IUBMB *Life*, 60(12): 824–833, December 2008

Critical Review

Joining the Loops: β-Globin Gene Regulation

Daan Noordermeer[1] and Wouter de Laat[1,2]
[1]*Department of Cell Biology and Genetics, Erasmus Medical Center, 3000 CA Rotterdam, The Netherlands*
[2]*Hubrecht Institute, Uppsalalaan, 3584 CT Utrecht, The Netherlands*

Summary

The mammalian β-globin locus is a multigene locus containing several globin genes and a number of regulatory elements. During development, the expression of the genes changes in a process called "switching." The most important regulatory element in the locus is the locus control region (LCR) upstream of the globin genes that is essential for high-level expression of these genes. The discovery of the LCR initially raised the question how this element could exert its effect on the downstream globin genes. The question was solved by the finding that the LCR and activate globin genes are in physical contact, forming a chromatin structure named the active chromatin hub (ACH). Here we discuss the significance of ACH formation, provide an overview of the proteins implicated in chromatin looping at the β-globin locus, and evaluate the relationship between nuclear organization and β-globin gene expression. **© 2008 IUBMB**
IUBMB *Life*, 60(12): 824–833, 2008

Keywords β-globin; locus control region; active chromatin hub; chromatin looping; nuclear organization.

Abbreviations ACH, active chromatin hub; CH, chromatin hub; CT, chromosome territory; HS, hypersensitive site; LCR, locus control region; MAR, matrix attachment region; MEL, murine erythroid leukemia; OR, olfactory receptor; WT, wild-type.

THE MAMMALIAN β-GLOBIN LOCUS

The vertebrate hemoglobin gene loci have been intensively studied as model systems for developmentally regulated multigene loci. In mammals, the α- and β-globin loci encode the proteins that form the heteromeric hemoglobin protein-complex involved in oxygen transport. Naturally occurring mutations in the loci show that coregulation of both loci is required, since imbalance between the different proteins may lead to anemia. Interestingly, despite the need for tight coregulation, the mammalian α- and β-locus are structurally very different, and both loci are located in different genomic environments. The α-globin locus is located in a region mainly containing actively expressed housekeeping genes. In contrast, the β-globin locus is embedded in a large region of inactive olfactory receptor (OR) genes. During the later stages of erythroid differentiation, the genes in both loci are expressed at exceptionally high rates. This is necessary to fill the terminally differentiated erythrocyte with hemoglobin. How the β-globin genes achieve their extremely high expression rates despite their location in a repressive chromatin environment has been the subject of intensive investigation.

The human and mouse β-globin loci are the most intensively studied mammalian globin loci. Like all other mammalian β-globin loci, they contain several globin genes, a large upstream regulatory element named the locus control region (LCR) and a number of additional regulatory elements (see Fig. 1A). The genes are positioned on the chromosome in the order of their expression during development (*1, 2*). Their expression is changed in a process called "switching." In the human locus, five expressed genes are present: ε, $^{G}\gamma$, $^{A}\gamma$, δ, and β. The ε-gene is predominantly expressed in primitive erythroid cells in the embryo, both γ-genes are expressed in primitive cells and during the fetal stage in definitive cells, and the β- and δ-genes are first activated in the fetal liver but mainly expressed perinatally (see Fig. 1B). The mouse locus contains four highly expressed genes (εy, βh1, β^{maj} and β^{min}), of which the εy and βh1 genes are expressed during primitive erythropoiesis in embryonic tissues, and the β^{maj} and β^{min} genes are expressed during definitive erythropoiesis in the fetus and adult mouse (see Fig. 1B). Intriguingly, in embryonic tissue, the βh1 gene initially appears more expressed than εy, suggesting that at defined stages of development gene expression may not strictly correlate with their order on the chromosome (*3*). The LCR is the main regulatory element in the β-globin locus and is required for high levels of expression of all the genes. It was identified in patients that were suffering from anemia (γδβ-thalassemia) because of

Received 9 July 2008; accepted 30 July 2008
Address correspondence to: Wouter de Laat, Hubrecht Institute, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands. Tel: +31 30-2121800; Fax: +31 30-2516464. E-mail: w.laat@niob.knaw.nl

