $\beta$ -globin locus in mammals, where the  $\beta$ -globin genes reside in a linear array and are usually arranged in order of their developmental expression (Fig. 1). The  $\beta$ -globin LCR consists of several DNase I hypersensitive sites (HSs) spread over a region of 20–30 kb; the DNA sequence associated with each of the HSs contains numerous binding sites for erythroid-specific and ubiquitous transcription factors (for review, see Grosveld et al. 1993; Orkin 1995; Martin et al. 1996; Hardison et al. 1997). Expression of stably integrated  $\beta$ -globin transgenes in the absence of the LCR occurs only at low levels and varies depending upon the site of integration, and so is said to be subject to position effects. The high level of expression driven by the LCR at ectopic sites appears to be the product of at least two separable activities, namely the establishment of an 'open' chromatin domain and direct gene activation.

Gene activation is also a property of enhancers, which are defined by their ability to direct high-level expression of linked genes in transient transfection assays. Although enhancers typically also improve the expression of transgenes integrated into the genome, such expression varies among integration sites due to negative position effects. Thus, unlike LCRs, enhancers are only capable of gene activation at a subset of genomic loci. Presumably, LCRs subsume the function of enhancers along with a more dominant chromatin 'opening' activity that can override negative effects from neighboring regions. In the case of the  $\beta$ -globin LCR, 5' HS2 is known to act as an enhancer in transient transfection assays, but elements in addition to HS2 are required for LCR activity in single-copy transgenes (Ellis et al. 1993, 1996). Evidence from one study indicates that an otherwise intact human  $\beta$ -globin locus translocated near centromeric heterochromatin may be subject to a position effect (Rees et al. 1994). Thus, the distinction between enhancers and

LCRs may only be one of degree, but in most transgenic assays is still measurable.

Given the apparent importance of LCRs in the regulation of many tissue-specific genes, the mechanism by which these elements function has been the subject of a great deal of debate. The currently predominant model for LCR function involves the establishment of an open chromatin domain encompassing the regulated genes, with subsequent gene activation accomplished by direct interactions between elements of the LCR and gene promoters by DNA looping. In this review, we will summarize models for the establishment of active chromatin domains by LCRs and describe in detail the evidence for the looping model of long-distance gene activation. Recent data, especially relating to the function of boundary elements and evidence for facilitators of enhancer and LCR function, indicate that simple LCR–promoter interactions are unlikely. Based on these data, we propose an alternative to the looping model, which we term the linking model.

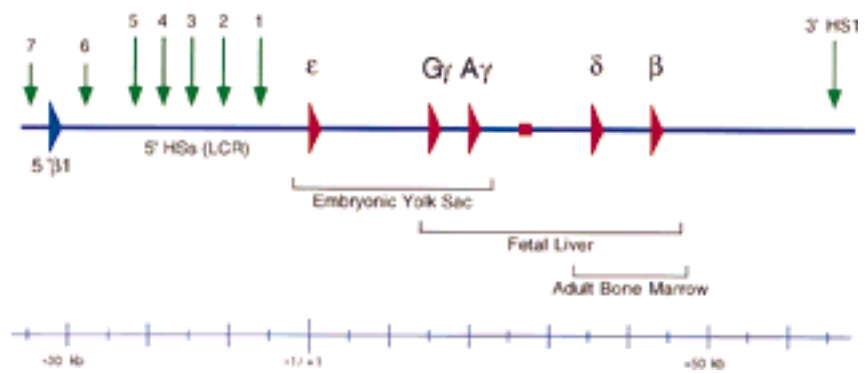
## LCRs and chromatin opening

Transcribed regions of metazoan genomes are more sensitive to nucleases as compared to nontranscribed regions (Weintraub and Groudine 1976; Stalder et al. 1980; Forrester et al. 1986; Hebbes et al. 1994). This sensitivity is not limited to the genes alone, but extends to flanking DNA, and so is termed generalized sensitivity to distinguish it from hypersensitivity displayed by localized regions of chromatin remodeling, such as enhancers and promoters. A region of the genome can be DNase I-sensitive in the absence of transcription (Reik et al. 1998), which suggests that the transition from a nuclease-insensitive to nuclease-sensitive structure is separable from the commitment to full gene activation.

The ability of LCRs to activate transcription at all sites of integration is commonly ascribed to a dominant chromatin opening activity that functions in a position-independent fashion. Presumably, the variable expression observed upon integration of simple enhancers indicates that they can provide this activity at only a subset of genomic loci. The majority of transgenic studies have not examined LCR-containing and LCR-lacking constructs at the same integration sites, so this deduc-

<sup>3</sup>Corresponding author.  
E-MAIL markg@fhcrc.org; FAX (206) 667-5994.

**Figure 1.** The human  $\beta$ -globin locus. Shown to scale are the human  $\beta$ -globin genes (red triangles) and prominent nucleosome HSs (green arrows). HS1–HS5 are commonly thought to comprise the  $\beta$ -globin LCR, although other HS both 5' of the LCR and 3' of the genes are known to exist. Brackets beneath the locus describe the expression pattern of the human  $\beta$ -globin genes in transgenic mice. The primary site of erythroid differentiation switches from the embryonic yolk sac, to the fetal liver, to adult bone marrow during mouse development. Concomitant with this, transcription from within a transgenic human locus switches from the  $\epsilon$ - and  $\gamma$ -globin genes, to the  $\gamma$ - and  $\beta$ -globin genes, to the adult globins alone, respectively.  $\delta$  is a minor globin gene and is transcribed at extremely low levels in the adult. 5'  $\beta$ 1 (blue triangle) is an olfactory receptor gene; the  $\beta$ -globin locus is known to be embedded within a cluster of such genes [Bulger et al. 1999].



tion usually is indirect. Based on transgenic studies, for example, it has long been thought that the  $\beta$ -globin LCR is able to establish an open chromatin domain at ectopic sites, as shown by the high levels of expression obtained [Grosveld et al. 1987]. Deletions performed at the endogenous  $\beta$ -globin loci of mouse and human, however, have demonstrated that in neither case is the LCR necessary for the initiation or maintenance of the nuclease-sensitive state [Epner et al. 1998; Reik et al. 1998]. This discrepancy implies either that chromatin opening at the endogenous  $\beta$ -globin locus is the result of multiple redundant activities, among which the LCR plays only a contributory role, or that the activities exhibited by the LCR at ectopic sites are in this sense fundamentally different from those it possesses at the location of the endogenous locus.

In support of the former case, enhancer-containing and -deleted transgenes examined at identical integration sites demonstrate that stability of expression in the presence of an enhancer is associated with the nuclease-sensitive state [C. Francastel, M. Groudine, and D.I.K. Martin, in prep.], providing a direct demonstration that an enhancer, in this case HS2 of the  $\beta$ -globin LCR, can be sufficient to open chromatin. It has also been suggested that the region of DNA that comprises the  $\beta$ -globin LCR may be broader than originally thought, and thus that the discrepancy between chromatin opening at ectopic and endogenous sites may be resolved by a larger deletion at the endogenous locus [Bulger et al. 1999]. Finally, it should be noted that nuclease sensitivity may provide only a crude measure of chromatin structure, and that both open and closed chromatin as defined by this measure could result from many different structures.

The phenomenon of LCR-mediated chromatin opening, as measured by relative sensitivity to nuclease digestion, raises two questions. First, what is the nature of the structural transition that occurs? Second, how does the LCR act to accomplish this transition?

### The structure of the active domain

Regions of the genome that contain a low density of active genes appear to be packaged with a distinct set of

chromatin-associated proteins that achieve a higher level of condensation. Presumably, this condensation prevents nonspecific gene activation by severely limiting the accessibility of the packaged DNA—as demonstrated by the decreased sensitivity of such regions to nuclease—and the process of gene activation must then be accomplished, at least in part, by a relative decondensation of large regions of chromatin. This decondensation or unfolding of chromatin is often thought to form a discrete domain within which the process of gene activation can take place, although examples of such clearly defined chromatin domains are actually quite rare and their extent may not be precisely determined in all cases.

Perhaps the best-characterized example of an active chromatin domain is the region containing the  $\beta$ -globin genes in chickens, where the transitions between nuclease-sensitive and -insensitive chromatin in erythroid cells have been defined on either side of the locus [Hebbes et al. 1994]. This domain is associated with an increase in levels of histone acetylation and with depletion of the linker histone H1 [Kamakaka et al. 1990; Hebbes et al. 1994]. Notably, the pattern of hyperacetylation ends where the transitions from nuclease sensitivity to insensitivity occur. These modifications are thought to lead to a more open chromatin structure by contributing to the unfolding of chromatin from a more condensed state.

The unfolded chromatin structure is most commonly portrayed as the 10-nm chromatin fiber, or 'beads-on-a-string,' but this view is unlikely to be correct. The appearance of the 10-nm fiber is known from electron micrographs of chromatin spread at low salt concentrations that do not reflect conditions in the nucleus [Thoma et al. 1979]. At physiological salt concentrations, even chromatin free of linker histones, or comprised of moderately acetylated core histones, spontaneously condenses to form a 30-nm chromatin fiber [Fletcher and Hansen 1996; Tse et al. 1998]. In addition, electron microscopic studies of the Balbiani ring genes of the moth *Chironomus tentans* indicate that the only observable decondensation below the level of the 30-nm fiber occurs between elongating RNA polymerases within the transcribed regions, and then only at high polymerase den-

sities (Andersson et al. 1982; Bjorkroth et al. 1988; Ericsson et al. 1989). When elongating polymerases were far enough apart, the chromatin structure of the transcribed region quickly reverted to the 30-nm fiber. These observations strongly suggest that the basal structure of an active chromatin domain is the 30-nm chromatin fiber. Finally, it is highly improbable that the DNA content of a cell could be stored in the nucleus if all of the active chromatin domains were to decondense to a 10-nm fiber, when this structure accomplishes only a sevenfold compaction of the DNA.

