Transcriptional regulation of the β-Globin locus

of mice and man

Regulatie van transcriptie van de β-globine loci

in muizen en in mensen

Proefschrift

Ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de Rector Magnificus
Prof. dr P.W.C. Akkermans
en volgens besluit van het College voor Promoties

De openbare verdediging zal plaatsvinden op
Vrijdag 11 september 1998 om 16.00 uur

door

Tolleiv Trimborn

Promotiecommissie

Promotor:	Prof. dr F. Grosveld
Overige Leden:	Prof. dr J. H. J. Hoeymakers
	Prof. dr B. Löwenberg
	Prof. dr H. Stunnenberg
Co-promotor:	Dr P. Fraser

Dit proefschrift is tot stand gekomen binnen de vakgroep Celbiologie van de Faculteit der Geneeskunde en Gezondheidswetenschappen van de Erasmus Universiteit te Rotterdam. De vakgroep maakt deel uit van het Medisch Genetisch Centrum Zuidwest Nederland. Het onderzoek is financieel ondersteund door de stichting voor Nederlands Wetenschappelijk Onderzoek (NWO).

Aan mijn ouders

Contents

Scope	7
General introduction	9
Chapter I: Chromatin	17
Euchromatin versus heterochromatin	
Chromatin structure and organisation	
Histones	19
Histone acetylation	
Histone deacetylation	20
Chromatin remodelling: swi/snf the crowbar of the nucleus	21
Heterochromatin and gene silencing	24
Polycomb and trithorax group proteins	25
Regulatory elements	26
Insulator	26
Locus control region	27
MARs/SARs	28
Chapter II: The human β-globin locus	31
General introduction	
Structure, organisation and expression of the β-globin genes	32
Epsilon-globin	
Gamma-globin	
Delta-globin	33
Beta-globin	
Locus Control Region	35
Properties of individual hypersensitive sites of the β-globin LCR	
Minor regulatory regions in the human β-globin locus	39
Epsilon silencer	
Stage selector element	40
3'γ-enhancer	
5'δ-region	41
3'β-enhancer	42
MARs/SARs	
Globin gene switching during development	43

Mouse model to study developmental regulation of the human β-globin locus 45
Erythroid specific transcription factors 45
EKLF 46
GATA-1 47
NF-E2
Replication of the human β-globin locus 48
Mouse β-globin locus 49
Models of LCR mediated gene-activation in the human β-globin locus 50
Scanning model
Random activation model
Looping model
Pimary transcript in situ hybridisation 52

Chapter III: Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. 57
Cell 1996 87: 105-114.

Chapter IV: The effect of distant on long-range chromatin interactions 69
Molecular Cell 1997 1: 131-139.

Chapter V: Chromatin interaction mechanism of transcriptional control *in vivo.* 79
in press

Chapter VI: Mechanisms of developmental control of transcription in the murine α and β globin loci 91
Submitted for publication

Concluding remarks 105
References 108
Summary 120
Samenvatting voor vrienden 121
List of abbreviations 124
Curriculum vitae 125
List of publications
Nawoord 126

Scope of this thesis

The aim of this thesis was to get more insight into the mechanism of transcriptional activation of the human β-globin locus.

Chapter 1 serves as an introduction to chromatin. The different aspects of chromatin are briefly described. **Chapter 2** reviews the work on the human β-globin locus over the last decade. This review aims to present some background information for the more recent experiments that are described in chapters 3 to 6. **Chapter 3** shows the results of individual hypersensitive site deletions of the β-globin LCR: the transgenes become subject to position effects and a novel position effect is described. In **chapter 4** one of the parameters that is important in the activation of the globin genes is described: the frequency of LCR-gene interactions. **Chapter 5** presents evidence that the locus control region activates only one gene at a time. Finally in **chapter 6** detailed analyses of the mouse β-globin locus is shown and the data suggest that the mechanism of gene activation is conserved during evolution. In concluding remarks I discuss some implications of our work and future experiments. For more detailed considerations of specific findings, the reader is referred to the discussion of chapters 3 to 6.

General introduction

GENERAL INTRODUCTION

In a multi-cellular organism haemoglobin facilitates the transport of oxygen from the lungs to the different organs and of carbon dioxide from the tissues back to the lungs. An erythroid cell mainly consists of soluble haemoglobin protein..

Haemoglobin is a tetrameric metallo-protein that consists of two alpha (α) and two beta (β) polypeptides each of which is combined with a haem-group and gives blood its characteristic red colour. Without the polypeptide chains the reaction of oxygen with the haem-group would not be reversible; the folding of the protein protects the oxygen molecule from being irreversibly linked to the haem-group.

A balanced production of α and β-globin proteins in a red blood cell is crucial. An imbalance, for example a lack of beta-chains, causes precipitation of the polymerising alpha chains. This subsequently deforms the cell membrane and destroys the erythrocyte. Tight regulation of the genes involved in producing these proteins is required; especially because a number of different globin genes are involved during mammalian development (3 alpha genes and 5 beta genes, Fig. 1).

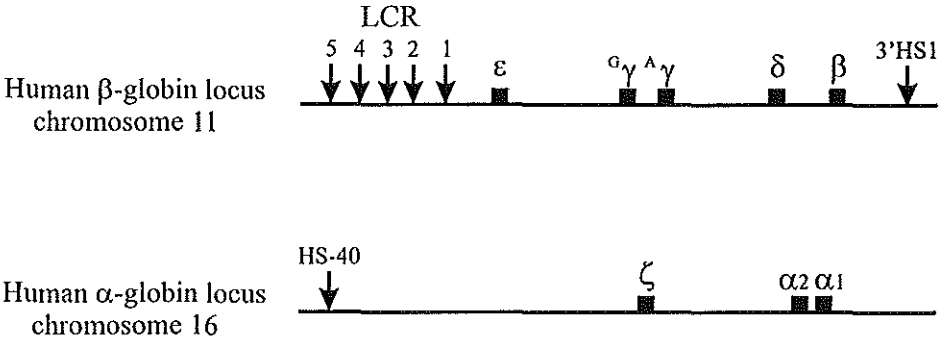


Figure 1. Schematic representation of the human alpha and bèta-globin loci. Arrows indicate the hypersensitive sites of the locus control region (LCR).

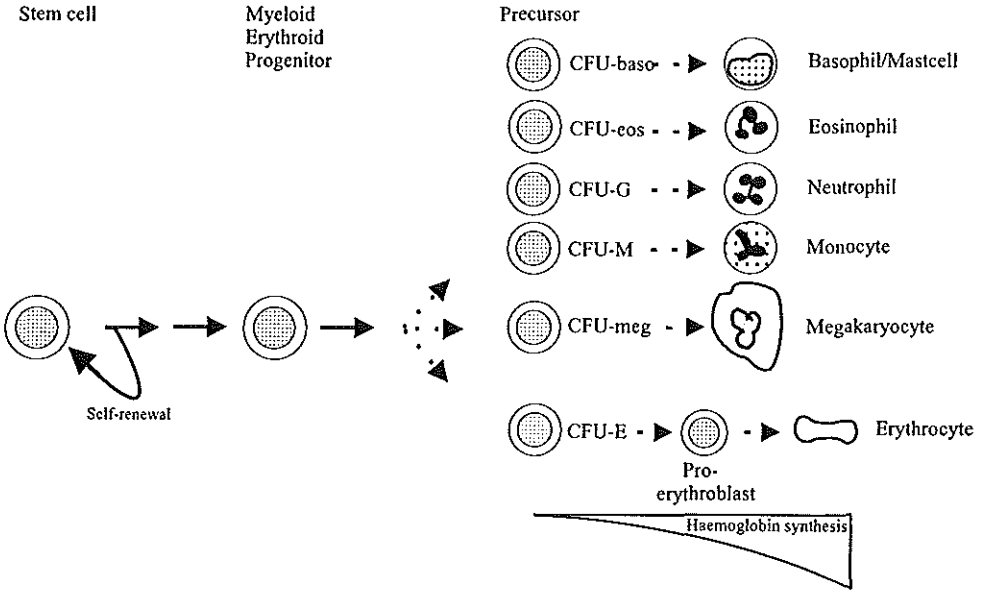


Figure 2. Schematic representation of human haematopoiesis.
Cells with their nucleus (shaded) are depicted as circles. Globin synthesis starts at the proerythroblast stage, the nuclear size of the erythrocyte decreases with ongoing differentiation and in the final step of differentiation, the cell becomes enucleated.

Haematopoiesis

The red cell originates from the hematopoietic stem cell and starts to express the globin genes in its terminal phase of differentiation (Fig. 2, reviewed in Stamatoyannopoulos, 1994). The characteristic feature that defines a stem cell is its unlimited capacity of self-renewal; in other words a stem cell is able to divide without differentiating into any of the potential blood cell lineages. An additional characteristic is its pluripotency to differentiate into progenitor cells of each of the different hematopoietic lineages. The different cell-types that arise from the hematopoietic stem cell are the following: early in development microglia and Kupfercells and subsequently megakaryocytes, macrophages, neutrophils, granulocytes, eosinophils and erythrocytes.

The progenitor cell is the second stage of the cell in erythropoiesis; this cell is committed to a lineage but it still retains considerable proliferative capacity. Progenitors are characterised by their ability to form precursor colonies in semisolid culture medium. The loss of proliferative capacity and the start of terminal maturation distinguish the transition of progenitor to precursor, this includes obtaining specialised function, like the production of haemoglobin (Fig. 2). In the final step of differentiation the cell undergoes several divisions from pro-erythroblast to acidophillic erythroblast to erythrocyte. The erythrocyte loses its nucleus, mitochondria and endoplasmatic reticulum and consists mainly of haemoglobin.

Haemoglobinopathies

Thalassemia

Alpha or beta-thalassemia (β-thal) refers to a heterogeneous group of conditions characterised by the decrease or deficiency of either the alpha or the beta globin polypeptide in a red blood cell. At the molecular level the mutation may vary from a single point mutation in the coding or regulatory sequences to large deletions of a particular gene (Fig. 3; Collins and Weissman 1984). For example, in case of β-thal the imbalance in the α to β-polypeptide ratio causes precipitation of the alpha proteins.

Sickle cell anaemia

Sickle cell anaemia is the result of one amino acid change in the β-globin protein. The sixth amino acid has changed from a glutamic acid to a valine, leading to the production of the sickle haemoglobin protein (HbS). HbS polymerises under deoxygenated conditions resulting in morphologically abnormal cell-types including typical sickle shaped cells. These cells are the cause of vaso-occlusion of endothelial tissues affecting several organs, in other words red cells in sickle cell patients loose their deformability and obstruct and destroy microcirculation (reviewed in Kaul et al, 1996)

Hereditary persistence of fetal haemoglobin (HPFH)

After birth the γ-globin genes are silenced, their expression levels drop gradually to a few percent of total globin expression and these levels are maintained throughout adulthood. In the case of HPFH γ-gene expression persists at higher levels in the adult stage, instead of being silenced properly. Several mutations have been identified that correlate with the HPFH phenotype; these mutations vary from point mutations in the γ-promoter region to large deletions at the 3'end of the globin locus (Fig. 3; Collins and Weissman, 1984)). The point mutations in the promoter region are thought to affect the binding of the putative repressors/silencers of the γ genes thereby keeping the genes active. Deletions 3' of the $^{A}\gamma$ gene are very heterogeneous and therefore not easy to interpret. A number of patients have been described (Fig. 3) with different deletions in the 3' ψβ area, some of these patients show an HPFH phenotype, others a δβ-thal phenotype (reviewed by Poncze et al.,1988).

The clinical importance of the HPFH phenotype is that it can partially rescue a β-thal or sickle cell anaemia. Patients that have both sickle cell anaemia (or β-thal) and HPFH have a much milder phenotype and in general do not need blood transfusions. The generation of an HPFH like phenotype, i.e. the reactivation of the γ-gene(s), is therefore an important goal.