ISSN 1521-6543 print/ISSN 1521-6551 online
DOI: 10.1002/iub.129

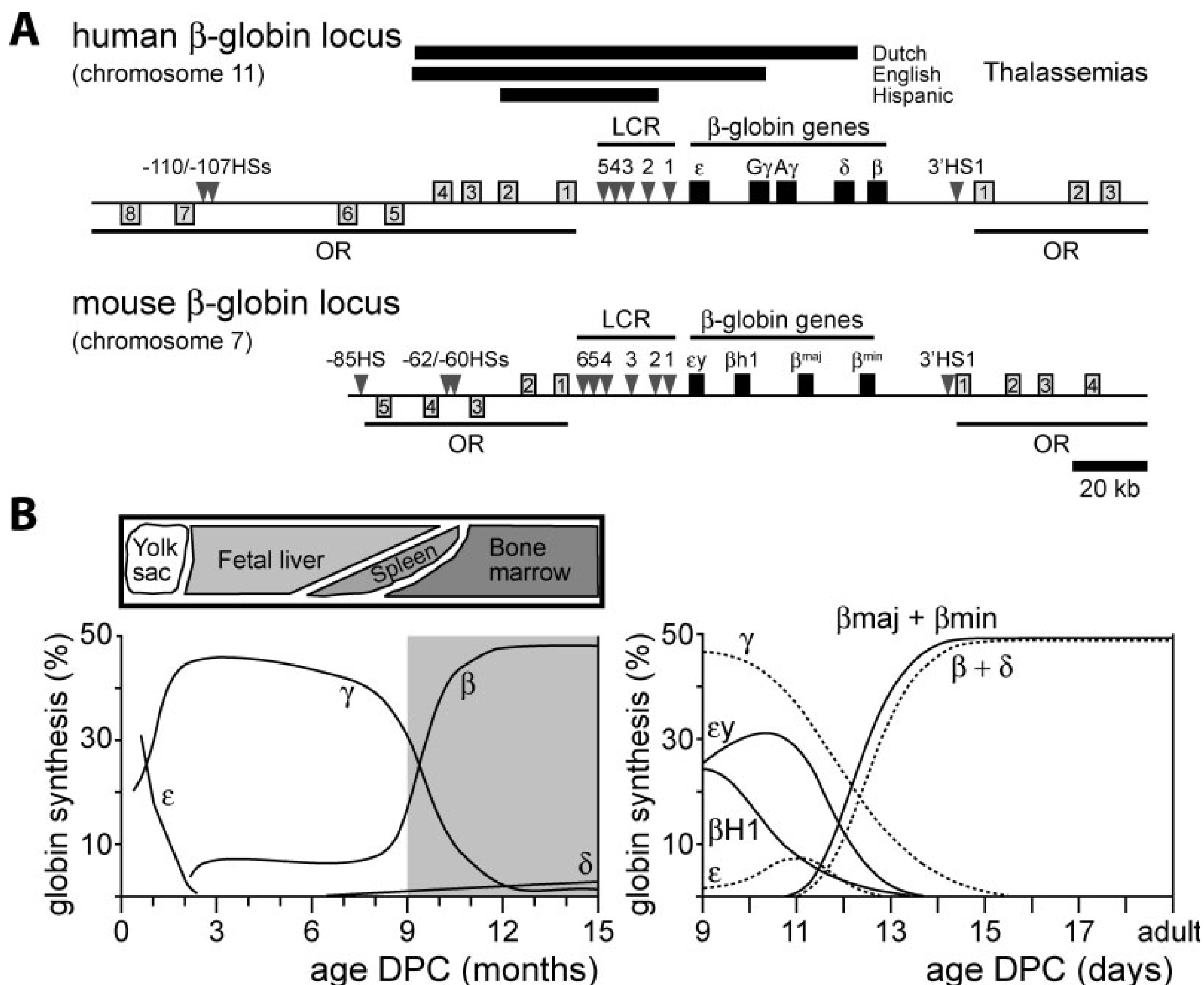

Figure 1. (A) The human and mouse β-globin locus. β-like genes are indicated by black boxes and hypersensitive sites by arrowheads. The genes in the surrounding olfactory receptor clusters (OR) are indicated by gray boxes and numbered according to their distance from the β-globin locus. Naturally occurring $\gamma\delta\beta$-thalassemia deletions are indicated above the human locus. (B) β-like globin protein content during development in human (left) and mouse (right) as the percentage of the total globin content in the cells. Above the graph for the human locus the globin producing tissues at the stage of development is indicated. In the graph for the mouse locus, both the globin content of the mouse globins (straight line) and human transgenically expressed globins (dashed line) are shown.

β-chain imbalance, even though they were carrying a normal β-globin gene (*4, 5*). Over the years, several other large deletions were identified that all comprised a region upstream of the ε-gene (see Fig. 1A). In this region, a number of sites showed strong DNase I-hypersensitivity in erythroid cells (*6*). The experiments by Grosveld et al. (*7*) in which globin transgenes coupled to the LCR were found to be highly expressed (independent of the site of integration) showed the absolute importance of the LCR for high levels of globin gene expression, and lead to a general definition of LCRs: elements that confer copy number dependent but position-of integration independent expression to transgenes. In agreement with this, deletion of the mouse LCR leads to a 25- to 100-fold reduction of globin gene expression (*8*). The number of HSs that comprise the LCR varies between different species and is not always agreed on, but in both human and mouse, the erythroid specific HS1–HS4 exhibit the strongest enhancer function. The constitutive HS5 has a more structural role but has little effect on expression. Outside the LCR, both upstream and downstream of the globin genes, a number of additional HSs are found that seem to also have a structural role, comparable to HS5 of the LCR (see Fig. 1A).

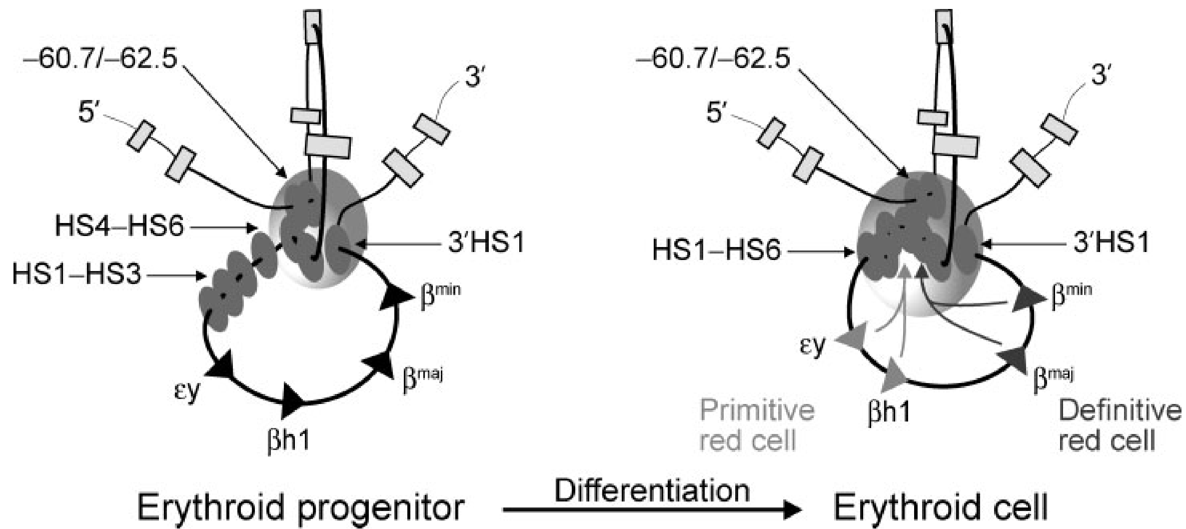

Figure 2. Structure of the mouse β-globin chromatin hub (CH, left) in erythroid progenitor cells and the active chromatin hub (ACH, right) in β-globin expressing erythroid cells. Inactive β-globin genes are depicted by black arrowheads. Embryonic genes, participating in the ACH in embryonic tissues, are depicted by light gray arrowheads. Fetal and adult genes, participating in the ACH in fetal and adult tissues, are depicted by dark gray arrowheads. Hypersensitive sites are shown as gray ovals, and olfactory receptor genes are depicted by gray boxes.

UPREGULATION OF THE β-GLOBIN GENES BY THE LCR

The discovery of enhancers and LCRs raised the question how a distant element could have such a large effect on the expression of its target genes. Many studies have been aimed at elucidating the mechanisms responsible for this long-range gene activation. Initially, a number of important discoveries were made that described the pattern of switching and demonstrated the relevance of gene order of the genes in the β-globin locus. In humans, the LCR upregulates only one gene at the time, and the genes compete with each other for activation by the LCR (*9*). In mice, the LCR seems to employ a similar mechanism, although nascent RNA signals from either εy and βh1 or β^{maj} and β^{min} are frequently detected simultaneously at one allele. A possible explanation for this could be LCR-mediated coregulation of the genes at the same allele, but after careful experiments, the authors conclude that the most likely explanation is a rapid switching of the LCR between the genes on the same allele (*10*). The main determinant for activation of a gene is the relative distance from the LCR. Introduction of a β-globin gene between the LCR and γ-gene leads to premature activation of the β-gene and a strong reduction in expression of the γ-gene (*11*). Introduction of a second β-gene in the locus causes the more proximal gene to be highest expressed, but the total output of the two β-genes together is not increased (*2*), similar to what has been found in patients that actively express the γ-globin genes in the adult stage (*12*). In human adult tissues, the switch leading to the expression of the further downstream located β- and δ-genes is achieved by active silencing of the embryonic and fetal genes (*13–15*).

The observed competition between the β-like genes for upregulation by the LCR led to speculations that the LCR functioned by physically interacting with the genes. The simultaneous development of two new techniques, Chromosome Conformation Capture (3C) and RNA TRAP, allowed the determination of the spatial organization of the β-globin locus and showed that indeed the LCR and activated globin genes are in spatial proximity (*16, 17*). The 3C technique proved very useful to study the changes during development in the β-globin locus. Using 3C, the capture of *in vivo* interacting DNA fragments via formaldehyde crosslinking and subsequent ligation can be quantified by qPCR across ligation junctions. As a result, this gives a measure for interaction frequencies of the fragments of interest (*18*). If 3C experiments are well controlled, they demonstrate the presence of chromatin loops (*18–21*). It is important to realize though that chromatin conformations as appreciated by 3C reflect steady-state averages measured across the population of cells. It may therefore well be that the structure of the β-globin locus at any given time is different from cell to cell (*16*).