Thus, the structure of an open or active domain consists of chromatin condensed to a 30-nm fiber or structures of even higher order. For example, in situ studies utilizing the reversible transcriptional inhibitor DRB suggest that condensation of DNA in open chromatin may proceed beyond the 30-nm fiber, perhaps by supercoiling of the chromatin fiber (Andersson et al. 1982). Although nuclease sensitivity was not addressed in this study, removal of DRB resulted in rapid restoration of transcription despite the high degree of condensation observed beforehand. The consequences of this are twofold: First, all mechanisms of gene activation within a specific domain, such as the  $\beta$ -globin locus, probably occur in the context of the 30-nm fiber; and second, the nuclease-insensitive or closed chromatin structure reflects a still higher-order level of DNA packaging.

### The mechanism of chromatin opening by LCRs

Based on the existing data, any model for the mechanism by which LCRs mediate the formation of an open chromatin structure must account for how biochemical modifications such as histone acetylation and linker histone depletion are accomplished over regions as broad as the  $\beta$ -globin locus. Studies of the intronic immunoglobulin enhancer ( $E_{\mu}$ ) suggest that chromatin opening reflects the establishment of local accessibility at the enhancer that is then propagated over longer distances (Jenuwein et al. 1997). The core  $E_{\mu}$  enhancer, which in transgenic assays is subject to position effects, confers accessibility to a bacterial promoter located immediately adjacent to it. When coupled with its flanking sequences, however, the full  $E_{\mu}$  regulatory element is relatively free from position effects and confers accessibility to a bacterial promoter 1 kb distal. In addition, it has been shown that the presence of an LCR on stably maintained episomes in cell lines results in histone hyperacetylation over the entire minichromosome (Madisen et al. 1998). The establishment of an active chromatin domain therefore devolves to how chromatin accessibility and associated modifications are propagated. Mechanisms for achieving this generalized structural alteration are suggested by data from several systems.

First, the phenomenon of mating-type silencing in yeast is known to involve the nucleation of a repressive chromatin structure at silencer elements with subsequent spreading of this structure throughout a large (>100 kb) domain (for review, see Loo and Rine 1995; Lustig 1998). Spreading of silenced chromatin occurs by

the direct binding over the entire domain of a complex made up of SIR proteins, and is associated with histone hypoacetylation. Similar nucleation and spreading effects have been postulated to account in part for Polycomb-mediated gene silencing in *Drosophila* and for the silencing of genes located in or near heterochromatin in most eukaryotes (for review, see Wakimoto 1998). In like fashion, the establishment of an open chromatin domain could involve the association of nonhistone chromosomal proteins in complexes along the entire domain. Studies of gene silencing also suggest another possible mechanism for achieving open chromatin: In some cases, silencing of transgenes has been correlated to their proximity to centromeric heterochromatin in the nucleus (Csink and Henikoff 1996; Dernburg et al. 1996; Brown et al. 1997; C. Francastel, M. Groudine, and D.I.K. Martin, in prep.). Thus, a large region of DNA may be condensed into inactive chromatin or decondensed into an active domain via localization within specific nuclear compartments enriched in histone deacetylases, acetyltransferases, or other factors.

Yet another possibility is presented by the observation of intergenic transcripts throughout the endogenous  $\beta$ -globin locus (Ashe et al. 1997). Notably, intergenic transcription only occurs in the same direction as  $\beta$ -globin gene transcription, and terminates near  $\beta$ -globin gene promoters with subsequent initiation just 3' of the genes. In transgenes under the control of 5' HS2 of the  $\beta$ -globin locus, non-mRNA transcription originates within the enhancer and proceeds toward the linked gene regardless of orientation or position of the enhancer (Kong et al. 1997). Based on these observations, it has been suggested that intergenic transcription may be a general feature of active loci. Such transcription could indicate a functional role for non-mRNA in the nucleus, as has been demonstrated for *XIST* RNA in X-chromosome inactivation (Panning and Jaenisch 1998). Alternatively, the process of transcription initiated within the LCR and elsewhere and proceeding through the  $\beta$ -globin locus could serve to open the region by disrupting chromatin structure.

As applied to the problem of chromatin opening by LCRs, each of these potential mechanisms is highly speculative. It should be noted, however, that these possibilities are not incompatible with each other. Silencing at the mating-type loci in yeast, for example, is associated both with a silenced chromatin structure and with localization of silenced regions near the nuclear periphery, suggesting that both nuclear compartments and spreading of specific chromatin structures are involved in the mechanism of regulation of the silenced domains (for review, see Cockell and Gasser 1999). In fact, it has been shown that a mutant silencer element can be rescued by tethering the DNA to the nuclear periphery; notably, silencing in this context still requires the SIR proteins and at least a partial silencer element (Andrulis et al. 1998). In like fashion, LCR-mediated establishment of a nuclear compartment, or localization of a domain within such a compartment, may be a necessary prerequisite for the propagation of an open chromatin structure

throughout the locus by the association of chromatin-bound complexes enriched in that compartment. Thus, it becomes important to examine if any experiments to date rule in or out any of these possible mechanisms of LCR-mediated chromatin opening.

The distinction between open (nuclease-sensitive) and closed (nuclease-insensitive) chromatin domains also raises a more formal problem. As the preceding discussion illustrates, the only generalized alterations in chromatin structure that have been shown to date to be active processes are those that lead to gene repression or silencing. Thus, it is possible that the role of LCRs and enhancers in the establishment of active chromatin domains has less to do with the propagation of a positive signal than with the inhibition or exclusion of negative ones. Prior to the commitment to a specific lineage, multipotential hematopoietic progenitor cells initiate gene expression from several different differentiation programs (Hu et al. 1997). Commitment to erythroid differentiation, for example, then presumably results in the loss of nonerythroid gene expression from these cells. The requirement for enhancer function is itself developmentally acquired, coincident with the onset of chromatin-mediated repression: In one-cell mouse embryos, transcription of injected reporter genes does not require an enhancer or LCR at all (Nothias et al. 1995). The data might suggest that in the context of development, the structure of most genomic loci is plastic until the cell is committed to a specific pathway of differentiation, whereupon expression from the majority of loci is prevented by the formation of repressive chromatin structures. Tissue-specific LCRs would then be required to prevent such structures from forming at the minority of loci from which expression is required for the particular cell type. Essentially, however, this 'repressive' view of general chromatin structure and differentiation merely restates the fundamental problem—how does an LCR prevent the formation of repressive chromatin structures over broad regions? The possible mechanisms described above also apply here.

### LCRs and gene activation

The phenomenon of chromatin opening appears to be separable from gene activation. For example, the  $\alpha$ -globin genes reside in a region that is open in a wide variety of tissues, yet their transcription is erythroid specific (Vyas et al. 1992). Deletion of the human  $\beta$ -globin HS2–5 results in complete loss of gene expression, but the chromatin structure in erythroid cells is still nuclease-sensitive (Reik et al. 1998). Thus LCR-mediated gene activation involves at least two distinct events, a potentiation event associated with chromatin opening and an activation event that is presumably related to the distance- and position-independent activity of enhancers.

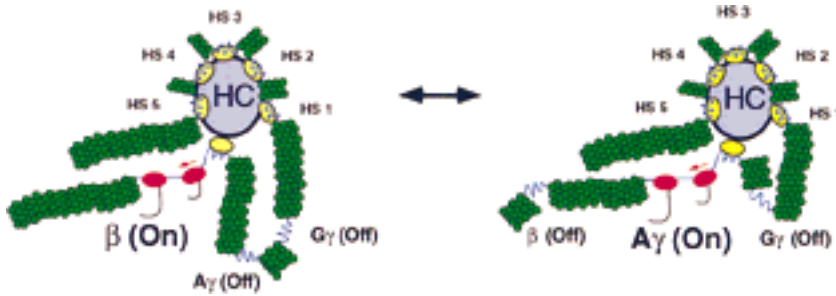
The long-distance activity of enhancers appears to involve an effect on the probability that a promoter will be active rather than on the rate at which transcription occurs (Moreau et al. 1981; Weintraub 1988; Moon and Ley 1991; Walters et al. 1996; Sutherland et al. 1997). The

presence of an enhancer in transgenes in mice or in cell lines is commonly observed to increase the proportion of expressing cells, rather than the levels of expression in each cell. This principle also applies to the  $\beta$ -globin LCR in transgenic mice where, depending on the site of integration, mutations in the LCR result in decreases in globin gene expression by two different mechanisms: a classic position effect in which expression is seen in only a subset of cells or a general decrease in expression in all cells that has been suggested to result from intermittent pulses of transcription (Milot et al. 1996). This binary activity of enhancers and LCRs could simply be related to chromatin opening, and in fact transcriptional activity of transgenes has been correlated to the nuclease-sensitive state (Garrick et al. 1996; C. Francastel, M. Groudine, and D.I.K. Martin, in prep.). On the other hand, the maintenance of nuclease sensitivity observed after deletion of the human  $\beta$ -globin LCR, even in the absence of detectable transcription (Reik et al. 1999), suggests that the switch between active and inactive states could also take place in the context of "open" chromatin. In support of this, expression of the *lacZ* gene under the control of HS2 or HS3 of the  $\beta$ -globin LCR at ectopic sites has been observed to oscillate between "on" and "off" states in the absence of detectable changes in nuclease sensitivity (Feng et al. 1999).