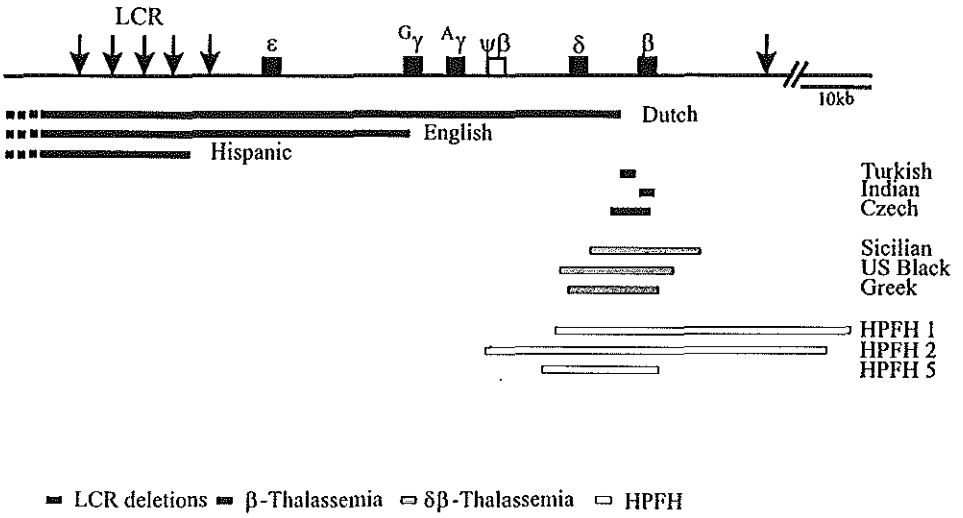


Figure 3. Schematic representation of the various naturally occurring deletions in the human β-globin locus resulting in thalassemias and HPFH mutations. The black bars indicate the γδβ thalassemias, these mutations delete the 5'end of the locus. The dark grey bars indicate a few informative β-thalassemia mutants. The light grey bars show the deletions leading to a δβ-thalassemia. The white bars indicate relatively small deletions that result in an HPFH phenotype.

Chapter I

Chromatin

CHROMATIN

Euchromatin versus Heterochromatin

Two states of chromatin have been described (Heitz, 1928) based on cytological studies of interphase nuclei: chromatin that stays in a condensed form throughout the cell cycle, called heterochromatin and the rest of the chromatin which undergoes a typical cycle of condensation and unravelling, named euchromatin[1]. It was later shown that in general heterochromatin is transcriptionally inactive and that it is located at the periphery of the nucleus, whereas euchromatin contains the active genes and is generally found more internally in the nuclear space. Heterochromatin replicates late in S phase and it consists of a high percentage of middle and highly repetitive DNA. Centromeres, the constriction in chromosomes that facilitates chromosome separation during mitosis, are known to contain a high number of repetitive sequences and are shown to be heterochromatic.

Chromatin structure and organisation.

Chromatin has to be organised in order to fit physically into a eukaryotic nucleus, several levels of organisation are described to explain the necessary compaction. The first level of compaction is realised by wrapping the DNA around nucleosomal core particles or nucleosomes to form a 10 nm fibre. The second level of compaction is the formation of a 30 nm fibre which is mainly dependent on histone H1 (Thoma et al., 1979; Dasso et al., 1994) and HMG (high mobility group) (Wolffe, 1995) proteins. The third level of compaction is achieved by the formation of loop-structures that are attached to the nuclear scaffold/matrix, however this higher order folding is as yet poorly understood and remains the subject of much debate (Cook and Brazel, 1976; Benyajati and Worcel, 1976; Paulson and Laemmli, 1982; Mirkovitch et al., 1984).

Nucleosomes are octamer protein complexes that consist of histone proteins, they are the fundamental building blocks of eukaryotic chromatin (Kornberg, 1974, 1977). An octamer contains two molecules of each core-histone H2A, H2B, H3 and H4, about 146 bp of DNA can be wrapped around the surface of an octamer. The first step in nucleosome assembly is the formation of a tetramer between H3 and H4, the DNA wraps around this tetramer such that two dimers of H2A and H2B can stably associate with it. The nucleosome-DNA interaction is further stabilised by linker histone H1 that binds at one side of the nucleosome

[1] It should be noted that all chromosomes are highly condensed during mitosis and that during interphase chromosomes become decondensed.

to the linker DNA (DNA spacing two nucleosomes). Nucleosome positioning on DNA is sequence independent but is determined by a number of other parameters. Experiments with *in vitro* reconstituted chromatin, using Drosophila extracts, have shown that the first step in nucleosome assembly is energy independent and the nucleosomes are positioned randomly on the DNA. The second step requires ATP and depends on a chromatin remodelling complex CHRAC (Varga-Weisz, 1997). Nucleosome position should be regarded as a stochastic rather than a fixed point on the DNA because nucleosomes exhibit short-range mobility.

Experimental evidence has indicated that the nucleosomal structure of chromatin restricts the access to the DNA template (for review see Wolffe 1994; Felsenfeld, 1996). Positioning of nucleosomes is thought to repress gene expression by placing regulatory sites in an inaccessible position on the nucleosome. Comparison of transcription initiation *in vitro* using free DNA and DNA reconstituted with nucleosomes showed that in the latter case transcription initiation is significantly reduced (O'Neil et al., 1992,1993,1995). If linker H1 is added to reconstituted chromatin, using Xenopus extracts, it was found to restrict nucleosome mobility and to repress transcription even more (Ura et al., 1995). However, it has been shown that the RNA polymerases are able to transcribe through nucleosome cores (Studitsky et al., 1994). These experiments indicate that chromatin conformation in general prevents the initiation of transcription, but not the subsequent elongation of transcription. How the transcription machinery "reads through" the assembled nucleosomes remain as yet unclear.

From these experiments it is clear that genes need to overcome the repressive effects of nucleosomal organisation to become activated. Nucleosomes at the promoter have to be disassembled, distorted or displaced in order to allow initiation of transcription. Several models and ideas have been put forward to explain this mechanism, among these are the co-operative mode of action of transcription factors, histone acetylation and chromatin remodelling complexes.

The co-operative mode of action model suggests that an array of nucleosomes which cover an area of transcription factor binding sites is displaced by an interplay of different transcription factors; suggesting that the first transcription factor that binds to its site on the DNA exploits the dynamic flexibility of a nucleosome[2] (Fig. 4). It was shown that the displacement can occur independent of DNA replication. The initial binding of a transcription factor would destabilise the nucleosome/DNA interaction allowing subsequent binding of additional transcription factors (Polach and Widom, 1995).

[2] It is assumed that DNA at the edge of an octamer is less tightly bound (Simpson, 1979)

Histones

Histones are small, basic proteins which are highly conserved in evolution especially H3 and H4, whereas H2A and H2B are less conserved (van Holde, 1995). Each core histone consists of two domains: an extended histone-fold domain at the carboxyl (C)-terminus of the protein through which histone-histone or histone–DNA interactions are established and charged tails at the amino (N)-terminal domain containing the bulk of the lysine residues. It is at these charged tails where important post-translational modifications take place. Modifications (e.g. acetylation, methylation, phosphorylation and ubiquitination) at these tails have important consequences for chromatin structure and function.

Histone acetylation

Histone acetylation was linked with active gene transcription in eukaryotic cells back in 1964 (Allfrey et al. 1964). Acetylation, a process that is reversible, occurs at lysine residues of the positively charged tails of the core-histones, mediated by enzymes named histone acetyltransferases (HATs). Neutralisation of the positive charge of the histone tails leads to a reduced affinity for DNA (Hong et al. 1993), thereby altering nucleosomal conformation and stability. This implies that acetylation is linked to transcriptional activity, since tight nucleosome binding is linked to transcriptional repression. Mapping experiments of acetylated histones indeed indicated a correlation between hyperacetylation and active genes/domains (Jeppesen and Turner, 1993; Hebbes et al., 1988, 1994).

Histones in euchromatin can be modified at all lysine residues that are present in their tail, however not every lysine residue of a particular core histone is acetylated. This seems to depend on the specificity of the HAT (Clarke et al., 1993; Kuo et al., 1996; Yang et al., 1996; Bannister and Kouzarides, 1996; Ogryzko et al., 1996). In yeast and *Drosophila* it was found that H4 in heterochromatin is hypoacetylated at residues K5 and K8, but was hyperacetylated at residue K12 (Turner et al., 1992). In contrast, H4 of the inactive mammalian X-chromosome is hypo-acetylated at all residues (K5, K8, K12, and K16) (Jeppesen and Turner, 1993). Strong evidence for the relationship between transcription and acetylation came from experiments in yeast where the yGCN5p, a co-activator of transcription was shown to exert HAT activity (Georgakopoulos and Thireos, 1992; Kuo et al., 1996). This complex interacts with specific transcription factors as well as the basal transcription machinery (Marcus et al., 1994; Georgakopoulos et al., 1995; Candau and Berger, 1996). Recently it was shown that yGCN5 is part of a multi-subunit complex of approximately 2 Mda, called SAGA (Spt-Ada-Gcn5-Acetyltransfease; Hampsey, 1997).

In mammals it was also shown that HATs are present in or linked closely to the basal transcription complex. The p300/CBP HAT activity (Bannister and Kouzarides, 1996; Ogryzko et al., 1996) was initially described as a co-activator that functions by interacting with a wide variety of enhancer binding proteins (reviewed in Janknecht and Hunter 1996).
Taf 130/250 HAT is another HAT associated with the basal transcription machinery; it is a subunit of the TFIID complex, a basic component of the Pol II transcription initiation complex (Mizzen et al., 1996). This could indicate that HAT activity should be regarded as a more general component of the transcription machinery.
Histone acetylation hints at additional levels of transcriptional regulation. Two basic models have been proposed to explain how the modifying activities of HATs can affect transcription activity in a gene specific manner (reviewed in Struhl 1998).
The first model proposes a general promoter targeting mechanism. The histone modifying activity is targeted in a non-specific manner to individual promoters. The TAF130/250 HAT activity argues in favour of such a general model, however physiological relevance of this activity has not yet been shown. Gene specific expression in this case depends on the intrinsic chromatin structure, protein-binding sites and rate-limiting steps for transcriptional initiation.
The second model suggests a gene-specific targeting model (Fig. 4). Such a model requires sequence specific targeting of the histone modifying activity, which could be accomplished by sequence specific, DNA-binding proteins. A number of experiments have been published that would fit this model. Kuo et al (1998 and references therein) have shown that yGcn5p acetylates histones *in vivo* with a strong preference for promoter regions. Furthermore, preferential histone acetylation at the promoter correlates with Gcn5-dependent transcription, indicating its relevance *in vivo*. The fact that more and more proteins that are involved in transcription regulation show intrinsic HAT activity combined with the fact that these HATs act on different templates indicates an additional level of gene regulation.

Histone deacetylation

As mentioned previously histone acetylation is a reversible modification and the enzyme that is responsible for removing the acetyl-group is called histone deacetylase. Deacetylation of histones has been correlated with repression of transcription. A number proteins that were implicated in transcriptional repression (e.g. Sin3 and YY1) have now been shown to be associated with a histone deacetylase such as Rpd3 in yeast (Yang et al., 1996; Kadosh and Struhl, 1997) or its human homologue HDAC1 (Taunton et al., 1996). Recent results indicate that Rpd3 histone deacetylase is important for transcriptional repression *in vivo*, but not absolutely required. All mutants of Rpd3 in this study showed some residual histone deacetylase activity (Kadosh and Struhl, 1998). Thus, the level of histone acetylation is

determined by the action of both histone acetyl transferases and histone deacetylases.
Histone deacetylation activity can be inhibited by several drugs, among these are trichostatin A, sodium butyrate[3](Vidale et al., 1978) and trapoxin. Experiments with transfected HeLa cells showed that in 80% of the clones the *lac*Z reporter gene is not transcriptionally active. After treatment with either trichostatin A or butyrate all clones analysed showed reactivation of *lac*Z expression (Chen et al., 1997). However, it should be noted that the secondary effects of these inhibitors are not well understood: It is known that trichostatin A affects cellular processes like differentiation, apoptosis and DNA synthesis.