In globin-expressing cells, depending on the tissue, the LCR colocalizes with the active genes, thereby forming a structure named the active chromatin hub (ACH, see Fig. 2 for the mouse locus) (*22*). In progenitor cells a substructure, the chromatin hub (CH), is present, consisting of the constitutive HS5 from the LCR and the upstream and downstream outer hypersensitive sites (*22*). The appearance and disappearance of DNA interactions in the β-globin locus strikingly correlates with β-globin gene expression levels. Additional genetic experiments should

reveal whether these loops are functional or not (*21*), but this does strongly suggest that ACH formation is important for the high-expression levels of the β-globin genes.

Important questions now are how these chromatin loops are established and changed in the switching process, how frequent and dynamic these interactions are, and which factors are involved. Furthermore, it is interesting to consider whether the specific chromatin conformation of the β-globin locus may change the nuclear environment, thereby potentially facilitating the high expression of the genes. In the remainder of this review, we will summarize the key players involved in chromatin looping at the β-globin locus and we will discuss the potential role of nuclear organization on the expression of the globin genes.

FACTORS INVOLVED IN CHROMATIN LOOPING AT THE β-GLOBIN LOCUS

Many proteins, both ubiquitously expressed and erythroid-specific, are known to be involved in β-globin gene regulation and chromatin looping. Most of these factors bind the promoters of the globin genes and/or the HSs of the LCR and often are present in protein-complexes at these sites. Here we will discuss the current knowledge of factors involved in chromatin looping at the β-globin locus.

GATA-1 and FOG-1

The most intensively characterized regulator of globin gene expression is the erythroid-specific GATA-1 transcription factor (previously also referred to as Gf-1 and NF-E1). GATA-1 can both activate and repress globin genes, but the study of this factor is complicated because of its involvement in activation and repression of many transcription factors involved in erythroid differentiation. In the mouse β-globin locus, GATA-1 binds the promoter of the β^{maj}-gene and HS1-HS4 of the LCR (*23–25*). Binding of GATA-1 to the promoter and HS2, but not to HS3 and HS4, depends on the GATA-1 interacting partner FOG-1 (*23, 24*). The interaction with FOG-1 is also a determining factor for the involvement of GATA-1 in different protein-complexes that can have activating or repressing functions during development (*26*). Therefore, FOG-1 interaction with GATA-1 may mediate recruitment of functionally different GATA-1 complexes with different regulatory functions to specific sites in the locus. Indeed site-specific epigenetic changes are observed in GATA-1 knockout cells expressing a GATA-1 mutant that is disrupted in its FOG-1 interaction. Acetylation of amino-acid residues in the tail of histone H3 are decreased at the β^{maj} promoter, where FOG-1 is required for GATA-1 binding. In contrast, no changes are observed at HS3 of the LCR, where GATA-1 binding is FOG-1-independent (*23*).

Next to targeting of regulatory protein-complexes, GATA-1 and FOG-1 binding is essential for loop formation between the mouse β^{maj} promoter and the LCR. 3C technology was applied to a cell line containing a GATA-1-fusion protein that is localized in the cytoplasm, but upon stimulation is rapidly recruited to the nucleus (*27*). When GATA-1 is located in the cytoplasm, looping between the promoter of β^{maj} and HS2 and HS3 is nearly absent. Relocation of GATA-1 to the nucleus strongly increases looping between the promoter and the LCR. The direct involvement of GATA-1 in looping is further supported by the fact that loops are still being formed in the absence of protein synthesis, indicating that activation or repression of other genes by GATA-1 is not required for chromatin looping. Disruption of FOG-1 binding to the GATA-1-fusion protein strongly decreases chromatin looping between HS2 and the gene, which is accompanied by a severe decrease in β^{maj} expression (*27*).

Interestingly, GATA-1 has a role in gene-repression involving chromatin looping at the erythroid-specific *Kit* locus. GATA-1 was found to be involved in chromatin loop switching, leading to silencing of the *Kit* locus in the cells previously used to examine GATA-1 inducible looping at the globin locus (*27, 28*). When the *Kit* locus is active, a number of GATA-sites in the locus are occupied by GATA-2, a factor closely related to GATA-1. By applying 3C, a loop is observed between the 5′-end of the *Kit*-gene and a distant upstream enhancer. When GATA-1 is relocated into the nucleus, GATA-2 is replaced by GATA-1 at the *Kit* locus and *Kit* transcription is repressed. The replacement of GATA-2 by GATA-1 could be due to direct competition between the two factors, but more likely is caused by GATA-1-induced repression of the *GATA-2* locus. Switching of GATA factors in the locus causes a change in chromatin conformation of the locus. The initial loop between the start of the *Kit* gene and enhancer is abolished, and instead, a new loop is formed between the 5′-end of the *Kit* gene and a region far more downstream in the gene (*28*). Whether GATA-1-induced chromatin loop switches occur in the β-globin locus needs to be established, but it is an intriguing thought that GATA-1 may be a driving factor for globin switching. An interesting observation in this respect is the recently reported GATA-1 mediated silencing of the γ-genes in the human β-globin locus. GATA-1 binds a region upstream of both the Aγ- and Gγ-promoter in a FOG-1 dependent manner, leading to recruitment of a component of the repressive NuRD-complex (*29*). Determination of whether binding of GATA-1 at these sites influences the structure of the ACH during development could potentially give further insight in how switching at the β-globin locus is achieved.

EKLF

A second well-characterized erythroid-specific factor involved in β-globin gene regulation and loop formation is the Krüppel-like zinc finger DNA-binding protein ELKF (encoded by the *KLF1* gene). EKLF is not required for the expression of the embryonal ε- and fetal γ-genes, but strongly affects the expression of the adult β-globin gene (*30*). In homozygous EKLF knockout mice containing a transgenic human β-globin locus, β-gene expression is absent and γ-globin expression is increased. In heterozygous EKLF mice, β-gene expression is

delayed during differentiation, accompanied by increased γ-gene expression, but β-globin levels are unaffected in adult mice (*30*). The exact mechanism of EKLF action in globin gene regulation is unclear because of several contradictory reports, most likely originating from different experimental designs. In mice, EKLF activates a human β-globin reporter construct containing HS3 of the LCR and is required for formation of this HS in the construct (*31*). Similarly, HS3 is not formed at the full transgenic human locus in EKLF knockout mice (*30*). Different results are obtained from the analysis of minichromosomes containing the β-globin gene coupled to HS2 and HS3 in human *K562* genes. Here upregulation by HS3 is not dependent on the presence of EKLF, but rather HS2 function is EKLF-dependent (*32*). Interestingly, upregulation by HS3 in this study is dependent on the structural integrity of the EKLF binding sites at HS2. A construct containing HS3 coupled to HS2 without EKLF binding sites totally ablates HS3 enhancer function, suggesting an interaction between the HSs that is dependent on EKLF binding (*32*).