The mechanism of activation of  $\beta$ -globin gene transcription by the LCR has been the subject of a great deal of speculation on the basis of studies utilizing transgenes in mice. In the predominant model for  $\beta$ -globin gene regulation, illustrated in Figure 2, individual globin gene promoters compete for the activity of an LCR-bound 'holocomplex,' and thus gene activation occurs via direct interactions between transcription factors bound to the LCR and those bound near the promoters, with intervening DNA looped out. The looping model represents an extension of the principle of protein-protein interactions that is known to underlie gene activation by promoter-proximal factors in bacteria and yeast (for review, see Ptashne and Gann 1997), but as will be seen, the behavior of metazoan enhancers and LCRs can be interpreted in other ways as well. An evaluation of the looping model necessitates an examination of each of its components in detail.

### Gene competition and alternating transcription

In the looping model for activation of  $\beta$ -globin genes, individual globin gene promoters are thought to compete for LCR activity. Gene competition was originally proposed to explain the properties of the chicken adult ( $\beta$ ) and embryonic ( $\epsilon$ ) globin genes in transient transfections (Choi and Engel 1988). When transfected with the chicken  $\beta$ -globin enhancer individually (i.e., on constructs in the arrangements  $E-\epsilon$  and  $E-\beta$ , where  $E$  represents the enhancer), both genes were transcribed in immortalized definitive (adult) erythroid cells, in which the endogenous  $\epsilon$ -globin gene is not expressed. Transfection of a single construct containing the enhancer and both genes (i.e.,  $\epsilon-E-\beta$ ) resulted in inhibition of  $\epsilon$ -globin tran-



**Figure 2.** The looping model of gene activation by the  $\beta$ -globin LCR. Shown here is a simplified representation of the transgenic human  $\beta$ -globin locus in mouse fetal liver, in which the  $\gamma$ - and  $\beta$ -globin genes are transcribed in most cells. For purposes of simplicity, we have only represented HS1–HS5 on this schematic representation. Protein complexes bound to each of the nuclease hypersensitive sites (HSs 1–5) contribute to the formation of a holocomplex (HC), which contains the site for gene activation. Only one

gene can interact with the holocomplex at any given time, so transcription alternates between  $\gamma$ -globin (*right*) and  $\beta$ -globin (*left*). Processive RNA polymerases are shown as red ovals, and promoter- or HS-bound complexes as yellow ovals. Regions not associated with nuclease hypersensitivity are shown here condensed to the 30-nm fiber. This schematic representation should not be interpreted to indicate that regions condensed to the 30-nm fiber possess a perfectly regular structure, nor that this structure is completely rigid or static. In this simplified schematic, the  $\epsilon$ -globin gene is assumed to be packaged in the region between HS1 and the  $G\gamma$ -globin gene. Because the boundaries or regions of transition from nuclease sensitivity to insensitivity for the human  $\beta$ -globin locus are unknown, the entire locus is shown in the open conformation.

scription, and further experiments demonstrated a strict inverse correlation between the activity of the  $\beta$ -globin promoter and that of the  $\epsilon$ -globin gene (Foley and Engel 1992). It was concluded that the enhancer provides a single limiting activity for which the two promoters compete.

The concept of gene competition has been extended to a number of situations in which multiple genes reside near a single enhancer. For the  $\beta$ -globin genes in mammals, it has been shown that the relative position of the genes with respect to the LCR contributes to their respective levels of expression. In transgenes, the presence of a  $\gamma$ -globin gene proximal to the LCR (i.e., LCR- $\gamma$ - $\beta$  vs. LCR- $\beta$ - $\gamma$ ) inhibits transcription of a more distal  $\beta$ -globin gene (Hanscombe et al. 1991; Peterson et al. 1993). In transgenic mice harboring the human  $\beta$ -globin locus, a marked  $\beta$ -globin gene inserted near the LCR is expressed at higher levels than the same gene when it is located more distal, and in both positions partially inhibits the expression of genes located downstream of it (Dillon et al. 1997). In another study, the entire region containing the  $\beta$ -globin genes was inverted with respect to the LCR. This configuration resulted in aberrant expression of adult  $\beta$ -globin in embryonic yolk sac, an inversion of the relative expression levels of the two  $\gamma$ -globin genes, and complete loss of  $\epsilon$ -globin expression (Tanimoto et al. 1999). Taken together, the data suggest that genes located in proximity to the LCR are preferentially activated by it, and thus that expression of the adult  $\beta$ -globin genes in earlier developmental stages is prevented by the competitive advantage of the more proximal  $\epsilon$ - and  $\gamma$ -globin genes.

Gene competition can only comprise part of the mechanism of developmental specificity, however. For example, the developmental specificity of  $\beta$ -globin gene transcription is unaffected upon deletion of the  $\beta$ -globin LCR in mouse, although the levels of transcription decrease significantly (Epner et al. 1999; M.A. Bender and M. Groudine, in prep.). Transgenic human embryonic ( $\epsilon$ ) and fetal ( $\gamma$ ) globin genes are silenced in adult mice both in the presence and absence of an adult  $\beta$ -globin gene,

indicating that gene competition is not required for the silencing of these genes in the adult (Raich et al. 1990; Dillon and Grosveld 1991). Thus it is only the suppression of adult ( $\beta$ ) globin transcription in the embryo that appears to involve a requirement for the presence of other genes. These results stand in direct contrast to the original transient transfections utilizing the chicken  $\beta$ -globin genes. In these experiments, the adult  $\beta$ -globin gene was required to prevent expression of the embryonic  $\epsilon$ -globin in definitive erythroid cells, whereas  $\epsilon$ -globin did not prevent expression of  $\beta$ -globin in primitive erythroid cells (Choi and Engel 1988). The results could reflect a fundamental difference between chickens and mammals in developmental regulation of the  $\beta$ -globin genes, but transgenic chicken  $\beta$ -globin loci in mouse are correctly regulated (Mason et al. 1995) and exhibit autonomous silencing of the embryonic  $\epsilon$ -globin gene in the adult (Foley et al. 1994). Thus, although the phenomena observed in transient transfections can be explained by gene competition, they may not reflect globin gene switching as it occurs during development.

Observations of gene regulation in *Drosophila* embryos suggest alternatives to gene competition in the regulation of expression from multigene loci. The promoter of the *Drosophila even-skipped (eve)* gene inhibits transcription of another *eve* promoter when located between it and an enhancer, but not when the two promoters are on opposite sides of the enhancer (Ohtsuki and Levine 1998). This boundary or insulator function is dependent upon binding of the product of the *trithorax-like* gene to a GAGA element located between the TATA box and transcription initiation site of the *eve* promoter. The demonstration that an active promoter can act as a boundary element has important implications for the regulation of expression from multigene loci such as the  $\beta$ -globin locus in mammals. Most investigations of  $\beta$ -globin gene competition have not distinguished between gene competition and boundary function, and intrinsic boundary activity of  $\beta$ -globin gene promoters would explain all of the existing data that is currently interpreted as  $\beta$ -globin gene competition. For example,

the observation that a gene more proximal to the LCR is preferentially activated, at the expense of distal genes, is easily interpreted as boundary activity. Boundary activity by active  $\beta$ -globin gene promoters also more readily accounts for the typical genomic organization of the  $\beta$ -globin genes in mammals, in which the genes are arranged in the order of their developmental expression.

With the exception of the chicken  $\beta$ -globin genes, and then only in transient transfections, gene competition has been demonstrated rigorously only between different classes of promoters. Enhancers tested thus far appear to favor promoters containing strong consensus TATA boxes over those with weak TATA elements and strong consensus initiator sequences (Ohtsuki and Levine 1998; Ohtsuki et al. 1998), whereas all of the mammalian  $\beta$ -globin genes possess similar TATA elements. On the other hand, the assays employed in these studies are not quantitative, so competition between genes with similar promoters cannot be ruled out. Gene competition and boundary activity by promoters are not incompatible, and thus both processes could be at work at any given locus. Regardless, the demonstration that promoters can act as boundaries introduces a possible new level of complexity to gene regulation at the  $\beta$ -globin locus and other multigene loci that needs to be examined.

In the competitive model for  $\beta$ -globin gene regulation, it logically follows that the LCR can only drive expression of one gene at a time. In support of this, it has been shown that in transgenic mice harboring the human  $\beta$ -globin locus, while most cells in the fetal liver possess mRNA for both  $\gamma$  and  $\beta$ -globin, only a relatively small proportion display unspliced primary transcripts indicative of ongoing transcription of both genes (Wijgerde et al. 1995). Further experiments have utilized DRB, a reversible inhibitor of transcriptional elongation (Gribnau et al. 1998; Trimborn et al. 1999). Upon release of fetal liver cells from DRB inhibition, a reproducible lag is observed in the reappearance of double signals, as compared to single signals indicating transcription from a single globin gene. The results have been interpreted to suggest that double signals arise from fortuitous overlap when a cell has just switched from transcription of one gene to transcription of the other, and therefore that transcription from within the  $\beta$ -globin locus dynamically alternates between individual genes.