Chromatin remodelling: SWI/SNF the crowbar of the nucleus

Several protein complexes have been described that are capable of remodelling chromatin conformation or that are able to disrupt nucleosome assembly to facilitate transcription factor binding at their recognition sites. These include SWI/SNF, NURF (Nucleosome Remodelling Factor; Tsukiyama et al., 1995), RSC (Remodelling the structure of chromatin, Cairns et al., 1996), CHRAC (Chromatin Remodelling Activity; Varga-Weisz et al., 1997) and ACF (Ito et al., 1997). The best studied complex is SWI/SNF, originally found through genetic mutants in yeast that cause a defect in transcription affecting both mating type switching (SWI, Peterson and Herskowitz, 1992)) and sucrose fermentation (SNF, Laurent et al., 1990). The complex consists of at least 11 different polypeptides, and it induces the transcription of a limited number of genes and thus is regarded as a regulator of transcription. It was shown that several core-histone mutations were able to suppress SWI/SNF defects and together with the observation that SWI/SNF mutations prevent the formation of 'active' chromatin conformation at the Suc2 promoter (Hirschhorn et al., 1992), this suggested that the SWI/SNF complex functions via interaction with chromatin structures. Indeed both human and yeast SWI/SNF complexes function *in vitro* by disrupting chromatin structure, facilitating factor binding (Coté et al., 1994; Imbalzano et al., 1994). Biochemical data have shown that the yeast SWI/SNF chromatin remodelling activity is ATP-dependent. The SWI2/SNF2 subunit contains an ATP-dependent ATPase domain that, at least *in vitro*, was shown to catalyse ATP hydrolysis in a DNA-dependent manner. Deletion of the ATP-binding site abolishes SWI/SNF activity both *in vitro* (Coté et al., 1994) and *in vivo* (Laurent et al., 1993).
The SWI/SNF complex is thought to exert its function directly on the nucleosomal DNA,

[3] Sodium butyrate is used on patients with a β-thal or sickle cell anaemia, to re-activate the γ-genes. The treatment is controversial since the patients suffer a number of unpleasant side effects.

rather than inducing nucleosomal displacement, thereby disrupting histone-DNA interaction.
To date, it is still not clear how SWI/SNF activity is targeted to specific DNA regions. Since it has been shown that SWI/SNF interacts in a sequence independent manner, targeting could be envisioned to be the result of possible interactions with transcriptional co-activators (Yoshinaga et al., 1992) or with the Pol II holoenzyme. In yeast it has been shown that SWI/SNF associates with the RNA Pol II holoenzyme (Wilson et al., 1996), thus coupling chromatin remodelling directly to transcription. This was also shown for the human SWI/SNF subunit BRG1 (Neish et al., 1998; Fig. 4).
A more general model, or "window of oppurtunity model" , assumes that there is no targetting of the remodeling complex; nucleosome displacement is a dynamic process that occcurs randomly and a transcription factor could use this window of opportunity to bind to the DNA when it is accessible.

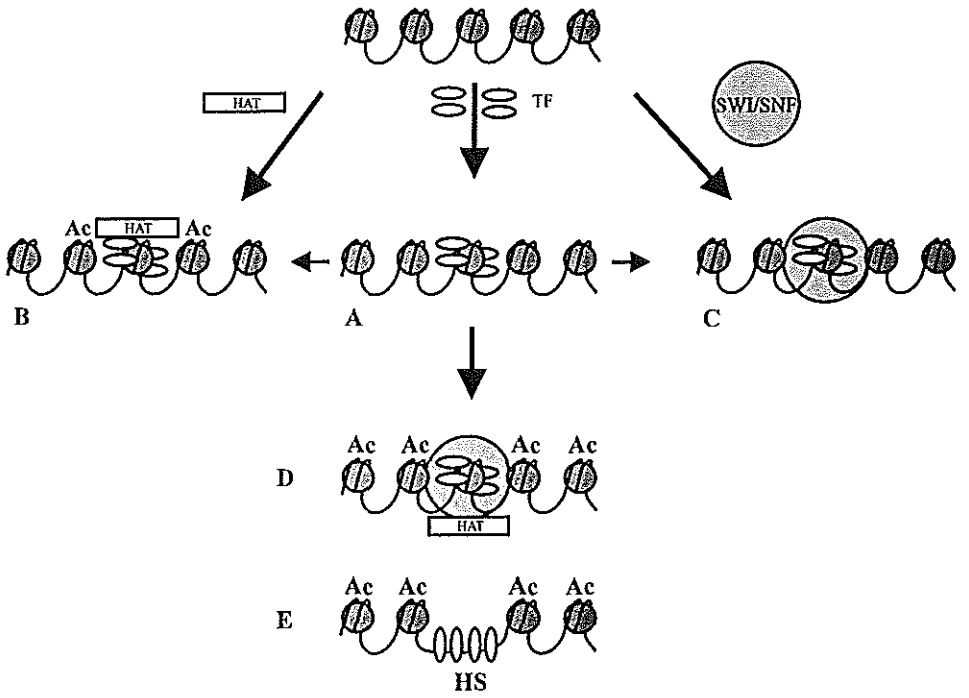


Figure 4. A schematic representation of the different possibilities to overcome the repressive nucleosome structure of chromatin.
This model is used to summarise the different mechanisms that have been described above that can disturb chromatin conformation to allow transcription. It should be noted that this model is highly speculative and that it does not necessarily represent genes become transcriptionally activated.
The initial step in the disruption of the nucleosomes is the sequence-specific binding of a transcription factor, causing destabilisation of the octamer. There are two possibilities either the nucleosomes are displaced by co-operative binding of the transcription factors (A) or HATs are recruited by e.g. transcription factors (TF) to acetylate the histone tails (B). This will destabilise nucleosome-DNA association, facilitating the transcription factor binding at the transcription factor binding-site. If the HAT is a co-activator of the basal transcription machinery this might be sufficient to initiate transcription.
Alternatively, the disruption of the nucleosomes only by transcription factors might not be sufficient and a chromatin remodelling activity is needed (C). In this case, recruitment of factor(s) is required, perhaps by one of the bound factors, that facilitates the displacement of the nucleosomes in an ATP-dependent manner. It is possible that all three mechanisms described above (A, B, C) act in concert in order to facilitate opening up of the chromatin structure (D). The DNA that now has become free of nucleosomes (E) will be occupied by the transcription factors that ensure a stable "active" promoter conformation (indicated by HS).

Heterochromatin and gene silencing

Heterochromatin is, except for a well-defined cytological description, poorly understood at the molecular level. The two best-studied examples of heterochromatic regions are the inactive X-chromosome in mammals and centromeric heterochromatin in *Drosophila*. It was shown that most genes which reside on the inactive X-chromosome are transcriptionally silent[4] and that this state is passed on from one cell to its progeny. Furthermore, it was shown that this inactivation was not restricted to genes on the X-chromosome. Upon translocation of an autosomal chromosome domain next to a heterochromatic X region, genes in this autosomal domain were repressed in a distance-dependent manner, implying a 'spreading of heterochromatin' mechanism. Similar spreading effects had been described earlier for *Drosophila* heterochromatin (reviewed by Karpen, 1994).

Position Effect Variegation (PEV) was first shown in *Drosophila*; using X-rays to create mitotic exchanges, Muller (1930) showed that upon artificial chromosomal rearrangements, euchromatic gene activity could be inhibited by juxtaposed heterochromatin resulting in the current definition of PEV: the clonally inherited variable pattern of gene expression observed in a subset of cells.

It turned out that transgenes introduced into the genomes of *Drosophila* (Wilson et al., 1990), mice (Festenstein et al., 1996; Milot et al., 1996) or plants (Matzke and Matzke, 1998) can be subject to PEV. The analysis of such transgenics serves as a handle to study PEV and the effect of heterochromatin structure and dynamics. PEV was also observed when multiple transgenes integrated as an array at the same genomic site (Dorer and Henikhoff, 1994; Garrick et al., 1998).

Several models have been proposed to explain PEV. The mass action model proposes that heterochromatin should be considered as a complex of DNA and proteins (Locke et al., 1988). The assembly of a functional unit of heterochromatin depends on the correct stoichiometry of each subunit. Heterochromatin formation may spread co-operatively along the DNA in a linear and polar fashion. This model implies that heterochromatin formation starts at a given initiator and is blocked at a terminator, which would reside in euchromatin. Evidence for such a model comes from experiments that have shown that there is polarity in the effect of a heterochromatic breakpoint on nearby euchromatic genes, in such a way that the genes closer to the heterochromatin are more severely affected (Spofford, 1976). Another model that was proposed to explain PEV in *Drosophila* suggested pairing-driven compartmentalisation, whereby the exchange of heterochromatic proteins in *trans* occurs as a

[4] Except for those genes that reside on a small part of the X-chromosome that is not inactivated

result of the tendency of heterochromatin domains to aggregate or pair (Wakimoto and Hearn, 1990). Recent experiments using FISH analysis by Sedat and colleagues have shown that a locus distal to the centromere associates with other heterochromatic regions on the chromosome in a stochastic manner, suggesting that the distal locus is silenced by specific contact with centromeric heterochromatin (Dernburg et al., 1996).

Polycomb and trithorax group proteins

In Drosophila a wide range of mutants have been identified genetically that either suppress PEV [su(var) mutants] or enhance PEV [e(var) mutants]. Only a few of these mutants encode structural proteins, of which heterochromatin protein 1 (HP1) is the best studied. HP1 is encoded by su(var)205 and contains a chromodomain that facilitates protein-protein interactions and which has been identified in other chromatin modulators (Paro and Hogness, 1991). Flies deficient in HP1 do not show proper chromosome condensation and this probably results in segregation defects because of centromere malfunction (Kellum and Alberts 1995).

Two groups of genes were initially identified as memory proteins, i.e. proteins that could maintain a 'determined chromatin state' over a number of cell-divisions/cell generations: the Polycomb Group proteins (Pc-G) and the Trithorax Group proteins (Trx-G). Both groups were identified as being transcriptional regulators of the homeotic selector genes (HOM). They maintain the differential expression pattern of these genes by keeping either the inactive (Pc-G) or active state (Trx-G) during development (Paro and Harte, 1996; Pirotta, 1997). It is important to note that these proteins are not involved in establishing a certain configuration, they only maintain an already existing state/structure. The Pc-G of proteins is a very heterogeneous set, although several conserved protein motifs were identified, suggesting that the Pc-G act via protein-protein interactions. One member, the polycomb protein (PC) was shown to have a chromodomain similar to HP1, raising the possibility that the Pc-G proteins could be involved in the formation of heterochromatin-like structures to maintain genes in a permanently silenced state. Based on a variety of genetic as well as biochemical data it has been suggested that Pc-G act as a multi-protein complex packaging large chromosomal domains into a repressed state, inaccessible to *trans*-acting factors (Zink and Paro, 1995; Pirotta, 1997). It is believed that the proteins are targeted by interacting with polycomb responsive elements (PREs), since PREs have the ability to correctly maintain the repressed state of a linked homeotic promoter in transgene constructs, even if this promoter is several kilobases away (Chan 1994; Zink and Paro, 1995). A number mammalian homologues have been identified over the last few years for Pc-G as well as Trx-G , among these are *bmi* 1 and *M33* (reviewed by Lohuizen, 1998).

The Trx-G of proteins is heterogeneous in sequence, structure, site of action and probably mechanism. Members of this group are Trithorax, brahma, GAGA factor and zeste (Pirotta, 1995, 1997). In many cases, flies carrying mutations in both trithorax and Pc-G genes have few phenotypic defects (Kennison et al., 1988), suggesting that trithorax function is needed to counteract Pc-G repression in those tissues in which expression is activated. Recent evidence suggests that Trx-G proteins act together with Pc-G complexes on common DNA elements (Strutt et al., 1997). Analysis of the Fab-7 cis regulatory element of the homeotic bithorax complex (BX-C) in Drosophila, showed that it is possible to relieve the determined, repressed state from the Fab-7 element. The derepressed state can now be stably transmitted throughout development and it was suggested that trx-G proteins are responsible for possible structural changes leading to an open chromatin conformation (Cavalli and Paro, 1998).

Regulatory Elements

Several elements have been described that are able to counteract the repressive effect of heterochromatin on gene transcription: insulators or boundary elements, Locus Control Regions (LCR) and Matrix attachment regions or Scaffold attachment regions (MAR/SAR).

Insulator

Insulators have been defined as elements that are able to separate active and inactive chromatin environments at distinct chromosomal positions. Insulator elements do not have intrinsic enhancer/repressor activities as has been shown for LCR's or MARs/SARs. A number of insulator elements have been characterised in several different species, the best known are the scs/scs' (Kellum and Schedl, 1991) and the su(Hw) [suppresser of hairy wing] (Roseman et al., 1993) elements from Drosophila and the 5'HS4 of the chicken β-globin locus (Chung et al., 1993). All elements were shown to exhibit classical insulator activity, tested by flanking a marker gene by two insulator elements and checking for position-independent expression upon integration. A second property indicative of insulator activity is that the candidate element should interfere with the expression of a gene when placed between its promoter and its enhancer.