Application of 3C technology on the mouse β-globin locus in EKLF knockout mice revealed that interactions between the LCR and the β^{maj}-gene are lost when EKLF is not present (*33*). To examine whether EKLF has a direct function in looping, 3C experiments were performed on EKLF knockout mice containing an EKLF-fusion protein that can be rapidly relocated to the nucleus. Recruitment of this EKLF-fusion protein to the nucleus in cells with blocked protein synthesis showed reformation of the ACH and partially restored β^{maj} expression, indicating that EKLF is directly involved in chromatin looping between the LCR and the β^{maj} promoter (*33*).

Together, these results indicate that EKLF is an important factor in the switching process. EKLF allows the β-globin gene to compete efficiently with the γ-globin genes, possibly by initiating loops between the β-globin gene promoter and HS2 and HS3 of the LCR. The decreased levels of EKLF in heterozygous mice reduce the competing ability of the β-globin promoter, leading to upregulation of the γ-globin genes. Later during development, the active silencing of the γ-globin genes relieves the β-globin gene from competition, allowing efficient loop formation despite the decreased EKLF levels, resulting in normal expression of the β-globin gene.

NF-E2

A third intensively studied factor involved in regulation of the globin genes is the NF-E2 heterodimer. The complex consists of two DNA binding subunits; the ubiquitously present MafK unit or p18 NF-E2 and its erythroid-specific partner p45 NF-E2. NF-E2 most prominently binds HS2 of the LCR in mouse erythroid fetal liver tissue and in several mouse and human erythroleukemia cell lines, although a lower level of binding is also observed at the promoter of the adult β-genes. Depending on the cell system used, the necessity for NF-E2 in β-globin gene expression greatly varies (*34*). In cell lines, p45 NF-E2 is required for β-globin gene expression, but in p45 NF-E2 knockout mice, β^{maj} expression levels are only slightly reduced (*34, 35* and references therein). An explanation for this difference could be that in mice, in the absence of p45 NF-E2, redundant factors may associate with MafK, thereby compensating for the absence of this erythroid specific factor. Evidence for this hypothesis may be found in two recent reports showing very different requirements for the two components that make up NF-E2. In the first study, ACH formation was analyzed in p45 NF-E2 knockout mice (*35*). In these mice, expression of the β^{maj} gene is slightly reduced [to a level similar as observed upon deletion of HS2 (*36*)], but the overall folding of the β-globin locus is not affected. Without p45 NF-E2, binding of the other NF-E2 subunit, MafK, is almost absent at the promoter and reduced to about 50% at HS2. MafK now seems associated with p45 NF-E2-related proteins like Nrf2 and to a lesser extent the repressing factor Bach1, which show strongly increased binding at HS2 (*35*). These results are in contrast to a study on an established mouse erythroid leukemia cell line, DS19, in which MafK was knocked down using an siRNA approach (*37*). MafK appears almost completely absent at the protein level, though binding of both MafK and p45 NF-E2 at HS2 and the promoter are only reduced to about one-third of WT levels. Nevertheless, β^{maj} expression is about 80% lower, and this is reflected by a reduction of looping between HS2 and the β^{maj} promoter to levels close to nonglobin expressing precursor cells. In contrast to the p45 NF-E2 knockout, in these cells no increase in Nrf2 binding is observed when MafK is knocked down (*37*). The latter study may be interpreted to suggest that NF-E2 is involved in loop formation and that *in vivo* function of p45 NF-E2 may be compensated by other dimerization partners like Nrf2. However, it should be noted that mice lacking both p45 NF-E2 and Nrf2, or p45 NF-E2 and Nrf3 show no erythroid phenotype beyond that seen with deletion of p45 NF-E2 alone (*38–40*), implying that compound knockout mice lacking all NF-E2-related factors need to be analyzed to unambiguously address this issue.

Other Transcriptional Regulators Involved in Looping

Recently, in addition to the previously discussed well-characterized transcription factors, a number of other proteins have been reported to have a potential role in chromatin looping at the β-globin locus. LDB1, a ubiquitous non-DNA binding protein, participates in many protein-complexes, which can both activate and repress the expression of target genes. Interestingly, the protein interacts with GATA-1 in murine erythroid cells (*26, 41*). In a recent study, both LDB1 and GATA-1 were reported to be present at the promoter of the β^{maj} gene and HS1–HS4 of the LCR in mouse fetal liver and murine erythroid leukemia (MEL) cells and at HS1–HS4 in human K562 cells (*25*). Truncation of LDB1 reduces the expression of the ε-gene on a minichromosome containing HS2 in K562 cells, while a knockdown of LDB1 leads to decreased β^{maj} expression in differentiated MEL cells. In these knockdown cells, interaction

between HS2 and the promoter of β^{maj} are reduced to background levels, showing that LDB1 is either directly or indirectly involved in long range chromatin interactions at the β-globin locus (*25*). Since LDB1 is a non-DNA-binding protein, but interacts with many factors known to be involved in the regulation of erythroid genes, it may function by attaching several transcription factors to each other, thereby structuring the ACH. It would be interesting to know whether depletion of LDB1 also abrogates binding of its complexing partners GATA-1 and MafK to the β-globin locus.

Two other proteins have recently been reported to have a potential role in the switching process. SATB1, a protein known to bind specific AT-rich sequences in the genome, has been proposed to serve as spatial "genomic organizer" in certain cell types. In K562 cells, the protein binds MARs in the promoter of the ε-globin gene and HS2 of the LCR. Overexpression of SATB1 increases ε-gene expression, while simultaneously downregulating γ-globin (*35*). Silencing of the ε-globin gene and activation of the γ-globin gene coincides with SATB1 repression, suggesting a potential relationship (*42*). In T-helper cells, SATB1 is involved in the formation of an intricate chromatin conformation at the T_H2 cytokine locus (*43*). This raises the question whether SATB1 may change chromatin loops during the switching process.

The Ikaros protein is also thought to be involved in switching. Expression of a dominant negative version of Ikaros in primary human cells decreases the expression of both the fetal and adult globin genes, but more interestingly, strongly alters the ratio between γ- and β-gene expression in favor of the γ-globin genes (*44*). In transgenic mice containing the human β-globin locus and expressing a DNA-binding impaired version of Ikaros, switching from γ- to β-gene expression is severely delayed, but β-globin gene expression is ultimately fully restored. Ikaros binds HS3 and to a lesser extent regions in the Gγ-gene and upstream of the δ-gene. In the mice containing the DNA-binding mutant of Ikaros, chromatin looping between a large fragment in the LCR and the β-gene is disturbed, and a new loop with the γ-globin genes is observed (*45*). Therefore, like EKLF, Ikaros may have a role in switching from fetal to adult globin gene expression. Whether the binding of Ikaros at both the Gγ-gene and the δ-gene may be a key event in directing the switch in chromatin looping will be an interesting topic for future research.