The existing data as yet do not require that transcription alternate between the  $\beta$ -globin genes. For example, if transcription of the  $\beta$ -globin genes is not continuous, but intermittent, it is not clear what proportion of double signals is to be expected. DRB can have effects on transcription initiation and is known to cause drastic changes in nuclear organization (Haaf and Ward 1996). Also, for many genes transcriptional pausing is a general rate-limiting step after transcription initiation (Krumm et al. 1995; for review, see Lis 1998). Although the possible regulation of  $\beta$ -globin transcription at the level of elongation has not been systematically examined, the mouse  $\beta_{maj}$  globin gene has been observed to possess paused RNA polymerase molecules in uninduced cells that lack full-length  $\beta_{maj}$  mRNA transcripts (Sheffery et

al. 1984; Cohen and Sheffery 1985). Thus the results from experiments using DRB could reflect differences in the elongation properties of different  $\beta$ -globin genes. Nevertheless, alternating transcription remains an interesting hypothesis. While such behavior fulfills the predictions of a competitive model, it should be noted that other models can accommodate it (see below). For example, if the active  $\gamma$ -globin promoter acts as a boundary element, the LCR would only activate the  $\beta$ -globin gene when the  $\gamma$ -globin gene was not transcribed.

### The holocomplex

If individual genes compete for a limiting activity provided by an LCR, the size and complexity of some LCRs becomes an issue of importance. The  $\beta$ -globin LCR is comprised of at least four HSs distributed over a span of 20–30 kb located 5' of the  $\beta$ -globin genes in mammals, and some other LCRs are similarly comprised of multiple elements spread over a broad region. Large (1–2 kb) deletions that encompass individual HSs within the endogenous mouse or transgenic human  $\beta$ -globin LCRs indicate that loss of any single site does not result in drastic decreases in gene expression, or in developmental- or gene-specific effects (Fiering et al. 1995; Hug et al. 1996; Peterson et al. 1996). In part to account for this, it has been proposed that factors bound to each of the LCR HSs contribute cooperatively to the formation of a larger physical entity termed the 'holocomplex' (Grosveld et al. 1993; Bungert et al. 1995, 1999). Individual HS deletions may lead to small decreases in the stability of this complex, but do not serve to abolish it. In this manner, the multiple dispersed elements of the LCR are believed to comprise a single limiting activity within the  $\beta$ -globin locus.

Recently, it has been shown that small (200–400 bp) deletions confined to the core regions of HS2, HS3, or HS4 in YAC transgenes containing the human  $\beta$ -globin locus result in more severe phenotypes than the larger HS deletions (Bungert et al. 1995, 1999; Navas et al. 1998). In addition, in some cases the effects of these deletions vary widely on individual globin genes and/or in different developmental stages. Based on these results, specific interactions between different HSs and promoters within the context of a holocomplex have been considered. These conclusions are not supported, however, by the fact that the phenotypes of the different core deletions are virtually identical, which suggests that these phenotypes are a general property of small core deletions as compared to larger deletions. In support of this, the HS core deletions caused the ablation of the remaining HSs (Li et al. 1998; Bungert et al. 1999), whereas in the larger deletion of core enhancer and flanking region, the remaining HSs still form (Fiering et al. 1995, and unpubl). Similarly, there is no evidence to suggest that formation of one LCR HS requires the presence of any of the others. The simplest explanation for this discrepancy is that in the absence of the core region of an HS the remaining flanking sequences behave as a dominant chromatin silencer. It is therefore notable that in some of the core

deletion experiments the most severe decreases in expression were observed for the normally active genes most proximal to the LCR (Navas et al. 1998; Bungert et al. 1999). In transgenic human  $\beta$ -globin loci,  $\epsilon$ - and  $\gamma$ -globin are expressed in the embryonic yolk sac and  $\gamma$ - and  $\beta$ -globin in the fetal liver (Fig. 1), whereas upon deletion of the core of either HS2 or HS3,  $\epsilon$ -globin is preferentially lost in the yolk sac and  $\gamma$ -globin in the fetal liver. These results are consistent with the hypothesis that a silenced chromatin structure spreads from the site of the core deletion. The lack of gene-specific effects in some cases (Bungert et al. 1995) could reflect position effects that favor repression across the entire locus.

The available data indicate that elements within the  $\beta$ -globin LCR cooperate and are at least partially redundant. As yet, however, a holocomplex formed by physical interactions between factors bound to different HSs is supported by no direct evidence. The formation of multiprotein complexes at the TCR $\alpha$  enhancer and the IFN $\beta$  promoter is highly sensitive to changes in the arrangement of the component binding sites (for review, see Grosschedl 1995). Since it is known that 'mini' and 'micro'  $\beta$ -globin LCRs, comprised of individual HS sequences artificially linked in tandem, are also capable of conferring high-level expression in transgenic assays (for example, see Enver et al. 1990), a putative LCR holocomplex would involve interactions between DNA-bound factors that are far more flexible and distance independent, and therefore of a different character than any transcription factor interactions that have been characterized. This description also applies to the hypothesis of DNA looping as the mechanism of enhancer-promoter interactions.

### Direct LCR-promoter interactions and DNA "looping"

The basic principle of gene activation by promoter-proximal DNA-binding proteins is recruitment (Ptashne and Gann 1997). This principle is derived from studies of activating regions in bacteria (also referred to as enhancers) and from eukaryotic promoters and upstream activating sequences (UASs) in yeast. Transcription factors bound to eukaryotic promoters are known to interact with chromatin modifying activities such as histone acetyltransferases and remodeling complexes such as SWI/SNF, and also appear to bind to components of the basal transcription machinery to facilitate the formation of active transcription complexes. The promoter of the IFN $\beta$  gene, among others, has been shown to depend upon a highly specific arrangement of interacting proteins in a structure termed an enhanceosome (Thanos and Maniatis 1995). Thus a wealth of evidence indicates that direct physical interactions are responsible for mediating short-range gene activation at promoters.

The looping model proposes that the action of enhancers and LCRs represents an extension of this basic principle. In support of this, coinjection experiments in *Xenopus* oocytes have demonstrated that an enhancer is capable of activating a promoter in trans when both elements are located on separate concatenated plasmids

(Dunaway and Droge 1989). Transactivation by enhancers has also been observed when enhancer and promoter are artificially bridged by a protein-linked biotin:streptavidin complex (Muller et al. 1989). At the  $\beta$ -globin locus, the LCR holocomplex is postulated to be sequestered by direct interactions with  $\beta$ -globin promoters, with looping out of the intervening DNA. Presumably, the holocomplex provides a limiting activity that the promoters alone are lacking.

Unfortunately, no experiment has yet been presented for which such DNA looping has predictive value. For example, gene competition and alternating transcription are commonly held as evidence for direct LCR-promoter interactions, but it is difficult to see how the absence of gene competition, or cotranscription of multiple genes, could be used to rule out such interactions. In addition, other models are equally adept at providing for gene competition and/or alternating transcription (see below). The experiments utilizing concatenated plasmids or a protein bridge demonstrate that enhancers can act in trans when they are artificially localized near a promoter, but the questions of if and how such colocalization occurs at endogenous loci have not been experimentally addressed.

There are indications that true enhancer activity cannot be accounted for by simple direct interactions. Bacterial enhancer and yeast UAS activities are drastically affected by increasing distance from the promoter, as would be expected from the probability of random interactions of freely diffusible elements, and do not function when located 3' of a gene or within an intron. In contrast, mammalian enhancers and LCRs are in part defined by their ability to function in a position-independent manner over large distances. For a looping mechanism to overcome the prohibitive effects of increasing distance, an activity must exist to facilitate the enhancer-promoter interaction. For the  $\beta$ -globin LCR, where more than one gene is activated within the same locus, perhaps by an alternating mechanism, LCR-promoter interactions must be both facilitated and dynamic. In essence, chromatin within an active  $\beta$ -globin locus cannot be freely diffusible, and a looping mechanism must involve a significant energy expenditure on the part of the cell.

An argument for a facilitator component to enhancer function is also presented by comparisons of yeast and mammalian transcriptional activators (Struhl 1989). Yeast activators bound to UASs can only activate transcription when located 5' of a promoter and in close proximity (<1500 bp). Similarly, mammalian factors expressed in yeast are also limited in this way. Both mammalian and yeast factors, however, can act as true position-independent enhancers in mammalian cells, suggesting that the absence of enhancer function in yeast does not reflect an intrinsic difference between activator proteins in yeast and mammals. Therefore, either yeast lack a fundamental activity present in mammals that facilitates long-range DNA looping, or gene activation by mammalian enhancers is facilitated by a mechanism entirely distinct from DNA looping. A candidate for an

enhancer facilitator has been provided by studies of boundary elements in *Drosophila*.