In a search for proteins that bind to insulator elements, a 32 kDa protein was identified, called the boundary element associating factor (BEAF-32). The protein binds to a palindromic sequence within the scs-element in *Drosophila* and this element acts as a typical insulator in the different assays (Zhao et al., 1995). Immunolocalisation of the protein showed that it is present at specific subnuclear regions and as expected at the scs'-containing border. Furthermore, it was also located at the edges of puffs typically found in the larval stage of *Drosophila* development, indicating that BEAF has a more general structural role in defining

boundaries within the *Drosophila* genome.
As yet, it is not clear how insulators exert their function but several models have been proposed to explain the observed features (Geyer, 1997). One model, the boundary model, implies that the element forms a nucleoprotein complex. Such a complex would associate with similar complexes, looping out the domains in between and establishing a chromatin domain of independent gene activity. An alternative is the decoy model, which predicts that the insulator element forms a complex that is very similar to a promoter-complex, thereby "fooling" enhancers or repressors by attracting them in non-productive interactions.

Locus Control Region
Locus control regions (LCR) are characterised as being a collection of elements that contain all the genetic information to give a (multi-)gene locus high level, developmental and tissue-specific expression upon stable integration, independent of the site of integration. The human β-globin LCR was the first LCR to be identified (Grosveld et al., 1987) and subsequently several LCR's have been described, the T-cell specific CD2 (Greaves et al., 1989); the chicken lysozyme (Bonifer et al., 1990); the B cell specific MHC class II Ea gene (Carson and Wiles, 1993); the liver specific LAP gene LCR (Talbot et al., 1994);). The human growth hormone gene (Jones et al., 1995). The β-globin LCR consists of 5 DNase I hypersensitive sites (see below) that are thought to form a holocomplex which interacts with the globin genes to provide appropriate expression.
Differences in chromatin conformation can be detected by DNase I endonuclease treatment. It was shown that transcriptionally active genes and their 5' and 3' flanking sequences were more accessible to DNase I than non-transcribed regions of the genome (Garel and Axel, 1976; Weintraub and Groudine, 1976). This type of DNase I sensitivity is referred to as general DNase I sensitivity and is attributed to more general conformational chromatin changes, for example acetylation or methylation of DNA.
Another type of sensitivity is called DNase hypersensitivity which is thought to be the result of small nucleosome-free chromatin regions where the DNA is digested at relatively low DNase I concentrations. It was shown that promoter regions where nucleosomes are disrupted or displaced by transcription factors (or SWI/SNF, see also section chapter 1) are hypersensitive to DNase I digestion (Wu et al., 1979; Felsenfeld, 1992).
Deletion of individual hypersensitive sites of the β-globin LCR resulted in position effects for the linked globin genes (Milot et al., this thesis). Similar deletions of hypersensitive sites of the CD2-LCR were also shown to give rise to position effects (Festenstein et al. 1996). This suggested that only a complete LCR is capable of overcoming the negative, repressive effects of heterochromatin. Several models have been proposed to explain how LCR's exert their function (see chapter II, models).

MARs/SARs

Scaffold Attachment Regions (SAR) were characterised in *Drosophila* as A/T-rich sequences of varying size (0.5-several kb). Based on the model that chromatin is organised at a higher level by the formation of loop-structures (loops), the sizes of the loops were estimated to range in the order of 30-100kb (Gasser and Laemmli, 1985). It was hypothesised that these loops would not be a random phenomenon but that the loops would be organised in the nucleus by attachment, at their bases, onto the nuclear scaffold or matrix (based on the isolation procedure using either low or high salt it is either called a scaffold or a matrix). In addition it was proposed that important functional elements such as centromeres, telomeres and origins of replication might also be associated with the nuclear scaffold (Gasser et al., 1989; Cook et al., 1994). It is therefore important to identify those regions/sequences in the genome that are associated with the scaffold. SARs/MARs have been shown to stimulate the expression of reporter constructs they flank, but only when the transgenes were allowed to integrate into the host genome (Laemmli et al., 1992; Jenuwein et al., 1997). It has been concluded that SARs/MARs do not confer position-independent expression to linked or flanked transgenes, although this has been the subject of debate (for review see Geyer, 1997). Interestingly, a SAR isolated from the hβ-interferon gene displayed tissue-specific boundary activity (Thompson et al., 1994), indicating that the effect of a SAR/MAR is not static but can vary during development, depending perhaps on trans-acting factors. Recent data by Jenuwein (1998) suggest that the MARs flanking the core element of the enhancer of the immunoglobulin μ gene can generate extended chromatin accessibility over a distance of ~1kb. The authors raise the possibility that the complete immunoglobulin μ gene enhancer can be regarded as an LCR and that upon deletion of any of its components it does not function properly.

Chapter II

The Human β-globin locus

The Human β-globin locus

General introduction

The human β-globin locus is a multi-gene locus located on the short arm of chromosome 11. The locus spans approximately 70 kb and consists of five functional genes. The genes are expressed specifically in erythroid tissue at high levels and are arranged in the same order as they are expressed during development (5'- ε, γ_G, γ_A, δ, β -3', Fig. 5). During embryonic development, when erythropoiesis takes place in the blood-islands located in the yolk sac, the ε gene is expressed. Later in development, the fetal liver becomes the major site of erythropoiesis and gene expression switches from the ε-gene to the γ genes (Pescle et al., 1985). Finally, around the time of birth, the bone marrow becomes the predominant erythropoietic centre and a second gradual switch in transcription takes place from the $^{G}\gamma$ and $^{A}\gamma$ genes to the adult δ and β genes. (Fig. 5, Wood, 1976).

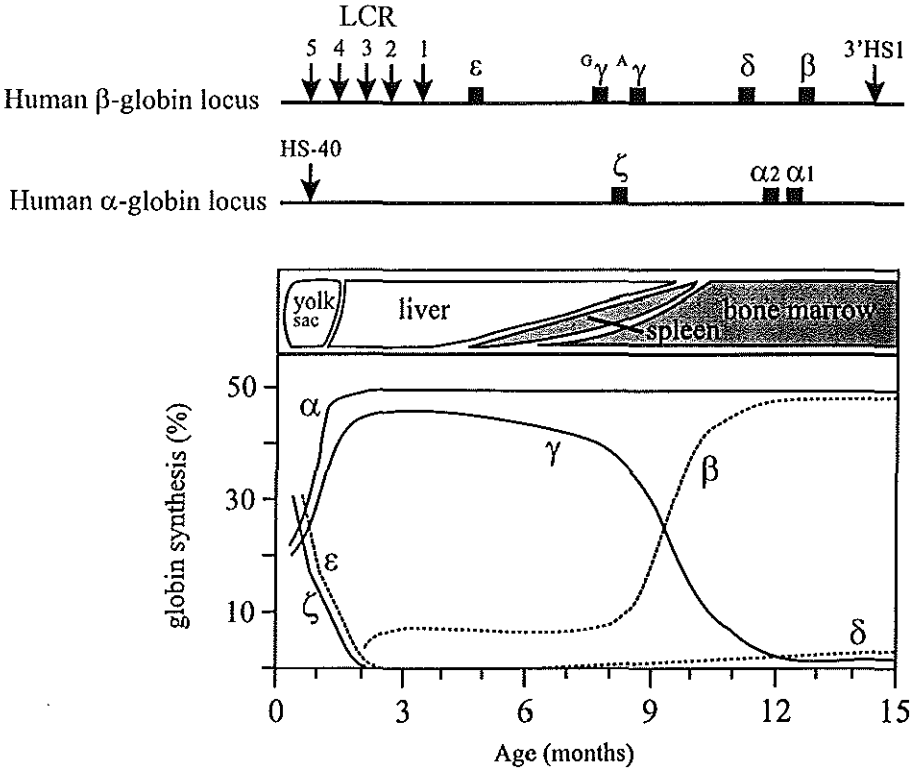


Figure 5. Organisation and developmental expression pattern of the human α-globin and β-globin loci. Top: Structure of the α-globin and β-globin loci. The arrows indicate the DNase hypersensitive sites and boxes represent the different genes. Bottom: Expression levels of the individual globin genes during development is shown. The sites of erythropoiesis are indicated.

Structure, organisation and expression of the β globin genes

The different β-globin genes are structurally organised in a very similar way. The genes are relatively small, about 1500 bp and all exhibit a typical three-exon, two-intron structure. The exons are about 200 bp long, the first intron is roughly 100 bp and the second intron is about 900 bp long.

The promoters of all the β-globin genes contain a TATA-box, which is found ~ 30 bp upstream of the CAP-site. They also have similar, but not identical transcription factor binding-sites, for example the CCAAT-box, the CACC-box, and GATA-1 sites (Fig. 6). Antoniou and Grosveld (1990) showed in stably transformed cell lines that the CCAAT and CACC-box elements are required for induced and tissue-specific expression levels, in agreement with mutation data of the CACC-box derived from the study of thalassemia patients (Orkin et al., 1984; Treisman et al., 1985;). Deletion of the TATA-box resulted in a dramatic decrease of expression levels, suggesting that the LCR has an effect on the formation of the initiation complex. Extended analysis of the TATA-box region revealed that TFIID is the only functional complex involved in the transcriptional initiation of the β-globin gene. In a TATA-less mutant transcription initiation is not restricted to the CAP-site and several start sites were observed, indicating a possible role for the upstream binding elements in positioning the TFIID complex (Antoniou et al., 1995).

Epsilon globin

The ε-gene has a CCAAT-box at –82 bp and a CACC-box at–110 bp (Cao et al., 1989; Fig. 6). When transgenic mice were made with a construct containing only the ε-gene with flanking sequences it was transcriptionally inactive; suggesting that the promoter region alone is not sufficient for its activation (Shih et al., 1990).

Gamma globin

The γ genes have a duplicated CCAAT-box. The proximal CCAAT box is located at –88 bp and the distal box is located at position –115 bp. There is also a CACC-box at –145 bp (Fig. 6). Point mutations in these binding sites (e.g. at –117 or at –158) are known to cause a non-deletion type HPFH [the γ gene is not properly silenced after the switch from γ to β gene expression] (Superti-Furga et al., 1988). However, not only mutations in known transcription factor binding sites cause an HPFH phenotype, but mutations at -196 bp and –202 bp are also well-documented point mutations causing HPFH (reviewed in Poncz et al., 1988). To date, it is not clear how the γ genes are suppressed during the adult stage of development. It has been suggested that silencer or repressor proteins bind to the γ promoter region thereby silencing

the gene. This could explain the observed persistence of γ-gene expression upon mutations of a binding site, assuming that the mutation abolishes the binding of a silencer/repressor (Ronchi et al., 19). Although it remains unclear what causes the reactivation of the γ gene in the more upstream mutations, experiments have indicated that chromatin structure plays an important role in the repression of the γ genes (Pissard et al., 1996).
When the human γ-gene was introduced into transgenic mice it was found to be expressed in a tissue-specific manner and to follow the expression pattern of the embryonic mouse genes (Chada et al., 1986; Kollias et al., 1986), albeit that at very low levels.

Delta globin

The promoter of the δ-gene has a mutated CCAAT-box (first C->G) at –70 bp, and it has an imperfect CACC box at –90 bp. Recent studies have shown that the imperfect CCAAT and CACC-box in the δ promoter are the cause of low expression levels of this gene. Replacement of this –90 bp region with the proximal CCAAC-box of the β-gene induces the expression of a reporter gene in a transient expression assay to a level that is comparable to that of the β-promoter (Donze et al., 1996; Tang et al., 1997). Similar experiments that restored the CCAAT box also resulted in increased expression of the δ-gene (Tang et al., 1997).

Beta globin

The β-gene promoter has a single CCAAT-box, which is located at –75 bp and two CACC boxes: the distal CACC–box is located at position –105 bp and the proximal CACC-box at –90 bp (Fig. 6). Extensive studies on the rabbit and mouse β-globin promoters have shown that deletions upstream of the CACC elements do not affect the transcription of a reporter gene. A deletion of the proximal or the distal CACC-element, the CCAAT-box or the TATA-box results in significant reduction in gene expression (Dierks et al, 1981, 1983; Grosveld et al., 1981,1982). When the human β-globin gene was assayed for activity in transgenic mice, expression was found mainly in adult erythroid tissues. The expression paralleled that of the endogenous adult mouse globin genes (Magram et al., 1985; Townes et al., 1985; Kollias et al., 1986). However, the expression levels of the human β-globin gene in transgenic mice were also very low, suggesting that the cis-acting elements that confer tissue-specific expression were closely associated with globin genes but that an important regulatory element was missing (see below).