CTCF and Cohesin

CCCTC-binding factor (CTCF) and the cohesin complex are present at the β-globin locus and implied in chromatin looping, but do not seem to be involved in β-globin transcriptional regulation. CTCF is the most intensively characterized mammalian insulator protein. It blocks the effect of an enhancer on a target gene by binding in between (*46, 47*). Enhancer blocking is proposed to be achieved by locating the gene and enhancer on different chromatin loops, which would preclude them from interacting with each other. CTCF binds HS5 of the LCR and the downstream 3′HS1 in the human and mouse β-globin loci. Additional binding sites have been reported at the upstream –62/−60 and –85 HSs in the mouse locus (see Fig. 1A). All these HSs, except the –62/−60 sites and possibly the –85 site, can act as enhancer blockers (*48, 49*). In nonglobin expressing erythroid precursor cells, the CTCF-binding sites are in spatial proximity, thereby forming a chromatin structure called the CH [(*22*) and Fig. 2]. In primary mouse CTCF knockout cells, association between these sites is severely decreased, showing that CTCF is indeed involved in the shaping of the CH (*50*). However, it is unclear what the function of these enhancer blocker sites in the β-globin locus is. It has been hypothesized that these sites either shield the LCR from activating the surrounding inactive OR-genes or that they block a potential signal from outside the locus to act on the globin genes. Deletion of the CTCF-binding site at 3′HS1 disturbs the participation of this site in the ACH. Surprisingly, looping between the LCR and the promoter of the β^{maj} gene at later stages of erythroid differentiation is not affected and expression of the β^{maj} gene is not changed. The inactivity of the further downstream located OR-genes is also not altered (*50*). Different studies supporting this finding showed that deletion of 3′HS1 and the upstream –62 HS does not affect εy and β^{maj} expression (*51*), that a human β-globin gene is upregulated when located upstream of the human LCR in transgenic mice (*52*) and that an integrated human β-globin LCR in an ectopic locus in mice upregulates genes upstream of HS5 from the LCR despite binding of CTCF (*53*). Together, these results indicate that in erythroid progenitor cells CTCF has an important role in structuring the CH, but that the transition to the ACH and activation of the globin genes is not dependent on the presence of CTCF. Furthermore, CTCF-binding is not required to prevent deregulation of genes in and directly outside the β-globin locus, arguing against a role for CTCF in insulating the locus from its surrounding. The functional relevance of β-globin chromatin loop formation by CTCF early during erythropoiesis therefore remains enigmatic.

Recently, a strong correlation between binding of CTCF and localization of cohesin to genomic loci was reported (*54, 55*). During replication, the cohesin complex keeps the replicated sister chromatids together by forming a ring structure around the two strands. A role for cohesin in gene regulation during interphase has also been established, but until recently was not well understood. The colocalization with CTCF makes it attractive to speculate that CTCF recruits cohesin, and that cohesin in turn stabilizes interactions between CTCF binding sites, possibly by formation of rings around the chromatin strands. In HeLa cells, cohesin binds transgenic chicken HS4, the HS that is functionally related to HS5 in the mammalian β-globin locus. Depletion of both CTCF or cohesin in these cells results in impaired insulator function of this HS, indicating that cohesin may play a role together with CTCF in CH formation (*54*).

THE ACH AND ITS NUCLEAR ENVIRONMENT

Gene expression and nuclear organization are tightly correlated mechanisms. Changes in the transcription state of a locus

often coincide with changes in localization versus nuclear substructures. Several studies have reported nuclear relocation of the β-globin locus during erythroid differentiation, when the expression status of the locus changes from inactive or moderately active, to very highly active (*56–58*). At the level of the β-globin locus, this process coincides with the transition from CH to ACH (*22*). An interesting question is how these changes relate to each other.

When inactive, the β-globin locus is mainly located at the periphery of the nucleus and associates with repressive centromeric heterochromatin. When the locus is activated, it relocates to a position more to the interior of the nucleus and away from centromeric heterochromatin (*58, 59*). Activation of the β-globin locus precedes the relocation to the nuclear interior, indicating that relocation is not a consequence of transcription at the locus. Rather, relocation may have a function in maintenance of transcriptional activity and may be driven by the association of the locus with nuclear entities known as transcription factories (*58*). Deletion of the LCR strongly impairs this association with transcription factories and the relocation away from the nuclear periphery, but had previously been reported to have no effect on movement away from centromeric heterochromatin (*58, 59*).

A recent topic of interest is the location of the β-globin locus versus other genomic loci in the nucleus. In a number of studies the location of loci in the chromosome territory (CT) has been determined. Depending on the cell system used for examination, different results are obtained. In mouse MEL cells, where the globin genes are poised for transcription, the locus was reported to be significantly more often located outside the CT compared to non erythroid cells (*56*). Furthermore, in this study MEL cells containing human chromosome 11 with or without the β-globin LCR were used to show that location of the β-globin locus outside the CT is strongly dependent on the presence of the LCR (*56*). In contrast, another study showed that in mouse anemic spleen cells and human erythroblasts sorted at various differentiation stages the β-globin locus is almost exclusively located inside the CT (*57*). To further explore the capacity of the LCR to relocate surrounding chromatin, we recently investigated an ectopically integrated human β-globin LCR integrated in a gene-dense region in the mouse genome. Localization outside the CT is strongly increased when the LCR is integrated in this region. The LCR upregulates many genes in the gene-dense region, most likely by chromatin looping, but not all genes respond to the integration of the LCR and upregulation is dependent on the orientation of this element (*53*). Together, these results show that looping out of the CT is not required for increased transcription, and that looping out does not automatically lead to increased transcription. It could be that the LCR locates the globin locus outside the CT to induce a poised state of the locus, as was suggested in the study of Ragoczy et al. (*56*). Whether this relocation is related to the spatial conformation of the locus will be an interesting topic for future studies.

The location of the β-globin locus versus other genomic loci has been characterized in two recent studies. In the first study, the location of the mouse β-globin locus was determined versus a selected number of other highly expressed loci on the same chromosome. These highly expressed genes are often in close spatial proximity when active. Colocalization was proposed to be at preassembled RNA polymerase II foci in the nucleus, where genes would need to migrate to for activation (*60*). As two out the four loci that interacted with the β-globin locus also carried erythroid-specific genes, the data may be interpreted to suggest that functionally related genes preferentially come together in the nuclear space. In another study, genome-wide interacting partners of the β-globin locus were identified by using 4C, an unbiased microarray-based method that allows screening the entire genome for regions that contact a locus of choice (*61*). It was found that the active β-globin locus colocalizes with many regions containing active genes on the same chromosome, including those found in the study by Osborne et al. [see reference (*60*)]. In the context of all interacting loci, no preference was found for the β-globin locus to contact other erythroid-specific genes (*61*). Interestingly, when 4C was applied to the inactive β-globin locus, it was found to associate with very different regions on the same chromosome, which mainly contained inactive genes (*61*). This suggests that the overall transcriptional status determines which chromosomal regions come together in the nucleus, with active chromatin separating from inactive chromatins. Whether clustering of the active β-globin locus with other active regions is only determined by transcription, or whether the LCR has a role in it is not known. Both studies have been repeated in the absence of transcription, and both studies reported no differences when transcription was blocked (*62, 63*). Therefore, it can be concluded that transcription itself is not necessary for the maintenance of these long-range associations, but this does not exclude that transcription is necessary for the initiation of these interactions.