### Boundary elements and enhancer facilitators

Boundary elements, or insulators, have been defined in part by their ability to block enhancer activity when located between an enhancer and a promoter (for review, see Geyer 1997). The most well-characterized examples of such elements include *scs/scs'*, originally identified in the *Drosophila* 87A7 heat shock locus, the *gypsy* retrotransposon, also in *Drosophila*, and in vertebrates 5' HS4 from the chicken  $\beta$ -globin LCR (Chung et al. 1993). Each element has been shown to possess a position-dependent ability to prevent promoter activation by enhancers, although some evidence exists to indicate that such activity may not be equivalent at all genomic loci (Walters et al. 1999).

Within the *gypsy* retrotransposon, boundary activity requires DNA binding sites for the protein encoded by the *suppressor of hairy-wing* [*su(Hw)*] gene, and these binding sites are also sufficient for this activity by themselves. Studies utilizing the *gypsy* retrotransposon or Su(Hw) binding sites have demonstrated that boundary activity does not involve inactivation or sequestration of either the enhancer or promoter, nor a generalized alteration of chromatin structure (Cai and Levine 1995; Scott and Geyer 1995). Thus, boundary elements appear to act purely by disrupting enhancer–promoter communication, and the mechanism by which they accomplish this should provide insight into the nature of enhancer function.

Within the context of looping models for enhancer function, two general mechanisms have been proposed to explain the activity of boundary elements. One, termed the “decoy” model (Geyer 1997), postulates that the boundary interacts with the enhancer directly and thus prevents it from interacting with a gene promoter farther away (the parallel with globin gene competition should be obvious). However, the observation that an enhancer blocked on one side by a boundary element still is capable of activating a promoter on the other side seems incomplete with this model, unless the enhancer actively scans or tracks along the DNA in search of a promoter, as has been proposed in some models for enhancer function (see, for example, Blackwood and Kadonaga 1998).

In the second proposed mechanism for boundary function, boundary elements serve to establish the limits of distinct chromatin domains that are then rendered inaccessible to each other (Gerasimova and Corces 1996; see also Felsenfeld 1999). The *scs/scs'* insulators, however, do not block site-specific recombination by the F1p recombinase, which involves the close apposition of two *ftt* sequences (Dunaway et al. 1997). This is not surprising, given that other site-specific recombinases such as Cre are known to be capable of acting interchromosomally, whereas enhancers are only known to activate promoters on separate chromosomes in the phenomenon of transvection, in which homologous pairing is required (Geyer et al. 1990). The data therefore suggest that

boundary elements function by blocking the transmission of a signal from enhancer to promoter that occurs along the intervening DNA. Phenomena such as transvection or enhancer–promoter interactions between concatenated plasmids then presumably involve the transfer or ‘jumping’ of this signal between DNA molecules that are brought close together artificially or by homologous pairing.

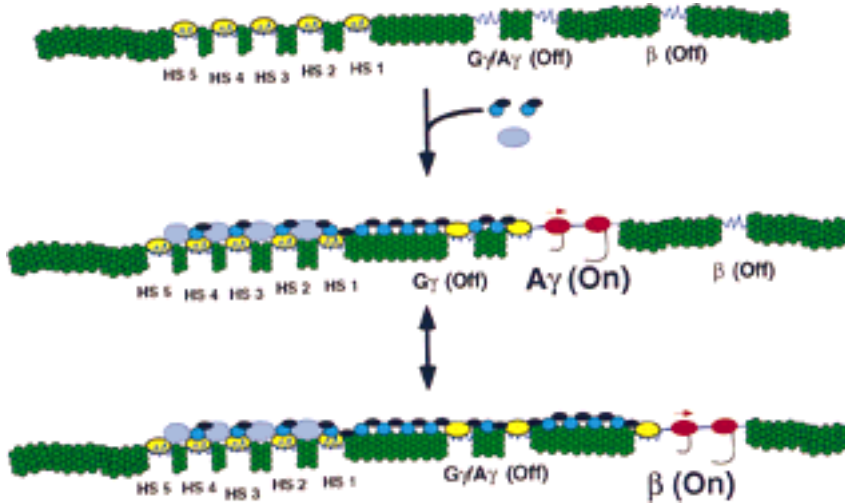
All boundary elements tested thus far also serve another function, that of chromatin insulators. When used to flank reporter constructs in *Drosophila*, boundary elements insulate randomly integrated transgenes from position effects (Kellum and Schedl 1991; Chung et al. 1993; Roseman et al. 1993). The insulation properties of these elements are believed to reflect the ability to block the spreading of inhibitory or closed chromatin structures from neighboring loci, and in fact in some cases boundary elements are known to coincide with changes in patterns of histone acetylation and/or nuclease sensitivity (Hebbes et al. 1994; Pikaart et al. 1998). This activity therefore might suggest that enhancer function is mediated by transmission of specific nucleoprotein structures between enhancer and promoter, which a boundary element is then capable of blocking in much the same way as it blocks the spread of silenced chromatin.

A potential candidate for a protein involved in this transmission is Chip, which was isolated originally in a screen for factors involved in the function of the wing margin enhancer of the *Drosophila cut* locus that might also be targets of the Su(Hw) insulator (Morcillo et al. 1996, 1997). Mutations in Chip reveal not only that this factor directly regulates the wing margin enhancer, but also that it displays a strong genetic interaction with *gypsy* insertions between the enhancer and the *cut* promoter. In fact, in mutants heterozygous for Chip, a *gypsy* insertion in one *cut* allele has a greater phenotypic effect than a heterozygous deletion of the enhancer, suggesting that in the presence of lower levels of Chip, a *gypsy* insertion in one allele is capable of blocking the function of the enhancer on the other allele in a phenomenon akin to transvection. Chip also appears to be a ubiquitous protein required for the proper regulation of a large variety of genes in *Drosophila*. On this basis, Chip has been suggested to act as a general facilitator of enhancer–promoter communication.

In the context of the looping model, the evidence that enhancer–promoter communication involves a signal propagated along the intervening DNA suggests that an enhancer facilitator such as Chip acts by organizing the intervening sequences to allow enhancer–promoter interactions to occur (Dorsett 1999). The characterization of Chip also suggests an alternative that is distinct from looping, which we term enhancer–promoter linking (Fig. 3). The basis for this model is provided by the behavior of Chip-like factors in vertebrates.

### Enhancer facilitators and the linking model

The closest known vertebrate homologs of Chip are fac-



**Figure 3.** The linking model of gene activation by the  $\beta$ -globin LCR. A simplified representation of a transgenic human  $\beta$ -globin locus in mouse fetal liver is shown, as in Fig. 2. Prior to induction in erythroid cells, chromatin is open (*top*). Upon induction of transcription during differentiation, a chain of CHIP-containing complexes (black ovals) bound to chromatin via interactions with homeodomain transcription factors or other proteins (blue circles) is nucleated at the LCR by higher-order protein complexes (purple ovals) and extends along the chromatin fiber between the LCR and the gene to be transcribed. In this figure, we have assumed that transcription alternates, and so when the  $A\gamma$  gene is transcribed (*middle*) the promoter-bound complex acts as a boundary to the further spread of the chain to more distal genes. When the  $A\gamma$  gene is inactive and  $\beta$  is transcribed (*bottom*), the  $A\gamma$  promoter is shown as another link in the chain. The silenced  $\epsilon$ -globin gene promoter is not shown but could also function in this way.

tors that are known to interact with LIM domains of nuclear proteins (Morcillo et al. 1997). LIM domains are zinc-finger motifs involved in a variety of protein-protein interactions (for review, see Dawid et al. 1998), and proteins possessing LIM domains include LIM-HD factors that also contain the homeodomain (HD) DNA-binding motif. The LIM domain of at least some homeodomain transcription factors inhibits DNA binding, and coinjection experiments in *Xenopus* have demonstrated that the Chip homolog Xldb1 relieves this inhibition (Breen et al. 1998). In addition, vertebrate Chip homologs can also associate with HD transcription factors that lack LIM domains. In erythroid cells, the Chip homolog Ldb1 has been characterized as part of a single DNA-binding complex that includes the LIM-domain protein Lmo2, the E-box transcription factor Tal-1, and the erythroid transcription factor GATA-1 (Wadman et al. 1997). Consistent with the behavior of its mammalian homologs, Chip also interacts directly with HD transcription factors, as well as with Su(Hw) (E. Torigoi, C. Rosen, and D. Dorsett, pers. comm.). The available evidence therefore indicates that Chip and homologous vertebrate LIM-domain binding proteins are generally involved in transcriptional regulation through interactions with many different factors.

Homeodomain proteins are known to bind in vitro to a broad range of different DNA sequences with similar affinity (Desplan et al. 1988; Walter and Biggin 1996; Carr and Biggin 1999). Consistent with this, in vivo cross-linking experiments have demonstrated that HD transcription factors bind uniformly throughout the gene loci they are known to regulate, and at lower levels to a large proportion of active genes in the *Drosophila* embryo (Walter et al. 1994; Orlando et al. 1997). This property contrasts with that of most non-HD transcription

factors, which commonly bind only to short regions such as enhancers and promoters. Consistent with this behavior, there is evidence that HD transcription factors are directly involved in the regulation of nearly all active genes in the *Drosophila* embryo (Liang and Biggin 1998; Carr and Biggin 1999). From these observations it has been suggested that multiple HD proteins binding throughout a chromatin domain serve to organize the structure of the locus and/or to communicate the effects of enhancers and other regulatory elements.