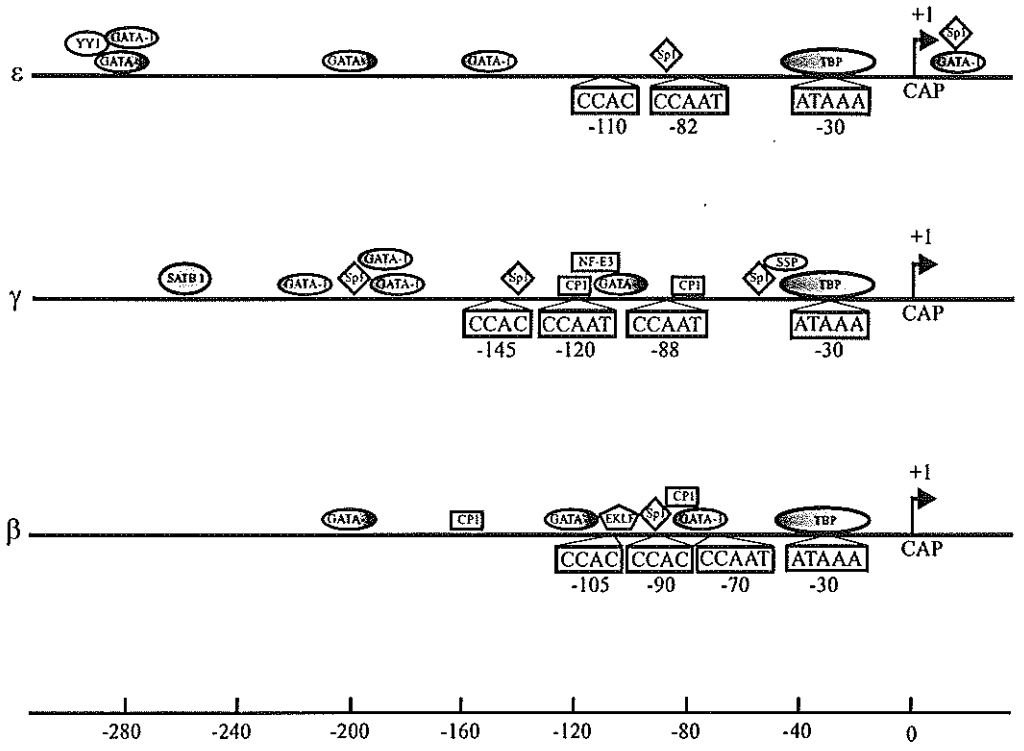


Figure 6. The promoters of the ε, γ and β-globin genes.
Shown are the locations of the conserved TATA, CCAAT and CCAC boxes and the binding-sites for the various transcription factors. The direction of the GATA motif is indicated by the shading of the ellipse.

Locus Control Region

Analysis of the Dutch γβ-thalassemia phenotype (van der Ploeg et al., 1980), in which thalassemia patients have a deletion of ~100kb that deletes the complete LCR-ε-γ-γ-δ region, leaving only the adult β-gene, revealed that the intact β-gene was inactivated due to *cis*-acting influences (Kiousis et al., 1983). Although it was shown that the β-gene was in an inactive chromatin conformation, it was not clear whether this cis-effect was due to the deletion of a positive regulatory region acting on the β-gene or to the addition of actively suppressing sequences that were translocated to the β gene. Studies using transgenic mice (see also section on gene-structure, Magram et al., 1985; Townes et al., 1985; Kollias et al., 1986) showed that closely linked cis-elements of the different globin genes were not able to give high levels of gene expression. Together these observations suggested that important regulatory elements were missing in both the γβ thalassemia patient and the transgenic mice studies. In an attempt to locate these possible regulatory elements of the human β-globin locus DNase I hypersensitive sites (HS) were mapped (Tuan et al, 1985; Forrester 1986; Grosveld et al., 1987) in different erythroid cell lines.

Minor and major hypersensitive sites were identified within the human β-globin locus. The minor hypersensitive sites (relatively high concentrations of DNase I were used to detect these) were located to the 5' regions of the different expressed globin genes.

Six erythroid cell specific and developmentally stable major hypersensitive sites were identified, of which 5 sites are located in a region of approximately 20 kb 5' to the ε gene. The sixth hypersensitive site was shown to be located at the 3' end of the β-globin locus, about 10 kb downstream of the β gene (3' HS1; Fig. 5).

The functional relevance of these hypersensitive sites became clear when mice were generated with a β-globin "minilocus". In this construct the 6 major hypersensitive sites were linked to the adult β-gene. Analysis of β-gene expression during development showed that the 5 hypersensitive sites from the 5'end of the locus conferred high level, tissue-specific, copy number dependent expression which was independent of the position of integration (Grosveld et al., 1987). This region was designated the Locus Control Region (LCR) (Fig. 5).

Properties of the individual hypersensitive sites of the β-globin LCR.

To investigate whether the individual hypersensitive sites have intrinsic transcription regulatory activity a number of different experiments were carried out in transgenic mice and in cell lines. It was shown that HS2, HS3 and HS4 are the most important hypersensitive sites of the LCR (Forrester et al., 1989; Collis et al., 1990; Fraser et al., 1993; Ellis et al., 1993,1996; Fig. 7).

The activity of the individual hypersensitive sites it was shown that the effective sequence of

the different sites resides in 200-300bp core-elements (Philipsen et al., 1990,1993; Talbot et al., 1990; Pruzina et al., 1991). These core-elements were shown to contain a number of binding-sites for DNA binding proteins / erythroid specific transcription factors (see below, Fig. 7). Interestingly, these binding sites closely resemble the motifs found in the promoters of the different globin genes.

HS1: In a naturally occurring deletion, called Spanish thalassemia, (Fig. 3; Driscoll et al., 1989; Forester et al., 1990) 35kb of sequence 5' to HS1 is deleted, leaving HS1 and the rest of the locus intact. The deletion results in a DNase I resistant configuration of a region approximately 200kb around the globin locus that is normally sensitive to DNase digestion. It was also shown that the replication timing of the globin locus changes from early S to late S phase(Aladjem et al., 1995). Analysis of transgenic mice in which HS1 is linked to a γγδβ cosmid showed that HS1 is not able to drive γ-gene expression in the fetal liver (Fraser et al., 1993), although it contributes significantly to the total expression of the locus as determined by deletions of HS1 from the mouse and human globin locus (Milot et al., chapter 3 and Trimborn et al., chapter 6).

HS2: This is the only site that shows classical enhancer activity in transient transfection assays in cell lines (Tuan et al., 1989). The major activity resides in a 300 bp core-element containing a number of binding sites for ubiquitous as well as erythroid-specific transcription factors (Talbot et al., 1990, 1991; see Fig. 7). The enhancer function has been attributed to the dimer AP-1 consensus binding site (Ney et al., 1990; Caterina et al., 1991; Talbot and Grosveld, 1991). However, a 1.5 kb HS2 fragment is not able to activate a linked β-gene when present as a single copy in transgenic mice. Activation is only achieved in multi-copy transgenic animals (Ellis et al., 1993; Ryan et al 1989). When HS2 was linked to a γγδβ cosmid it was shown that this site is able to contribute to the overall expression of the locus (Fraser et al., 1993).
Pawlik and Townes (1995) have suggested that the formation of HS2 is an intrinsic property and is not dependent on a linked globin gene. Injection of a 1.9 kb HS2 fragment into mice resulted in hypersensitive site formation in 9 out of the 10 lines examined. The integration sites of these transgenes were not analysed and it can not be excluded that some of them may have integrated near endogenous transcription units.

HS3: The core element (225bp) contains a triple repeat of GATA-1 binding sites and G-rich sequences (Philipsen et al., 1990, 1993). HS3 is the only site that can drive expression of both γ and β-globin genes in fetal liver (in a γγδβ cosmid, Fraser et al., 1993) and is the most active

site in the embryonic and fetal stages of development. HS3 consistently drives expression of single copy β-globin transgenes, suggesting a dominant chromatin-opening function for HS3 (Ellis et al., 1996).

HS4: The basic properties of HS4 are retained within a 280-bp core element (Pruzina et al., 1991; Lowrey et al., 1992), with similar transcription factor binding sites compared to those present in HS2 and HS3. When linked to the γγδβ cosmid, HS4 is very active in the adult stage driving β-globin expression. However, as a single copy transgene linked to the β-globin gene, it is not able to activate the linked β-globin gene (Ellis et al., 1996).

HS5: This hypersensitive site was originally described as a ubiquitous hypersensitive site (Tuan et al., 1985) with possible boundary or insulator activities (Li and Stamatoyannopoulos, 1994). Careful analysis using transgenic mice indicated that neither effects were reproducible. HS5 is a tissue-specific hypersensitive site that is only detectable in red cells and does not have any boundary or insulator activity (Zafarana et al., 1996). When HS5 is linked to a β-globin gene it is not able to enhance transcription of this gene or to shield it from position effects (Raguz, personal communication).

These experiments showed that although the hypersensitive sites share similar transcription factor binding-sites (Fig. 7), each site behaves quite differently when linked to a (set of) reporter gene(s) and suggests that the HS elements must interact with one another to obtain wild-type levels of globin gene expression. The distinct intrinsic properties of the individual hypersensitive sites became even more evident from studies in transgenic mice with YAC constructs in which HS3 was deleted or replaced by HS4 and *Vice Versa* (Bungert et al., 1995).

The deletion of HS3 or HS4 caused a dramatic reduction in the expression of the globin genes in all stages of development (see also chapter 3 of this thesis). In the transgenic mouse lines that have a duplicated HS4, there is a competitive increase in γ-expression at the expense of the ε-gene while β-gene expression is severely ablated. Duplication of HS3 did not result in an aberrant expression pattern, indicating that HS3 can complement the loss of HS4 in these mice. These data suggest that the LCR acts as an integral functional unit, and that the dissection of this complex leads to the disruption of its structure and function. Similar experiments, in which individual hypersensitive sites were deleted, showed that upon deletion of any one hypersensitive site in the LCR; the transgenes were no longer expressed independently of the integration site. The ability of the LCR to confer position-independent expression to linked genes was lost. DNA-FISH analysis revealed that those transgenes that

were integrated in or near the centromere showed a position effect (Milot et al., 1996), while mice with transgenes that were not integrated in or close to the centromere showed only minor changes in globin expression levels. The latter results were confirmed by Peterson et al. (1996), who also showed that deletion of HS2 or HS3 of the LCR did not lead to dramatic changes in expression of the different globin genes during development[5].

Experiments with the endogenous mouse β-globin LCR (deleting the complete HS2 or HS3) resulted in a reduction of expression levels of the adult β-globin genes of approximately 30 % (Fiering et al., 1995; Hug et al., 1996). .Deletion of HS1, HS4 or HS5 resulted in a decrease of 10% in expression levels of the adult β-genes (Trimborn et al., unpublished). Summarising these experiments it can be concluded that the LCR exerts its function as a holocomplex and it should be regarded as a single regulatory element.

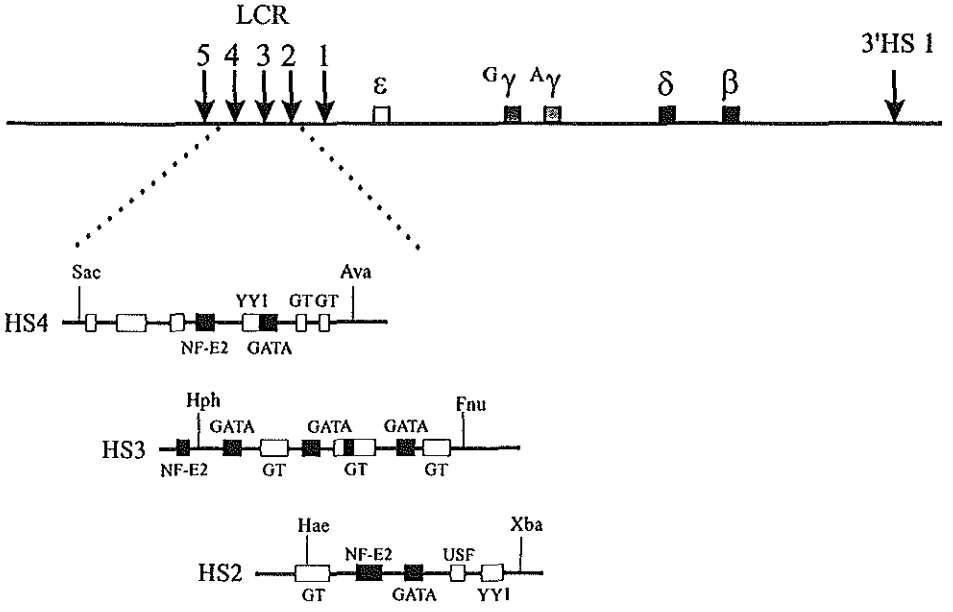


Figure 7. The human β-globin Locus Control Region (LCR).
The human β-globin locus is shown, indicated are the 6 DNase I hypersensitive sites (arrows), the 5' hypersensitive sites are confined to a region of ~20 kb, upstream of the ε-gene and are referred to as the LCR. The binding motifs of various proteins within individual hypersensitive sites 2, 3 and 4 are depicted.