Together, these data show that in the process of development, and possibly also during erythroid differentiation, the β-globin locus is relocated to a very different nuclear environment. To understand the functional significance of the interactions formed by the active β-globin locus, it would be interesting to apply 4C technology to the locus with and without the LCR. Likewise, it would be interesting to apply 4C technology to the experimental system containing the ectopically integrated LCR and investigate the DNA contacts formed with and without the LCR. Question is: does the LCR search for erythroid-specific or functionally related genes elsewhere in the genome? If this were true, it would strongly support the idea of dedicated transcription factories exclusively transcribing subsets of functionally related genes.

THE ACTIVE CHROMATIN HUB, FROM STEADY STATE TO DYNAMICS

Since the elucidation of the structure of the ACH, considerable effort has been dedicated to determine the factors involved

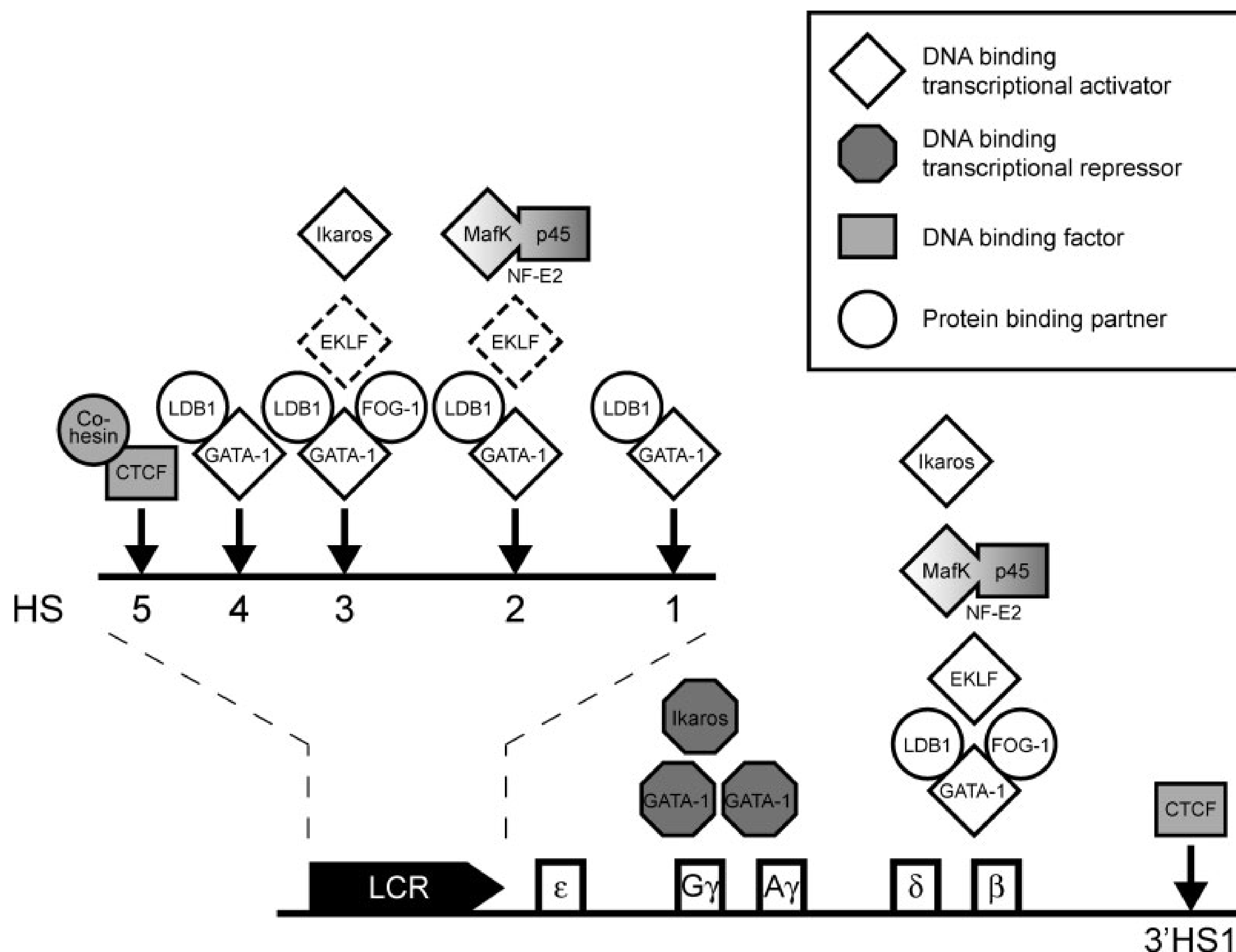

Figure 3. Binding of factors known or proposed to be involved in chromatin looping to different sites in the β-globin locus. Binding sites for factors from studies on the adult human and mouse locus have been combined in this figure. Globin genes are depicted by boxes, and hypersensitive sites are indicated by arrows. Reports on binding of EKLF to sites in the LCR are contradictory and therefore the factor is shown with a dashed line.

in structuring this 3D chromatin structure. In this review, we have tried summarizing the current factors known or proposed to be involved in long-range chromatin interactions at the β-globin locus and how the structure of the locus relates to its nuclear environment. In Fig. 3 a summary of factors present at sites in the β-globin locus is shown. Even though a considerable number of factors have already been shown to be involved in looping at the globin locus, we have little doubt many more proteins play a role in chromatin looping at the locus. Probably the biggest challenge is to move from the current steady state studies that are based on cell populations to single cell studies and the visualization of changes in the locus during differentiation and the switching process. Indeed, it is to be expected that the ACH describes a structure that in each locus is dynamically formed and destabilized, with different HSs making and breaking contacts over time. At the molecular level, it is important to investigate how these interactions affect the process of transcription initiation, reinitiation and possibly also elongation. At the level of nuclear organization the main challenge will be to determine where exactly the activated locus moves to and whether such movement is a prerequisite for proper expression of the β-globin genes or a mere consequence of gene activation. For this, the nuclear environment of the β-globin locus needs to be characterized at different stages of differentiation, both in terms of protein content and in terms of the DNA loci present. 4C technology, in combination with immuno-FISH experiments, will help addressing these issues.

ACKNOWLEDGEMENTS

We thank Robert-Jan Palstra for sharing pictures and Frank Grosveld and members of the lab for useful discussions. Wouter

de Laat is funded by grants 016-006-026 and 912-04-082 from the Dutch Scientific Organization (NWO).