On the basis of the properties of Chip, its vertebrate homologs, and the proteins with which they interact, it has been proposed that communication between an LCR and a promoter is mediated by a chain of complexes containing Chip-related molecules that are anchored to the chromatin fiber by interactions with HD proteins (Marcillo et al. 1997; Dorsett 1999). A simple extension of this model is that mammalian LCRs and promoters communicate in a similar fashion. Alternatively, Chip-containing complexes might bind directly to relatively hyperacetylated chromatin in the absence of HD transcription factors, perhaps by recognizing a feature of active loci such as histones acetylated in a specific pattern. This possibility presents an interesting counterpoint to repression of transcription from the mating-type loci in yeast, in which the binding of complexes of SIR proteins to relatively hypoacetylated chromatin may also involve specific histone acetylation patterns (for review, see Grunstein 1998; Lustig 1998). In either case, the linking model proposes that a promoter by itself is unable to recruit a crucial, limiting activity that is only provided after initiation of a chain of these Chip-containing complexes by an LCR. In loci where regulatory elements are dispersed across large regions, each distinct element serves to reinforce an overall structure that is propagated

throughout the locus. Gene promoters can also serve this function, although by themselves they have little to no enhancer activity. Importantly, the only looping that occurs is that which is accomplished by the packaging of DNA into the chromatin fiber.

### Conclusions: LCRs and model building

The linking model accounts for all of the features of LCR-mediated gene regulation that we have described. With a chain of protein complexes between enhancer and promoter, distance is far less a factor than with simple looping interactions. However, it should be noted that the assembly of such a structure on the chromatin fiber could involve at least as great an energy expenditure as would the facilitation of direct enhancer–promoter contacts. In this scenario, boundary elements function via interactions with Chip that prevent the extension of the chain to the promoter. Similarly, the function of certain gene promoters as boundaries depends upon the ability to prevent extension of Chip complexes to more distal genes; in the absence of promoter boundary function, a complex at the proximal promoter may actually function as another link in the chain. Alternating transcription, if it occurs, thus reflects a switch between these two states at the proximal promoter.

The linking model also suggests a reason for the lack of true long-distance enhancer activity in yeast. Based on comparisons of in vitro and in vivo DNA-binding patterns, it is known that homeoproteins in yeast bind DNA in vivo only in conjunction with cofactors (Johnson 1992), whereas *Drosophila* homeoproteins may not require such cofactors, which has been suggested as an explanation for their more generalized binding patterns (Carr and Biggin 1999). This difference in in vivo binding activities could reflect the absence of the appropriate LIM-domain binding proteins, such as Chip, that in *Drosophila* may serve to allow the binding of HD or other factors throughout an active locus, or may directly bind to chromatin throughout the locus themselves.

Recently, another putative enhancer facilitator, termed Nipped-B, has been isolated from the same genetic screen that resulted in the discovery of Chip (Rollins et al. 1999). Nipped-B, like Chip, is directly involved in regulation of the *Drosophila cut* gene by the wing margin enhancer, and genetically interacts with the *gypsy* insulator. Nipped-B encodes a homolog of proteins from yeast and fungi that function in chromosome condensation and DNA repair, which has led to the suggestion that this factor plays a structural role in enhancer–promoter interactions, although the precise nature of this role is as yet not clear.

In the end, a useful hypothesis for the function of LCRs in gene regulation must be a testable one. It is perhaps ironic that while the more specific molecular mechanisms of LCR and enhancer function (such as histone modification and nucleosome remodeling) are currently the subject of important and continuing advances in our understanding, the overall framework within which such mechanisms work (looping, linking, or other

models we have not discussed) remains refractory to direct experimentation. The fundamental problem with both the looping and linking models is not that experimental evidence exists that one or the other model cannot explain, but rather that no experimental evidence yet exists that could be used to distinguish between them. Perhaps the most promising avenue toward this end is provided by the identification of Chip and Nipped-B as general facilitators of enhancer function in *Drosophila*. The unique combination of genetic and molecular approaches to enhancer function provided in *Drosophila*, as applied to the question of enhancer facilitators, may allow us to address for the first time the fundamental basis of long-distance gene regulation.

### Acknowledgments

We thank Dale Dorsett, Rohinton Kamakaka, M.A. Bender, Dan Cimbara, Tony Krumm, Andreas Reik, and Dirk Schubeler for critical reading and discussion of the manuscript. M.B. is supported by the Helen Hay Whitney Foundation. Work in the laboratory of M.G. is supported by National Institutes of Health grants DK44746, HL57620, and CA54337.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked 'advertisement' in accordance with 18 USC section 1734 solely to indicate this fact.

### References

- Andersson, K., R. Mahr, B. Bjorkroth, and B. Daneholt. 1982. Rapid reformation of the thick chromosome fiber upon completion of RNA synthesis at the Balbiani ring genes in *Chironomus tentans*. *Chromosoma* **87**: 33–48.
- Andrulis, E.D., A.M. Neiman, D.C. Zappulia and R. Sternglanz. 1998. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* **394**: 592–595.
- Ashe, H.L., J. Monks, M. Wijgerde, P. Fraser, and N.J. Proudfoot. 1997. Intergenic transcription and transinduction of the human  $\beta$ -globin locus. *Genes & Dev.* **11**: 2494–2509.
- Bell, A.C. and G. Felsenfeld. 1999. Stopped at the border: Boundaries and insulators. *Curr. Opin. Genet. Dev.* **9**: 191–198.
- Bjorkroth, B., C. Ericsson, M.M. Lamb, and B. Daneholt. 1988. Structure of the chromatin axis during transcription. *Chromosoma* **96**: 333–340.
- Blackwood, E.M. and J.T. Kadonaga. 1998. Going the distance: A current view of enhancer action. *Science* **281**: 60–63.
- Breen, J.J., A.D. Agulnick, H. Westphal, and I.B. Dawid. 1998. Interactions between LIM domains and the LIM domain-binding protein Ldb1. *J. Biol. Chem.* **273**: 4712–4717.
- Brown, K.E., S.S. Guest, S.T. Smale, K. Hahm, M. Merken-schlager, and A.G. Fisher. 1997. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* **91**: 845–854.
- Bulger, M., J.H. von Doorninck, N. Saitoh, A. Telling, C. Farrell, M.A. Bender, G. Felsenfeld, R. Axel, and M. Groudine. 1999. Conservation of sequence and structure flanking the mouse and human beta-globin loci: The beta-globin genes are embedded within an array of odorant receptor genes. *Proc. Natl. Acad. Sci.* **96**: 5129–5134.
- Bungert, J., D. Utpal, K.-C. Lim, K.H. Lieuw, J.A. Shavit, Q. Liu, and J.D. Engel. 1995. Synergistic regulation of human  $\beta$ -globin gene switching by locus control region elements HS3 and