[5] Note that neither Bungert et al. nor Peterson et al. determined the integration sites of their transgenes.

Minor regulatory regions in the human β-globin locus

A number of additional regulatory regions or elements have been described in the human β-globin locus. These elements are thought to play an important role in the developmental expression patterns of the genes.

ε-Silencer

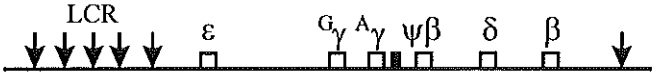
A region 5' to the ε-gene was identified as conferring for silencing activity on the ε-gene *in vitro* as well as *in vivo*. It was shown by transient transfection assays that a small region (-117 bp to –392 bp) silences a CAT reporter gene, irrespective of its orientation or location (Cao et al., 1989). It was also shown in transgenic mice, using a heterologous construct (HS1 and HS2 linked to the 5'-ε and the Aγ gene), that if this region is deleted, expression of ε is not silenced and is therefore no longer restricted to embryonic blood cells (Shih et al., 1993). In addition a deletion of an even smaller fragment in this region (Liu et al., 1997) made in a β-globin YAC suggested that this region is not only a silencer of ε, but is also required for its activation and surprisingly, is also required for γ-gene activation in the yolk sac.

Transgenic mice generated with constructs containing specific point mutations at potential binding site for nuclear proteins revealed that the transcription factors YY1 and GATA-1 are candidate factors responsible for the repression of the ε-gene in the adult stage of development (Raich et al., 1995; Li et al., 1997). The mechanism of this repression is not clear but from these results it was suggested that either YY1 is a direct repressor or that it hinders the binding of GATA-1 and thus preventing transcriptional activation. It was recently shown that YY1 associates with the histone-deacetylase Rpd3 (Yang et al., 1996). This indicates that modification of nucleosomal histones assembled on the ε-promoter may play an important role in the transcriptional regulation of the gene.

Stage selector element (SSE)

In the chicken β-globin locus a competition model was proposed to explain the observed developmental gene switching of the ε and β-genes (Choi and Engel, 1988). In this model the ε and the β genes compete for the same enhancer in definitive erythrocytes[6]. The SSE provides the β-gene with a competitive advantage over the ε-gene, resulting in the silencing of the ε-gene. In an analogy to this model, Jane et al. (1992) described a region in the γ-promoter (-53 to –35) that is suggested to be essential for the γ genes to compete effectively with the β-gene during early development. Furthermore a protein was characterised that binds to this region: -50γ or stage specific protein (SSP). The designation SSE is not a very accurate one as it is obvious that this region should be regarded merely as part of the γ-gene promoter where a number of different elements play a role in the switch from γ to β-gene expression. .
Analysis of transgenic mice with a deletion of this region has revealed a significant change in the γ to β mRNA ratios. The human β-gene, which is transcriptionally silent in cells of day 10 embryonic blood carrying the wild type γ-β minilocus construct, has become transcriptionally active in animals that have a mutated –50 region of the γ-gene. These results were confirmed by primary transcript *in situ* hybridisations using intron specific oligonucleotide probes (Ristaldi and Grosveld, in prep).

3'γ enhancer



The 3' $^{A}\gamma$-enhancer is located 400bp downstream of the poly-adenylation site of the $^{A}\gamma$-gene and is less than 750 bp in length (Bodine and Ley, 1987). Patient data indicated that large deletions in the human β-globin locus that deleted the region immediately 3'of the γ gene resulted in a δβ thalassemia. But patients with a similar deletion which did not include the 3'γ-enhancer showed an HPFH phenotype. These data suggest the presence of a possible regulatory element in this region.
Analysis of the 3'γ enhancer, by transient transfection assays, showed that it has classical enhancer activity being independent of its orientation and induces the expression of a linked reporter gene. Furthermore, this region also induced expression from heterologous promoters in erythroleukemia cells as well as in non-erythroid cells and it showed erythroid-specific DNase I hypersensitive sites. Footprint analysis of this region revealed that there are three possible binding sites for the erythroid transcription factor GATA-1, but sites for more

[6] Note that the β-gene is silent in primitive erythrocytes

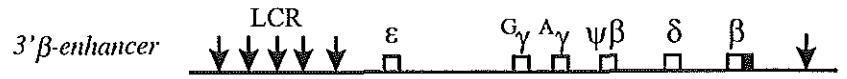
ubiquitously expressed transcription factors were also identified e.g. AP-2 and Sp-1 (Cunningham et al., 1994). In the same study it was shown that SATB1 binds to the 3' γ-region, a protein that was originally identified to bind to MAR/SARs (Jarman and Higgs, 1988; see below). The 3' γ region did not show up in earlier biochemical experiments that were designed to identify the MAR/SARs within the globin locus (see below), but it remains possible that this region is attached to the nuclear matrix/scaffold and therefore functions as a regulatory element in separating topological domains.

From experiments using micro-LCR-γ constructs with different, increasing deletions in the 3' γ sequences it was concluded that the 3'enhancer is necessary for position-independent expression of these transgenes (Stamatoyannopoulos et al., 1997). The results of these experiments are not easy to interpret since the lines analysed were not single copy and it is not clear what the significance of these data is in the context of the complete β-globin locus.

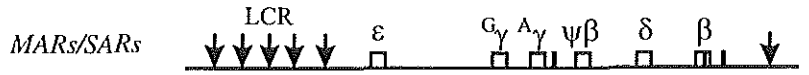
Several naturally occurring deletions suggested that the region 3'of the ψβ gene and ~3kb upstream of the δ gene might be of interest for further studies, since they result in different phenotypes with respect to γ-globin expression. The Italian HPFH-5 deletion, which removes 15 kb of sequence starting 3.2 kb upstream of δ to the 3β'enhancer causes an HPFH phenotype (Fig. 3). A similar deletion that removes the same sequence, but differs only in that the deletion starts 2.5kb upstream of δ results in a thalassemic phenotype (US Black, Fig. 3).

A 1.7 kb fragment (2.3 - 4.0 kb 5' to δ) was used in a transient transfection assay in MEL cells and showed an inhibitory effect on both the γ and the β-globin promoters but not on a non-globin promoter (Vitale et al., 1994).

Studies using transgenic mice made with YAC constructs that delete this 5'δ-region show that this transgene no longer exhibits position-independent expression. Two transgenic lines were obtained, DNA-FISH analysis have revealed that one transgene is integrated in the centromere and is therefore located in heterochromatin. The integration site of the second transgene is near the centromere (Grosveld, personal com.). Both lines are subject to classical PEV, as was shown using primary transcript *in situ* hybridisation (Milot et al., 1996), mRNA was only detected in a subset of erythrocytes that were transcriptionally active. A second phenomenon detected for both lines is the fact that the mRNA levels, analysed using S1-nuclease protection assays, were lower than the quantitated primary transcript signals. Experiments performed thus far always showed comparable levels of mRNA and number of transcriptionally active genes (Wijgerde et al., 1995; 1996; Dillon et al., 1997; Trimborn et al., this thesis) suggesting that in this case the number of polymerases per transcription unit has decreased.



The human β-globin gene, when introduced in transgenic mice, is expressed in a correct developmentally specific manner albeit at very low levels (Magram et al., 1985; Townes et al., 1985; Kollias 1986). Expression was not detectable during embryonic development but is turned on in the fetal liver and continues into adulthood. The presence of downstream elements in the chicken β-globin gene were indicative of the presence of a 3' regulatory element for the human β-globin gene. Kollias et al. (1986) showed that a γ-β hybrid gene is expressed in a tissue and developmentally specific manner, again suggesting the presence of a regulatory element. Experiments in transgenic mice and in MEL cells proved that the erythroid and developmentally specific activity resides in a ~1 kb region 3'of the poly-A site of the β-gene (Kollias et al., 1987; Trudel et al., 1987; Behringer et al., 1987; Antoniou et al., 1988). Transgenic mice made with YAC constructs that have a deletion of the 3' β-enhancer (Liu et al., 1997) showed that there is a significant decrease in β-gene expression in the fetal liver and the spleen. The effect was very specific and only on the β gene.



In the human β-globin locus three SAR/MAR sequences have been described, two based on the biochemical isolation protocols (Jarman and Higgs, 1988) the third via an indirect strategy (Cunningham et al., 1994). The first SAR/MAR is located in the intervening sequence (β-IVSII) of the β-gene, this region was also shown to have tissue-specific enhancer activity (Antoniou, 1988). The second SAR is located 3' of the β gene approximately 500 bp downstream of the 3'β-enhancer. The third SAR was described more recently, using affinity chromatography. A protein was isolated that associated with the 3'$^{A}\gamma$ enhancer region which turned out to be SATB1 (special A/T binding protein; Dickinson et al., 1992). SATB1 has been shown to bind selectively to SARs/MARs.

All three regions described above exhibit a typical SAR/MAR phenomenon; they co-localise with cis-regulatory elements. This has also been observed for several genes in *Drosophila* where it was called co-habitation (Gasser and Laemmli, 1986). However, the MAR/SARs that were identified in *Drosophila* never resided within a transcriptional unit and this raises the question whether the SAR shown in IVSII of the β-gene is a true SAR, all the more so because there has been much discussion about the different protocols used to identify SARs/MARs. A new protocol, using FISH analysis, described by Bickmore and Oghene (1996) is probably going to set the new standard of how SARs should be localised *in vivo*. In this particular experiment a 450 kb YAC was used to hybridise to an area around the PAX6 gene. In double hybridisation experiments using specific cosmids, it could be determined where the site of attachment to the nuclear scaffold was located.

Globin gene switching during development

The genes of the human β-globin locus are differentially expressed during development and this involves tight regulation of these genes to assure correct activation and repression at each stage. Using transgenic mice the mechanism of developmental switching in the human β-globin locus was studied. First of all it was shown that the cis-elements closely linked to the ε gene are able to repress the expression of this gene after the switch from embryonic to fetal erythropoiesis (Raich et al., 1990; Shih et al., 1990). Transgenic mice were made with a μLCR-ε construct and it was shown that the ε gene was only expressed in embryonic blood cells, with no significant expression of the gene in 14.5 day fetal liver. Secondly, experiments using μLCR-γ, μLCR-β and μLCR-γδβ constructs (Enver et al., 1990) showed that the μLCR-γ and μLCR-β did not show correct developmental expression. The γ-gene was expressed throughout development, in embryonic blood, but also in fetal liver until at least 18 days of gestation and was not silenced correctly in the fetal period. Data for the β-construct were very similar with β-gene expression observed early in embryonic erythrocytes and staying on during fetal development. Similar experiments by Dillon and Grosveld (1991) showed that if the 3' γ enhancer is included in the mini-LCR-γ construct, the γ gene is shut down correctly during fetal development, suggesting that the enhancer might be essential for autonomous silencing of the γ-genes. However, it is difficult to compare these two sets of data since there were also differences in the LCR-fragments [Dillon and Grosveld used a mini-LCR (1991) whereas Enver et al. used a μ-LCR[7] (Forrester et al., 1989)]. Results with the μLCR-γδβ constructs (Enver et al., 1990) demonstrated that by using this construct correct developmentally regulated globin gene switching was achieved and suggested that the fetal and adult genes compete for activation by the LCR, as was proposed for the chicken β-globin locus (Choi and Engel, 1988).