REFERENCES

1. Grosveld, F., Dillon, N., and Higgs, D. (1993) The regulation of human globin gene expression. *Baillieres. Clin. Haematol.* **6**, 31–55.
2. Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P., and Grosveld, F. (1997) The effect of distance on long-range chromatin interactions. *Mol. Cell.* **1**, 131–139.
3. Kingsley, P. D., Malik, J., Emerson, R. L., Bushnell, T. P., McGrath, K. E., Bloedorn, L. A., Bulger, M., and Palis, J. (2006) "Maturational" globin switching in primary primitive erythroid cells. *Blood* **107**, 1665–1672.
4. Van der Ploeg, L. H., Konings, A., Oort, M., Roos, D., Bernini, L., and Flavell, R. A. (1980) gamma-beta-Thalassaemia studies showing that deletion of the gamma- and delta-genes influences beta-globin gene expression in man. *Nature* **283**, 637–642.
5. Kioussis, D., Vanin, E., deLange, T., Flavell, R. A., and Grosveld, F. G. (1983) Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. *Nature.* **306**, 662–666.
6. Tuan, D., Solomon, W., Li, Q., and London, I. M. (1985) The "beta-like-globin" gene domain in human erythroid cells. *Proc. Natl. Acad. Sci. U S A* **82**, 6384–6388.
7. Grosveld, F., Blom van Assendelft, G., Greaves, D. R., and Kollias, G. (1987) Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell.* **51**, 975–985.
8. Epner, E., Reik, A., Cimbora, D., Telling, A., Bender, M. A., Fiering, S., Enver, T., Martin, D. I., Kennedy, M., Keller, G., and Groudine, M. (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.
9. Wijgerde, M., Grosveld, F., and Fraser, P. (1995) Transcription complex stability and chromatin dynamics in vivo. *Nature* **377**, 209–213.
10. Trimborn, T., Gribnau, J., Grosveld, F., and Fraser, P. (1999) Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. *Genes Dev.* **13**, 112–124.
11. Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N., and Grosveld, F. (1991) Importance of globin gene order for correct developmental expression. *Genes Dev.* **5**, 1387–1394.
12. Giglioni, B., Casini, C., Mantovani, R., Merli, S., Comi, P., Ottolenghi, S., Saglio, G., Camaschella, C., and Mazza, U. (1984) A molecular study of a family with Greek hereditary persistence of fetal hemoglobin and beta-thalassemia. *EMBO J.* **3**, 2641–2645.
13. Raich, N., Enver, T., Nakamoto, B., Josephson, B., Papayannopoulou, T., and Stamatoyannopoulos, G. (1990) Autonomous developmental control of human embryonic globin gene switching in transgenic mice. *Science* **250**, 1147–1149.
14. Dillon, N., and Grosveld, F. (1991) Human gamma-globin genes silenced independently of other genes in the beta-globin locus. *Nature* **350**, 252–254.
15. Sabatino, D. E., Cline, A. P., Gallagher, P. G., Garrett, L. J., Stamatoyannopoulos, G., Forget, B. G., and Bodine, D. M. (1998) Substitution of the human beta-spectrin promoter for the human agamma-globin promoter prevents silencing of a linked human beta-globin gene in transgenic mice. *Mol. Cell. Biol.* **18**, 6634–6640.
16. Tolhuis, B., Palstra, R. J., Splinter, E., Grosveld, F., and de Laat, W. (2002) Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol. Cell.* **10**, 1453–1465.
17. Carter, D., Chakalova, L., Osborne, C. S., Dai, Y. F., and Fraser, P. (2002) Long-range chromatin regulatory interactions in vivo. *Nat. Genet.* **32**, 623–626.
18. Dekker, J., Rippe, K., Dekker, M., and Kleckner, N. (2002) Capturing chromosome conformation. *Science* **295**, 1306–1311.
19. Dekker, J. (2003) A closer look at long-range chromosomal interactions. *Trends Biochem. Sci.* **28**, 277–280.
20. Splinter, E., Grosveld, F., and de Laat, W. (2004) 3C technology: analyzing the spatial organization of genomic loci in vivo. *Methods Enzymol.* **375**, 493–507.
21. Simonis, M., Kooren, J., and de Laat, W. (2007) An evaluation of 3C-based methods to capture DNA interactions. *Nat. Methods* **4**, 895–901.
22. Palstra, R. J., Tolhuis, B., Splinter, E., Nijmeijer, R., Grosveld, F., and de Laat, W. (2003) The beta-globin nuclear compartment in development and erythroid differentiation. *Nat. Genet.* **35**, 190–194.
23. Letting, D. L., Chen, Y. Y., Rakowski, C., Reedy, S., and Blobel, G. A. (2004) Context-dependent regulation of GATA-1 by friend of GATA-1. *Proc. Natl. Acad. Sci. USA* **101**, 476–481.
24. Im, H., Grass, J. A., Johnson, K. D., Kim, S. I., Boyer, M. E., Imbalzano, A. N., Bieker, J. J., and Bresnick, E. H. (2005) Chromatin domain activation via GATA-1 utilization of a small subset of dispersed GATA motifs within a broad chromosomal region. *Proc. Natl. Acad. Sci. USA* **102**, 17065–17070.
25. Song, S. H., Hou, C., and Dean, A. (2007) A positive role for NLI/Ldb1 in long-range beta-globin locus control region function. *Mol. Cell.* **28**, 810–822.
26. Rodriguez, P., Bonte, E., Krijgsveld, J., Kolodziej, K. E., Guyot, B., Heck, A. J., Vyas, P., de Boer, E., Grosveld, F., and Strouboulis, J. (2005) GATA-1 forms distinct activating and repressive complexes in erythroid cells. *EMBO J.* **24**, 2354–2366.
27. Vakoc, C. R., Letting, D. L., Gheldof, N., Sawado, T., Bender, M. A., Groudine, M., Weiss, M. J., Dekker, J., and Blobel, G. A. (2005) Proximity among distant regulatory elements at the beta-globin locus requires GATA-1 and FOG-1. *Mol. Cell.* **17**, 453–462.
28. Jing, H., Vakoc, C. R., Ying, L., Mandat, S., Wang, H., Zheng, X., and Blobel, G. A. (2008) Exchange of GATA factors mediates transitions in looped chromatin organization at a developmentally regulated gene locus. *Mol. Cell.* **29**, 232–242.
29. Harju-Baker, S., Costa, F. C., Fedosyuk, H., Neades, R., and Peterson, K. R. (2008) Silencing of Agamma-globin gene expression during adult definitive erythropoiesis mediated by GATA-1-FOG-1-Mi2 complex binding at the −566 GATA site. *Mol. Cell. Biol.* **28**, 3101–3113.
30. Wijgerde, M., Gribnau, J., Trimborn, T., Nuez, B., Philipsen, S., Grosveld, F., and Fraser, P. (1996) The role of EKLF in human beta-globin gene competition. *Genes Dev.* **10**, 2894–2902.
31. Gillemans, N., Tewari, R., Lindeboom, F., Rottier, R., de Wit, T., Wijgerde, M., Grosveld, F., and Philipsen, S. (1998) Altered DNA-binding specificity mutants of EKLF and Sp1 show that EKLF is an activator of the beta-globin locus control region in vivo. *Genes Dev.* **12**, 2863–2873.
32. Jackson, D. A., McDowell, J. C., and Dean, A. (2003) Beta-globin locus control region HS2 and HS3 interact structurally and functionally. *Nucleic Acids Res.* **31**, 1180–1190.
33. Drissen, R., Palstra, R. J., Gillemans, N., Splinter, E., Grosveld, F., Philipsen, S., and de Laat, W. (2004) The active spatial organization of the {beta}-globin locus requires the transcription factor EKLF. *Genes Dev.* **18**, 2485–2490.
34. Forsberg, E. C., Downs, K. M., and Bresnick, E. H. (2000) Direct interaction of NF-E2 with hypersensitive site 2 of the beta-globin locus control region in living cells. *Blood* **96**, 334–339.
35. Kooren, J., Palstra, R. J., Klous, P., Splinter, E., von Lindern, M., Grosveld, F., and de Laat, W. (2007) Beta-globin active chromatin hub formation in differentiating erythroid cells and in p45 NF-E2 knock-out mice. *J. Biol. Chem.* **282**, 16544–16552.
36. Fiering, S., Epner, E., Robinson, K., Zhuang, Y., Telling, A., Hu, M., Martin, D. I., Enver, T., Ley, T. J., and Groudine, M. (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.