- HS4. *Genes & Dev.* **9**: 3083–3096.
- Bungert, J., K. Tanimoto, S. Patel, Q. Liu, M. Fear, and J.D. Engel. 1999. Hypersensitive site 2 specifies a unique function within the human  $\beta$ -globin locus control region to stimulate globin gene transcription. *Mol. Cell Biol.* **19**: 3062–3072.
- Cai, H. and M. Levine. 1995. Modulation of enhancer-promoter interactions by insulators in the *Drosophila* embryo. *Nature* **376**: 533–536.
- Carr, A. and M.D. Biggin. 1999. A comparison of *in vivo* and *in vitro* DNA-binding specificities suggests a new model for homeoprotein DNA binding in *Drosophila* embryos. *EMBO J.* **18**: 1598–1608.
- Choi, O.-R.B. and J.D. Engel. 1988. Developmental regulation of  $\beta$ -globin gene switching. *Cell* **55**: 17–26.
- Chung, J.H., M. Whiteley, and G. Felsenfeld. 1993. A 5' element of the chicken  $\beta$ -globin domain serves as an insulator in human erythroid cells and protects against position effect in *Drosophila*. *Cell* **74**: 505–514.
- Cockell, M. and S.M. Gasser. 1999. Nuclear compartments and gene regulation. *Curr. Opin. Genet Dev.* **9**: 199–205.
- Cohen, R.B. and M. Sheffery. 1985. Nucleosome disruption precedes transcription and is largely limited to the transcribed domain of globin genes in murine erythroleukemia cells. *J. Mol. Biol.* **182**: 109–129.
- Csink, A.K. and S. Henikoff. 1996. Genetic modification of heterochromatic association and nuclear organization in *Drosophila*. *Nature* **381**: 529–531.
- Dawid, I.B., J.J. Breen, and R. Toyama. 1998. LIM domains: Multiple roles as adaptors and functional modifiers in protein interactions. *Trends Genet.* **14**: 156–162.
- Dernburg, A.F., K.W. Broman, J.C. Fung, W.F. Marshall, J. Phillips, D.A. Agard, and J.W. Sedat. 1996. Perturbation of nuclear architecture by long-distance chromosome interactions. *Cell* **85**: 745–759.
- Desplan, C., J. Theis, and P.H. O'Farrell. 1988. The sequence specificity of homeodomain-DNA interaction. *Cell* **54**: 1081–1090.
- Dillon, N. and F. Grosveld. 1991. Human gamma-globin genes silenced independently of other genes in the beta-globin locus. *Nature* **350**: 252–254.
- Dillon, N., T. Trimborn, J. Strouboulis, P. Fraser, and F. Grosveld. 1997. The effect of distance on long-range chromatin interactions. *Mol. Cell* **1**: 131–139.
- Dorsett, D. 1999. Distant liaisons: Long range enhancer-promoter interactions in *Drosophila*. *Curr. Opin. Genet Dev.* (in press).
- Dunaway, M. and P. Droge. 1989. Transactivation of the *Xenopus* rRNA gene promoter by its enhancer. *Nature* **341**: 657–659.
- Dunaway, M., J.Y. Hwang, M. Xiong, and H.-L. Yuen. 1997. The activity of the scs and scs' insulator elements is not dependent on chromosomal context. *Mol. Cell Biol.* **17**: 182–189.
- Ellis, J., D. Talbot, N. Dillon, and F. Grosveld. 1993. Synthetic human  $\beta$ -globin 5'HS2 constructs function as locus control regions only in multicopy transgene concatemers. *EMBO J.* **12**: 127–134.
- Ellis, J., K.C. Tan-Un, A. Harper, D. Michalovich, N. Yannoutsos, S. Philipsen, and F. Grosveld. 1996. A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human  $\beta$ -globin locus control region. *EMBO J.* **15**: 562–568.
- Enver, T., N. Raich, A.J. Ebens, T. Papayannopoulou, F. Constantini, and G. Stamatoyannopoulos. 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* **344**: 309–313.
- Epner, E., A. Reik, D. Cimbara, A. Telling, M.A. Bender, S. Fiering, T. Enver, D.I.K. Martin, M. Kennedy, G. Keller, and M. Groudine. 1998. The  $\beta$ -globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse  $\beta$ -globin locus. *Mol. Cell* **2**: 447–455.
- Ericsson, C., H. Mehlin, B. Bjorkroth, M.M. Lamb, and B. Daneholt. 1989. The ultrastructure of upstream and downstream regions of an active Balbiani ring gene. *Cell* **56**: 631–639.
- Feng, Y.-Q., R. Alami, and E.E. Bouhassira. 1999. Enhancer-dependent transcriptional oscillations in mouse erythroleukemia cells. *Mol. Cell Biol.* **19**: 4907–4917.
- Fiering, S., E. Epner, K. Robinson, Y. Zhuang, A. Telling, M. Hu, D.I.K. Martin, T. Enver, T.J. Ley, and M. Groudine. 1995. Targeted deletion of 5'HS2 of the murine  $\beta$ -globin LCR reveals that it is not essential for proper regulation of the  $\beta$ -globin locus. *Genes & Dev.* **9**: 2203–2213.
- Fletcher, T.M. and J.C. Hansen. 1996. The nucleosomal array: Structure/function relationships. *Crit. Rev. Eukaryot. Gene Expr.* **6**: 149–188.
- Foley, K.P. and J.D. Engel. 1992. Individual stage selector element mutations lead to reciprocal changes in  $\beta$ - versus  $\epsilon$ -globin gene transcription: Genetic confirmation of promoter competition during globin gene switching. *Genes & Dev.* **6**: 730–744.
- Foley, K.P., S. Pruzina, J.D. Winick, J.D. Engel, F. Grosveld, and P. Fraser. 1994. The chicken  $\beta/\epsilon$  globin enhancer directs autonomously regulated, high-level expression of the chicken  $\epsilon$ -globin gene in transgenic mice. *Proc. Natl. Acad. Sci.* **91**: 7252–7256.
- Forrester, W.C., C. Thompson, J.T. Elder, and M. Groudine. 1986. A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proc. Natl. Acad. Sci.* **83**: 1359–1363.
- Garrick, D., H. Sutherland, G. Robertson, and E. Whitelaw. 1996. Variegated expression of a globin transgene correlates with chromatin accessibility but not methylation status. *Nucleic Acids Res.* **24**: 4902–4909.
- Gerasimova, T.I. and V.G. Corces. 1996. Boundary and insulator elements in chromosomes. *Curr. Opin. Genet Dev.* **6**: 185–192.
- Geyer, P.K. 1997. The role of insulator elements in defining domains of gene expression. *Curr. Opin. Genet Dev.* **7**: 242–248.
- Geyer, P.K., M.M. Green, and V.G. Corces. 1990. Tissue-specific transcriptional enhancers may act in *trans* on the gene located in the homologous chromosome: The molecular basis of transvection in *Drosophila*. *EMBO J.* **9**: 2247–2256.
- Gribnau, J., E. de Boer, T. Trimborn, M. Wijgerde, E. Milot, F. Grosveld, and P. Fraser. 1998. Chromatin interaction mechanism of transcriptional control *in vivo*. *EMBO J.* **17**: 6020–6027.
- Grosschedl, R. 1995. Higher-order nucleoprotein complexes in transcription: Analogies with site-specific recombination. *Curr. Opin. Cell Biol.* **7**: 362–370.
- Grosveld, F., G. Blom van Assendelft, D.R. Greaves, and G. Kollias. 1987. Position-independent, high-level expression of the human  $\beta$ -globin gene in transgenic mice. *Cell* **51**: 975–985.
- Grosveld, F., M. Antoniou, M. Berry, E. de Boer, N. Dillon, J. Ellis, P. Fraser, O. Hanscombe, J. Hurst, A. Imam, M. Lindenbaum, S. Philipsen, S. Pruzina, S. Strouboulis, S. Raguz-Bolognesi, and D. Talbot. 1993. The regulation of human globin gene switching. *Philos. Trans. R. Soc. Lond. Ser. B* **339**: 183–191.
- Grunstein, M. 1998. Yeast heterochromatin: Regulation of its assembly and inheritance by histones. *Cell* **93**: 325–328.
- Haaf, T. and D.C. Ward. 1996. Inhibition of RNA polymerase II