The competition model for the human β-globin locus was confirmed by using constructs with a different gene-order relative to the LCR (Hanscombe et al., 1991; Peterson et 1993). The results indicated that a combination of local control elements in conjunction with polar competition[8] regulates the developmental expression in the human β-globin locus. In other words the ε and the γ genes are silenced autonomously after the switch from embryonic to fetal erythropoiesis (Raich et al., 1990; Dillon et al., 1991), whereas the β gene is silenced in the

[7] The μ-LCR is a 2.5 kb cassette containing four of the erythroid-specific hypersensitive sites from the LCR (Forrester et al., 1989); the mini-LCR is a 22 kb LCR fragment containing all 5 hypersensitive sites (Grosveld et al., 1987).

[8] This implies that the relative position of a gene in relation to the LCR is important.

embryonic stage of development via a competition mechanism. Since the ε and γ genes can not be activated in the late fetal period, the β gene is now the only gene that can be expressed. In chapter 4 of this thesis additional evidence for the competition model is presented and the importance of the relative position of a gene in the locus is discussed (Dillon et al., 1997).

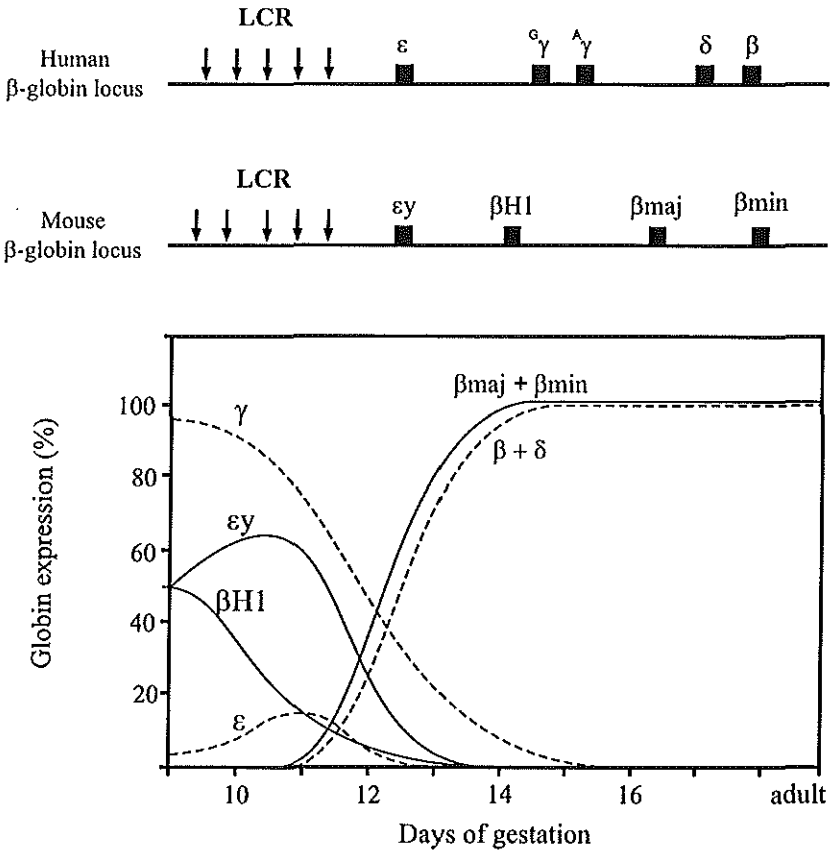


Figure 8. The organisation and developmental expression profiles of the human and the mouse β-globin genes. The graphs represent the mRNA levels of the individual human β-globin genes from a single copy transgenic mouse line (Line 72; Strouboulis et al., 1992a) compared with the mRNA levels of the endogenous mouse β-globin genes.

Mouse model to study developmental regulation of the human β-globin locus

It was clear from earlier experiments that the study of the multi-gene human β-globin locus would be much more complicated than a single gene-promoter study. It became evidently that a multi-gene locus has to be studied in conditions as close as possible to the endogenous situation. However, a major problem was the limited size of DNA that could be cloned and manipulated in bacteria. By linking together two cosmids spanning the 70 kb DNA region of the human β-globin locus, our lab first succeeded in generating single copy transgenic mice carrying the complete human β-globin locus (Strouboulis et al., 1992b). The developmental expression pattern of these single copy, human β-globin locus mice was examined. Yolk-sac RNA from 8.5, 10.5 and 11.5 days and fetal liver RNA from day 12.5, 14.5 and 16.5 was analysed using S1 nuclease protection assays (Strouboulis et al., 1992a). The results of the individual human globin genes compared to the endogenous mouse genes showed that during embryonic erythropoiesis, the two γ genes are most abundantly expressed, approximately five-fold higher than the ε-gene (Strouboulis et al., 1992a; Fig. 8). Expression of the ε-gene was not detected in fetal liver whereas low levels of γ-gene expression were detectable until 16.5 days of gestation. Expression of the β-gene was detected starting at 12.5 day when the fetal liver is the main site of erythropoiesis. These data showed that the expression of the human genes in transgenic mice mimics the developmental expression pattern of the human β-globin locus closely and that therefore these mice provide an excellent model to study the molecular mechanism of gene regulation in a multi-gene locus. Recently it has been demonstrated that with the use of yeast artificial chromosomes (YAC) it is also possible to manipulate large sequences of DNA (ranging from 100-500 kb) and to generate transgenic mice (Peterson et al., 1993; Gaensler et al., 1993). These mice with YACs containing either 150 or 250 kb of DNA including the human β-globin locus showed an expression pattern similar to that obtained with the linked-cosmid strategy (Strouboulis et al., 1992a/b).

Erythroid Specific Transcription Factors

Several transcription factors that are involved in the regulation of the β-globin locus have been identified; for example AP1, EKLF, GATA-1, NF-E2, NF-E6, SP1 and YY1. These *trans*-acting factors were not only present in the promoters of the five different β-globin genes (Fig. 6), but also in the hypersensitive sites of the LCR (Fig. 7). A number of these factors is discussed more extensively below.

EKLF

The erythroid Krüppel like factor (EKLF), is a zinc-finger protein of approximately 38 kDa. It was cloned using a subtractive hybridisation technique in a search for new erythroid-specific transcription factors (Miller and Bieker, 1993). EKLF contains 3 zinc-fingers and was shown to bind to the CACC-box, which can be found in all globin promoters (as shown above). EKLF is an erythroid-specific transcription factor that is only expressed in erythroid tissues. Knocking-out EKLF in ES cells yielded mice that developed a fatal anaemia in the early fetal liver (Nuez et al., 1995; Perkins et al., 1995). From these experiments it was concluded that EKLF is indispensable for the final steps of definitive erythropoiesis.

It was shown that EKLF activates the adult β-globin gene to a much greater extent than it does the fetal γ-genes in cultured cells (Donze et al., 1995). The generation of mice containing a single copy of the human β-globin locus (Strouboulis et al., 1992a) as well as an EKLF deficiency (+/- or -/-) enabled the study of the effect of EKLF on the developmental regulation of the human globin genes (Wijgerde et al., 1996). It was found that without EKLF there is no β-gene expression at all, this coincides with the disappearance of the 5'HS of the β-promoter. It was suggested that this might result from changes in the interaction between the LCR and the upstream β-elements (Wijgerde et al., 1996). Furthermore, it was shown that a decrease in EKLF (in heterozygous animals) leads to a higher ratio of γ to β transcription during the switching period. Conversely, overexpression of EKLF results in a lower γ to β ratio of transcription (Tewari et al., 1998). Thus, these experiments suggested that EKLF might help to complete the switch from γ to β-gene transcription by ensuring strong interaction between the β-gene and the LCR (Wijgerde et al., 1996; Perkins et al., 1996).

Experiments in which μ-LCR and HS3 constructs linked to a heterologous promoter were analysed in an EKLF -/- background revealed that EKLF is also required to drive expression of a heterologous promoter linked to HS3 (Tewari et al., 1998). This suggests that EKLF indeed interacts with HS3 *in vivo*. Furthermore, in another *in vivo* experiment, using transgenic mice in which the EKLF binding site of the 5' HS3 was mutated, it was shown that expression of the linked β-gene was abolished. Concomitantly matching amino acid changes in the zinc-fingers of either Sp1 or EKLF polypeptides with changes in their respective recognition sites, showed that these changes indeed resulted in specific binding of the changed protein to the mutant binding site. More interesting was the result that only in combination with the changed EKLF protein the expression of the β-gene was restored demonstrating that EKLF activates HS3 by direct interaction (Gillemans et al., in press).

GATA-1

The DNA-binding protein GATA-1 contains two zinc-fingers and is abundantly expressed in erythrocytes and megakaryocytic cells. The most important GATA consensus sequence (T/A)GATA(A/G) is present in the cis-elements of nearly all genes expressed in the erythroid lineage including all β-globin gene promoters (Philipsen et al., 1990) and the GATA-1 promoter itself (Nicolis et al., 1991). Its essential role in erythroid development has been established by gene-targeting studies (Pevny et al., 1991), loss of GATA-1 results in apoptosis. Loss of another family member, GATA-2, leads to a decrease in hematopoietic precursors (Tsai et al., 1994) indicating that this protein is essential in early development, whereas overexpression of an estrogen inducible GATA-2 protein results in proliferation of precursors and inhibition of differentiation (Briegel et al., 1994).

These data resulted in the following model for erythroid differentiation (Weiss and Orkin, 1995): GATA-2 is necessary and sufficient for the proliferation and differentiation of the erythroid lineage up to the proerythroblast stage. Subsequently, GATA-1 is activated by an unknown mechanism, leading to the down-regulation of GATA-2. It has been suggested that GATA-1 levels increase through a positive autoregulatory loop[9], thus reaching levels sufficiently high for the activation of erythroid-specific genes. However, studies in MEL cells have shown that overexpression of GATA-1 (~3 fold) gives the opposite result. Potential target genes are not induced and MEL cells are not directed towards differentiation but start to proliferate (Whyatt et al., 1997). Recent experiments show that GATA-1 acts directly on the cell cycle, possibly by regulating cyclin A2 which regulates phosphorylation of the Retinoblastoma protein (Rb). Interestingly erythropoiesis is also disrupted in the Rb knock-out mouse and its phenotype is very similar to the GATA-1 knock-out mouse (Maandag et al., 1994).

NF-E2

The transcription factor NF-E2 is a heterodimer consisting of an erythroid-specific subunit p45 and a ubiquitous DNA-binding subunit called p18-maf (Andrews et al., 1993). Both subunits are basic leucine zipper proteins, the heterodimer binds to tandem AP1 binding-sites (Mignotte et al., 1989; Andrews et al. 1993). Binding sites are found throughout the β-LCR (Fig. 7). *In vitro* studies, using reconstituted chromatin, have shown that NF-E2 can disrupt the nucleosome structure of chromatin (Armstrong and Emerson, 1996).

[9] Note that a consensus GATA binding site is present in the promoter of GATA-1.

Furthermore, GATA-1 is only able to bind to the reconstituted DNA if both subunits of NF-E2 are added. Thus it was suggested that NF-E2 facilitates the binding of GATA-1 to the DNA by replacing the nucleosomes on the template. Recent biochemical data have shown that *in vitro* NF-E2 might interact with the transcription machinery via a direct interaction with TAF-130 (Amrolia et al., 1997).

Disruption of the p45 subunit of NF-E2 in mice revealed only a very mild effect on globin chain synthesis. Thus, the regulation of globin gene transcription *in vivo* is not dependent on p45NF-E2. When the p18 subunit was disrupted null mutant animals developed normally and the disruption did not have any effect on globin expression, indicating that p18NF-E2 is not an essential component of the heterodimer (Kotkow and Orkin, 1996).

Replication of the human β-globin locus

Studies in MEL cells have shown that the β-globin locus is replicated in early S phase during the cell cycle (Epner et al., 1981). More extensive experiments, looking at the direction of replication, showed that the origin of replication is located in a 2 Kb element between the δ and β gene (Kitsberg et al., 1993). Analysis of a cell line of a Lepore patient, which has an 8 Kb deletion spanning this δ-β region showed that the replication becomes unidirectional (Kitsberg et al., 1993) and suggests that it originates from another initiation region (IR). Recently it was shown that in a cell-line of a Hispanic thalassemia patient which has a deletion of 25 Kb 5' to HS1 (Fig. 3), the replication fork changes direction (Aladjem et al., 1995). This change in direction was attributed to the deletion of LCR sequences. However, the authors do not consider the possibility that due to the loss of the LCR the chromatin conformation of the locus changes, as indicated by the change in DNase I digestion (see also section "properties of individual hypersensitive sites"),. Since it is known that heterochromatin replicates late during S-phase, it could be that the globin locus has been packaged into heterochromatin.