37. Du, M. J., Lv, X., Hao, D. L., Zhao, G. W., Wu, X. S., Wu, F., Liu, D. P., and Liang, C. C. (2008) MafK/NF-E2 p18 is required for beta-globin genes activation by mediating the proximity of LCR and active beta-globin genes in MEL cell line. *Int. J. Biochem. Cell. Biol.* **40**, 1481–1493.
38. Martin, F., van Deursen, J. M., Shivdasani, R. A., Jackson, C. W., Troutman, A. G., and Ney, P. A. (1998) Erythroid maturation and globin gene expression in mice with combined deficiency of NF-E2 and nrf-2. *Blood* **91**, 3459–3466.
39. Kuroha, T., Takahashi, S., Komeno, T., Itoh, K., Nagasawa, T., and Yamamoto, M. (1998) Ablation of Nrf2 function does not increase the erythroid or megakaryocytic cell lineage dysfunction caused by p45 NF-E2 gene disruption. *J. Biochem.* **123**, 376–379.
40. Derjuga, A., Gourley, T. S., Holm, T. M., Heng, H. H., Shivdasani, R. A., Ahmed, R., Andrews, N. C., and Blank, V. (2004) Complexity of CNC transcription factors as revealed by gene targeting of the Nrf3 locus. *Mol. Cell. Biol.* **24**, 3286–3294.
41. Meier, N., Krpic, S., Rodriguez, P., Strouboulis, J., Monti, M., Krijgsveld, J., Gering, M., Patient, R., Hostert, A., and Grosveld, F. (2006) Novel binding partners of Ldb1 are required for haematopoietic development. *Development* **133**, 4913–4923.
42. Wen, J., Huang, S., Rogers, H., Dickinson, L. A., Kohwi-Shigematsu, T., and Noguchi, C. T. (2005) SATB1 family protein expressed during early erythroid differentiation modifies globin gene expression. *Blood* **105**, 3330–3339.
43. Cai, S., Lee, C. C., and Kohwi-Shigematsu, T. (2006) SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat. Genet.* **38**, 1278–1288.
44. Dijon, M., Bardin, F., Murati, A., Batoz, M., Chabannon, C., and Tonnelle, C. (2008) The role of Ikaros in human erythroid differentiation. *Blood* **111**, 1138–1146.
45. Keys, J. R., Tallack, M. R., Zhan, Y., Papathanasiou, P., Goodnow, C. C., Gaensler, K. M., Crossley, M., Dekker, J., and Perkins, A. C. (2008) A mechanism for Ikaros regulation of human globin gene switching. *Br. J. Haematol.* **141**, 398–406.
46. Bell, A. C., West, A. G., and Felsenfeld, G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**, 387–396.
47. Wallace, J. A., and Felsenfeld, G. (2007) We gather together: insulators and genome organization. *Curr. Opin. Genet. Dev.* **17**, 400–407.
48. Farrell, C. M., West, A. G., and Felsenfeld, G. (2002) Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol.* **22**, 3820–3831.
49. Bulger, M., Schubeler, D., Bender, M. A., Hamilton, J., Farrell, C. M., Hardison, R. C., and Groudine, M. (2003) A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. *Mol. Cell. Biol.* **23**, 5234–5244.
50. Splinter, E., Heath, H., Kooren, J., Palstra, R. J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006) CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. *Genes. Dev.* **20**, 2349–2354.
51. Bender, M. A., Byron, R., Ragoczy, T., Telling, A., Bulger, M., and Groudine, M. (2006) Flanking HS-62.5 and 3′ HS1, and regions upstream of the LCR, are not required for beta-globin transcription. *Blood* **108**, 1395–1401.
52. Wai, A. W., Gillemans, N., Raguz-Bolognesi, S., Pruzina, S., Zafarana, G., Meijer, D., Philipsen, S., and Grosveld, F. (2003) HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. *EMBO J.* **22**, 4489–4500.
53. Noordermeer, D., Branco, M. R., Splinter, E., Klous, P., van Ijcken, W., Swagemakers, S., Koutsourakis, M., van der Spek, P., Pombo, A., and de Laat, W. (2008) Transcription and chromatin organization of a housekeeping gene cluster containing an integrated beta-globin locus control region. *PLoS Genet.* **4**, e1000016.
54. Wendt, K. S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T., Yahata, K., Imamoto, F., Aburatani, H., Nakao, M., Imamoto, N., Maeshima, K., Shirahige, K., and Peters, J. M. (2008) Cohesin mediates transcriptional insulation by CCCTC-binding factor. *Nature* **451**, 796–801.
55. Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H. C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., Cobb, B. S., Yokomori, K., Dillon, N., Aragon, L., Fisher, A. G., and Merkenschlager, M. (2008) Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell* **132**, 422–433.
56. Ragoczy, T., Telling, A., Sawado, T., Groudine, M., and Kosak, S. T. (2003) A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. *Chromosome Res.* **11**, 513–525.
57. Brown, J. M., Leach, J., Reittie, J. E., Atzberger, A., Lee-Prudhoe, J., Wood, W. G., Higgs, D. R., Iborra, F. J., and Buckle, V. J. (2006) Coregulated human globin genes are frequently in spatial proximity when active. *J. Cell. Biol.* **172**, 177–187.
58. Ragoczy, T., Bender, M. A., Telling, A., Byron, R., and Groudine, M. (2006) The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. *Genes Dev.* **20**, 1447–1457.
59. Schubeler, D., Francastel, C., Cimbora, D. M., Reik, A., Martin, D. I., and Groudine, M. (2000) Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev.* **14**, 940–950.
60. Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J. A., Lopes, S., Reik, W., and Fraser, P. (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* **36**, 1065–1071.
61. Simonis, M., Klous, P., Splinter, E., Moshkin, Y., Willemsen, R., de Wit, E., van Steensel, B., and de Laat, W. (2006) Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture-on-chip (4C). *Nat. Genet.* **38**, 1348–1354.
62. Mitchell, J. A., and Fraser, P. (2008) Transcription factories are nuclear subcompartments that remain in the absence of transcription. *Genes Dev.* **22**, 20–25.
63. Palstra, R. J., Simonis, M., Klous, P., Brasset, E., Eijkelkamp, B., and de Laat, W. (2008) Maintenance of long-range DNA interactions after inhibition of ongoing RNA Polymerase II transcription. *PLoS ONE* **3**, e1661.