- transcription causes chromatin decondensation, loss of nucleolar structure, and dispersion of chromosomal domains. *Exp. Cell Res.* **224**: 163–173.
- Hanscombe, O., D. Whyatt, P. Fraser, N. Yannoutsos, D. Greaves, N. Dillon, and F. Grosveld. 1991. Importance of globin gene order for correct developmental expression. *Genes & Dev.* **5**: 1387–1394.
- Hardison, R., J.L. Slightom, D.L. Gumucio, M. Goodman, N. Stojanovic, and W. Miller. 1997. Locus control regions of mammalian beta-globin gene clusters: Combining phylogenetic analyses and experimental results to gain functional insights. *Gene* **205**: 73–94.
- Hebbes, T.R., A.L. Clayton, A.W. Thorne, and C. Crane-Robinson. 1994. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *EMBO J.* **13**: 1823–1830.
- Hu, M., D. Krause, M. Greaves, S. Sharkis, M. Dexter, C. Heyworth, and T. Enver. 1997. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes & Dev.* **11**: 774–785.
- Hug, B.A., R.L. Wesselschmidt, S. Fiering, M.A. Bender, E. Epner, M. Groudine, and T.J. Ley. 1996. Analysis of mice containing a targeted deletion of  $\beta$ -globin locus control region hypersensitive site 3. *Mol. Cell. Biol.* **16**: 2906–2912.
- Jenuwein, T., W.C. Forrester, L.A. Fernandez-Herrero, G. Laible, M. Dull, and R. Grosschedl. 1997. Extension of chromatin accessibility by nuclear matrix attachment regions. *Nature* **385**: 269–272.
- Johnson, A. 1992. A combinatorial regulatory circuit in budding yeast. In *Transcriptional regulation* (ed. S.L. McKnight and K.R. Yamamoto), pp. 975–1006. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Kamakaka, R.T. and J.O. Thomas. 1990. Chromatin structure of transcriptionally competent and repressed genes. *EMBO J.* **9**: 3997–4006.
- Kellum, R. and P. Schedl. 1991. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* **64**: 941–950.
- Kong, S., D. Bohl, C. Li, and D. Tuan. 1997. Transcription of the HS2 enhancer toward a *cis*-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. *Mol. Cell. Biol.* **17**: 3955–3965.
- Krumm, A., L.B. Hickey, and M. Groudine. 1995. Promoter-proximal pausing of RNA polymerase II defines a general rate-limiting step after transcription initiation. *Genes & Dev.* **9**: 559–572.
- Li, G., K.-C. Lim, J.D. Engel, and J. Bungert. 1998. Individual LCR hypersensitive sites cooperate to generate an open chromatin domain spanning the human  $\beta$ -globin locus. *Genes Cells* **3**: 415–429.
- Lis, J. 1998. Promoter-associated pausing in promoter architecture and postinitiation transcriptional regulation. *Cold Spring Harbor Symp. Quant. Biol.* **63**: 347–356.
- Liang, Z. and M.D. Biggin. 1998. *Eve* and *ftz* regulate a wide array of genes in blastoderm embryos: The selector homeoproteins directly or indirectly regulate most genes in *Drosophila*. *Development* **125**: 4471–4482.
- Loo, S. and J. Rine. 1995. Silencing and heritable domains of gene expression. *Annu. Rev. Cell Dev. Biol.* **11**: 519–548.
- Lustig, A.J. 1998. Mechanisms of silencing in *Saccharomyces cerevisiae*. *Curr. Opin. Genet Dev.* **8**: 233–239.
- Madisen, L., A. Krumm, T.R. Hebbes, and M. Groudine. 1998. The immunoglobulin heavy chain locus control region increases histone acetylation along linked *c-myc* genes. *Mol. Cell. Biol.* **18**: 6281–6292.
- Martin, D.I., S. Fiering, and M. Groudine. 1996. Regulation of beta-globin gene expression: Straightening out the locus. *Curr. Opin. Genet Dev.* **6**: 488–495.
- Mason, M.M., E. Lee, H. Westphal, and M. Reitman. 1995. Expression of the chicken beta-globin gene cluster in mice: Correct developmental expression and distributed control. *Mol. Cell. Biol.* **15**: 407–414.
- Milot, E., J. Strouboulis, T. Trimborn, M. Wijgerde, E. de Boer, A. Langeveld, K. Tan-Un, W. Vergeer, N. Yannoutsos, F. Grosveld, and P. Fraser. 1996. Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* **87**: 105–114.
- Moon, A.M. and T.J. Ley. 1991. Functional properties of the  $\beta$ -globin locus control region in K562 erythroleukemia cells. *Blood* **77**: 2272–2284.
- Morcillo, P., C. Rosen, and D. Dorsett. 1996. Genes regulating the remote wing margin enhancer in the *Drosophila cut* locus. *Genetics* **144**: 1143–1154.
- Morcillo, P., C. Rosen, M.K. Baylies, and D. Dorsett. 1997. Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in *Drosophila*. *Genes & Dev.* **11**: 2729–2740.
- Moreau, P., R. Hen, B. Wasyluk, R. Everett, M.P. Gaub, and P. Chambon. 1981. The SV40 72 base pair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic Acids Res.* **9**: 6047–6068.
- Muller, H.-P., J.M. Sogo, and W. Schaffner. 1989. An enhancer stimulates transcription in *trans* when attached to the promoter via a protein bridge. *Cell* **58**: 767–777.
- Navas, P.A., K.R. Peterson, Q. Li, E. Skarpidi, A. Rohde, S.E. Shaw, C.H. Clegg, H. Asano, and G. Stamatoyannopoulos. 1998. Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion. *Mol. Cell. Biol.* **18**: 4188–4196.
- Nothias, J.Y., S. Majumder, K.J. Kaneko, and M.L. DePamphilis. 1995. Regulation of gene expression at the beginning of mammalian development. *J. Biol. Chem.* **270**: 22077–22080.
- Ohtsuki, S. and M. Levine. 1998. GAGA mediates the enhancer blocking activity of the *eve* promoter in the *Drosophila* embryo. *Genes & Dev.* **12**: 3325–3330.
- Ohtsuki, S., M. Levine, and H.N. Cai. 1998. Different core promoters possess distinct regulatory activities in the *Drosophila* embryo. *Genes & Dev.* **12**: 547–556.
- Orlando, V., H. Strutt, and R. Paro. 1997. Analysis of chromatin structure by in vivo formaldehyde cross-linking. *Methods* **11**: 205–214.
- Orkin, S.H. 1995. Regulation of globin gene expression in erythroid cells. *Eur. J. Biochem.* **231**: 271–281.
- Panning, B. and R. Jaenisch. 1998. RNA and the epigenetic regulation of X chromosome inactivation. *Cell* **93**: 305–308.
- Peterson, K.R. and G. Stamatoyannopoulos. 1993. Role of gene order in developmental control of human  $\gamma$ - and  $\beta$ -globin gene expression. *Mol. Cell. Biol.* **13**: 4836–4843.
- Peterson, K.R., C.H. Clegg, P.A. Navas, E.J. Norton, T.G. Kimbrough, and G. Stamatoyannopoulos. 1996. Effect of deletion of 5'HS3 or 5'HS2 of the human beta-globin locus control region on the developmental regulation of globin gene expression in beta-globin locus yeast artificial chromosome transgenic mice. *Proc. Natl. Acad. Sci.* **93**: 6605–6609.
- Pikaart, M.J., F. Recillas-Targa, and G. Felsenfeld. 1998. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes & Dev.* **12**: 2852–2862.
- Ptashne, M. and A. Gann. 1997. Transcriptional activation by recruitment. *Nature* **386**: 569–577.
- Raich, N., T. Enver, B. Nakamoto, B. Josephson, T. Papayan-

- nopoulou, and G. Stamatoyannopoulos. 1990. Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science* **250**: 1147–1149.
- Rees, M.I., M. Worwood, P.W. Thompson, C. Gilbertson, and A. May. 1994. Red cell dimorphism in a young man with a constitutional chromosomal translocation t(11;22)(p15.5; q11.21). *Br. J. Haematol.* **87**: 386–395.
- Reik, A., A. Telling, G. Zitnik, D. Cimborra, E. Epner, and M. Groudine. 1998. The locus control region is necessary for gene expression in the human beta-globin locus but not the maintenance of an open chromatin structure in erythroid cells. *Mol. Cell. Biol.* **18**: 5992–6000.
- Rollins, R.A., P. Morcillo, and D. Dorsett. 1999. Nipped-B, a *Drosophila* homologue of chromosomal adherins, participates in activation by remote enhancers in the *cut* and *Ultrabithorax* genes. *Genetics* **152**: 577–593.
- Roseman, R.R., V. Pirrotta, and P.K. Geyer. 1993. The *su(HW)* protein insulates expression of the *Drosophila melanogaster white* gene from chromosomal position-effects. *EMBO J.* **12**: 435–442.
- Scott, K.S. and P.K. Geyer. 1995. Effects of the *su(HW)* insulator protein on the expression of the divergently transcribed *Drosophila* *yolk protein* genes. *EMBO J.* **14**: 6258–6279.
- Sheffery, M., P.A. Marks, and R.A. Rifkind. 1984. Gene expression in murine erythroleukemia cells: Transcriptional control and chromatin structure of the  $\alpha 1$ -globin gene. *J. Mol. Biol.* **172**: 417–436.
- Stalder, J., M. Groudine, J.B. Dodgson, J.D. Engel, and H. Weintraub. 1980. Hb switching in chickens. *Cell* **19**: 973–980.
- Struhl, K. 1989. Molecular mechanisms of transcriptional regulation in yeast. *Annu. Rev. Biochem.* **58**: 1051–1077.
- Sutherland, H.G.E., D.I.K. Martin, and E. Whitelaw. 1997. A globin enhancer acts by increasing the proportion of erythrocytes expressing a linked transgene. *Mol. Cell. Biol.* **17**: 1607–1614.
- Tanimoto, K., Q. Liu, J. Bungert, and J.D. Engel. 1999. Effects of altered gene order or orientation of the locus control region on human  $\beta$ -globin gene expression in mice. *Nature* **398**: 344–348.
- Thanos, D. and T. Maniatis. 1995. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell* **83**: 1091–1100.
- Thoma, F., T. Koller, and A. Klug. 1979. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *J. Cell Biol.* **83**: 403–427.
- Trimborn, T., J. Gribnau, F. Grosveld, and P. Fraser. 1999. Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. *Genes & Dev.* **13**: 112–124.
- Tse, C., T. Sera, A.P. Wolffe, and J.C. Hansen. 1998. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase II. *Mol. Cell. Biol.* **18**: 4629–4638.
- Vyas, P., M.A. Vickers, D.L. Simmons, H. Ayyub, C.F. Craddock, and D.R. Higgs. 1992. Cis-acting sequences regulating expression of the human alpha-globin cluster lie within constitutively open chromatin. *Cell* **69**: 781–793.
- Wadman, I.A., H. Osada, G.G. Grutz, A.D. Agulnick, H. Westphal, A. Forster, and T.H. Rabbitts. 1997. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *EMBO J.* **16**: 3145–3157.
- Wakimoto, B.T. 1998. Beyond the nucleosome: Epigenetic aspects of position-effect variegation in *Drosophila*. *Cell* **93**: 321–324.
- Walter, J. and M.D. Biggin. 1996. DNA binding specificity of two homeodomain proteins in vitro and in *Drosophila* embryos. *Proc. Natl. Acad. Sci.* **93**: 2680–2685.
- Walter, J., C.A. Dever, and M.D. Biggin. 1994. Two homeo domain proteins bind with similar specificity to a wide range of DNA sites in *Drosophila* embryos. *Genes & Dev.* **8**: 1678–1692.
- Walters, M.C., W. Magis, S. Fiering, J. Eidemiller, D. Scalzo, M. Groudine, and D.I.K. Martin. 1996. Transcriptional enhancers act in cis to suppress position-effect variegation. *Genes & Dev.* **10**: 185–195.
- Walters, M.C., S. Fiering, E.E. Bouhassira, D. Scalzo, S. Goeke, W. Magis, D. Garrick, E. Whitelaw, and D.I.K. Martin. 1999. The chicken beta-globin 5'HS4 boundary element blocks enhancer-mediated suppression of silencing. *Mol. Cell. Biol.* **19**: 3714–3726.
- Wijgerde, M., F. Grosveld, and P. Fraser. 1995. Transcription complex stability and chromatin dynamics in vivo. *Nature* **377**: 209–213.
- Weintraub, H. 1988. Formation of stable transcription complexes as assayed by analysis of individual templates. *Proc. Natl. Acad. Sci.* **85**: 5819–5823.
- Weintraub, H. and M. Groudine. 1976. Chromosomal subunits in active genes have an altered conformation. *Science* **193**: 848–856.