Mouse β-globin locus

The developmental expression pattern of the mouse locus was studied and compared with the expression of the human locus in transgenic mice. The organisation of the mouse β-globin locus is very similar to the human locus and it is therefore important to know whether the activation of the genes is regulated in a similar way; i.e. if the regulation of the β-globin genes is conserved during evolution (see chapter 6).

The murine β-globin locus consists of 4 functional genes: εy, βH1, βmajor and βminor and is located on chromosome 7 (Fig. 8). In analogy to the human locus a locus control region was identified upstream of the εy gene. The mouse β-LCR also consists of 5 hypersensitive sites that are erythroid specific and developmentally stable. The core elements of the hypersensitive sites were shown to be highly homologous to their human counterparts not only at the level of structure and organisation but also at the level of DNA sequence where a high degree of conservation is observed (Moon and Ley, 1990; Hug et al., 1992; Jimenez et al., 1992; Enver et al., 1994; Fiering, unpublished).

The genes are expressed during development in the same order as they appear in the locus with εy and βH1 expression restricted to embryonic blood. The βmaj and βmin genes are expressed already in the embryonic period albeit at low levels and are the only genes expressed in the fetal liver (Whitelaw et al., 1990).

Analysis on single cell level, using the primary transcript *in situ* hybridisation technique suggests that the regulation of the mouse β-globin locus is very similar to the human β-globin locus, in that the LCR activates only one gene at a time by a looping mechanism (Trimborn et al., this thesis).

Models of LCR mediated gene-activation in the human β-globin locus

Scanning model

Three models have been proposed to explain the mechanism of multi-gene activation in the human β-globin locus (Fig. 8). All models have in common the implication that the LCR provides an "open" or activateable chromatin conformation. The mechanism of gene activation is different for the individual models. The scanning model suggests that a "transcription activation machinery"[10] or chromatin structure, spreads along the DNA in a linear fashion until it encounters an activateable gene (Tuan et al., 1992; Herendeen, 1992). Another version of the tracking model envisages that the LCR or more specifically, HS2 might function as an entry site for polymerases, which again would traverse the DNA in a linear fashion, to activate the first gene encountered (Kong et al., 1997.) This scanning model explains the polar competition, observed for the genes in the globin locus. The blocking or attenuation of such a "transcription activation machinery" complex by the first gene that it encounters results in an advantage of an LCR-proximal gene over a more distally located gene.

Random activation model

The random activation model suggests that the β-globin locus is in an activateable chromatin environment in a red blood cell. The individual genes are activated at random, facilitated by stage-specific transcription factors that can access the different genes (Martin et al., 1996). A change in the transcription factor environment together with the presence of gene-specific silencers/repressors would result in a progressive increase in adult β-transcription. To explain the polar competition evident in the globin locus an additional aspect in this model involves transcriptional interference. This is thought to cause the disadvantage of a distal gene over a more proximal gene. The mechanism of such interference is unknown, but may involve transmission of topological change, downstream of a transcription complex, or compartmentalisation within the nucleus.

Looping model

The suggestion of the looping model (Hanscombe et al., 1991) was based on the work of Ptashne (1986), who showed that looping can occur over small distances in a prokaryotic

[10] This transcription machinery could be envisioned as a polymerase complex either in association with the LCR or on its own which tracks along the DNA in an unidirectional fashion.

system and also on the work of the chicken β-globin enhancer (Choi and Engel, 1988). The model proposes that the LCR forms dynamic, direct chromatin interactions with regulatory elements of a single gene looping out the intervening DNA, thus activating that particular gene. This model explains the observed polar competition within the β-globin locus; a gene located closer to the LCR has a higher chance of interaction with the LCR then a gene located further away (Hanscombe et al., 1991; Dillon et al., 1997). The model implies that there are two important parameters in the activation of the different human β-globin genes: the frequency of interaction between the LCR and a gene and the time that this interaction lasts (i.e. the stability of the interaction).

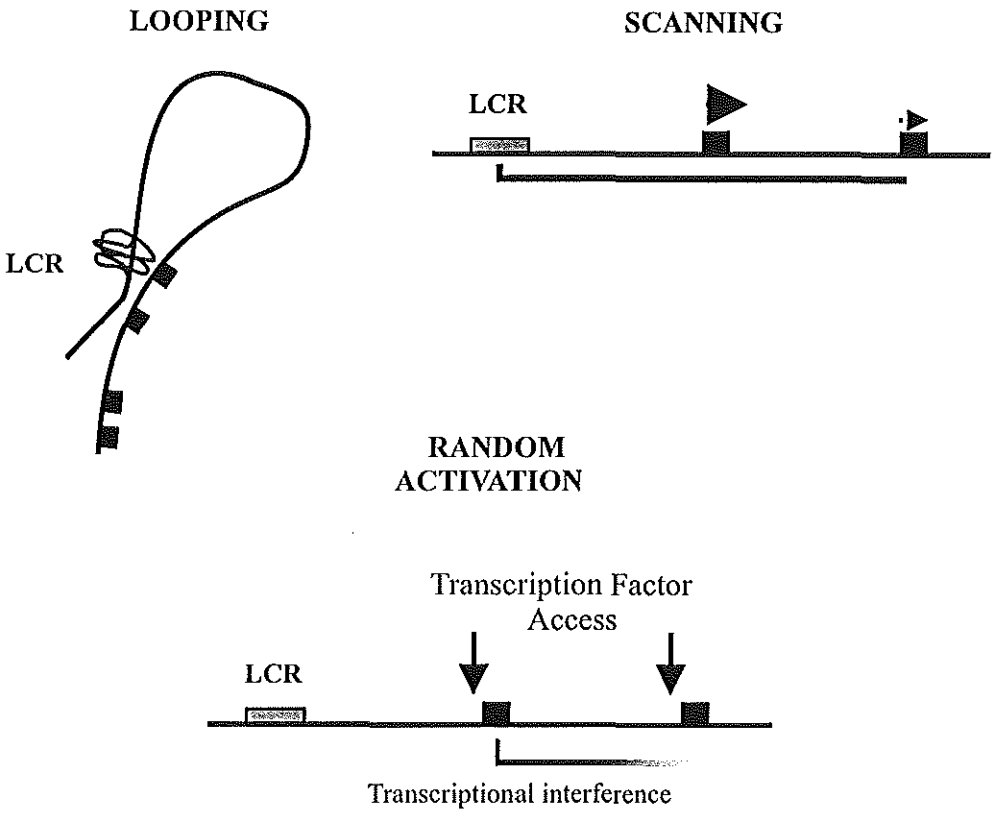


Figure 9. Models that have been proposed to explain LCR-mediated transcriptional activation of the human β-globin genes during developmen.

Primary transcript *in situ* hybridisation.

Previously it was shown by immunofluorescence, using antibodies that detect the different polypeptides, that the γ and β-genes are co-expressed in 16 day fetal liver cells (Fraser et al., 1993). This clearly showed that mouse erythrocytes are capable of expressing two genes of the same multi-gene locus. From these experiments it was not clear whether the LCR was committed in activating only one gene on the same chromosome or whether it could activate all developmentally appropriate genes. The first experiment to address this question took advantage of the fact that transcriptionally active genes have a chromatin conformation that is sensitive to nuclease digestion and restriction enzyme activity (Bresnick and Felsenfeld, 1994). It was shown that K562 cells (cells that express the γ genes) indeed gave a clear restriction fragment upon Apa I digestion as opposed to HeLa cells in which globin genes are not expressed. This result led the authors to conclude that the LCR was capable of sharing its activity among genes of the same locus. However, there is no evidence to suggest that hypersensitivity is a valid argument for actual gene transcription; in particular since hypersensitivity can be present prior to transcription (Blom van Assendelft et al., 1989; Jimenez et al., 1992). More informative results were obtained with experiments using fluorescent *in situ* hybridisation to detect primary transcripts of the individual globin genes (Wijgerde et al., 1995). Oligonucleotides (oligos) were designed to hybridise to the intron sequences of the nascent transcripts coming from a specific globin gene, allowing the detection of active globin transcription within a single red cell.

Fetal liver cells of a homozygous transgenic mouse containing a single copy of the human β-globin locus (Strouboulis et al., 1992a) were analysed using the primary transcript *in situ* hybridisation technique. It was shown that upon hybridisation with oligos specific for the γ and β primary transcripts, every possible combination of transcriptional foci was observed in erythrocytes (Wijgerde et al., 1995), i.e. one nucleus would show two β-gene spots, another would show a γ and a β spot on the different chromosomes and yet another one would show a β signal and a γ and β signal (double signal.) on the other chromosome Quantitation of the different combinations of transcription foci in 12.5 day fetal liver cells showed that approximately 85% of the foci were single signals with a γ versus β ratio of 8:2. This was the first indication that the LCR most likely activates only one gene at any time, reasoning that if the LCR would be capable to activate more than one gene at a time we would expect to see a much higher percentage of double signals on one chromosome, than the 15% observed. The question remained as to whether one locus first transcribes the γ-gene and subsequently (later in differentiation) switches to activate the β-gene or whether this activation of the individual genes was a more dynamic process (almost all 16 day fetal liver cells have γ and β proteins,

Fraser et al., 1993) in which the different genes can be activated randomly. The answer to this question came from an experiment in which fetal liver cells were hybridised with oligos specific for γ and β primary transcripts and in addition oligos that hybridised to the β mRNA (exon-specific oligos). The authors showed that a subset of cells showed transcription of the γ gene in a cell that previously had been transcribing the β-gene, as indicated by a halo of β mRNA in the cytoplasm (Wijgerde et al., 1995). This result suggested that the activation of the human β-globin genes along the locus is not by gradual progression whereby the most 5' gene is activated first, followed by the activation of the more 3' located genes. The results argued that the activation of individual β-globin genes is a dynamic process in which the genes can be alternately activated.

Chapter III

Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription.

Cell 1996 87: 105-114.

Heterochromatin Effects on the Frequency and Duration of LCR-Mediated Gene Transcription

Eric Milot,* John Strouboulis,†‡ Tolleiv Trimborn,* Mark Wijgerde,* Ernie de Boer,* An Langeveld,* Kian Tan-Un,§ Wilma Vergeer,*‖ Nikos Yannoutsos,* Frank Grosveld,* and Peter Fraser*
*Erasmus University
Medical Genetics Center, Department of Cell Biology
3000 DR Rotterdam
The Netherlands
†National Institute for Medical Research
Department of Gene Structure and Expression
The Ridgeway, Mill Hill
London NW7 1AA
United Kingdom

Summary

Locus control regions (LCRs) are responsible for initiating and maintaining a stable tissue-specific open chromatin structure of a locus. In transgenic mice, LCRs confer high level expression on linked genes independent of position in the mouse genome. Here we show that an incomplete LCR loses this property when integrated into heterochromatic regions. Two disruption mechanisms were observed. One is classical position-effect variegation, resulting in continuous transcription in a clonal subpopulation of cells. The other is a novel mechanism resulting in intermittent gene transcription in all cells. We conclude that only a complete LCR fully overcomes heterochromatin silencing and that it controls the level of transcription by ensuring activity in all cells at all times rather than directly controlling the rate of transcription.

Introduction

Nuclear processes such as transcription, replication, and recombination can be affected by chromatin structure (e.g., Pillus and Grunstein, 1995). The genome can be roughly divided in two major chromatin states based on cytological observation (Heintz, 1928; Eissenberg et al., 1995). Euchromatin appears decondensed in interphase, whereas heterochromatin is highly condensed throughout the cell cycle. Euchromatic regions that contain most of the genes and unique sequences replicate early in S phase. Heterochromatin replicates late in S phase (Holmquist, 1987) and contains middle and highly repetitive sequences (John and Miklos, 1979). The largest regions of heterochromatin appear near the centromeres and in the inactivated X chromosome. Studies in Drosophila and yeast have shown that heterochromatin can have effects on juxtaposed euchromatic regions, suggesting that chromatin condensation can spread from heterochromatin into regions of euchromatin (Wilson et al., 1990). Stochastic heterochromatinization of juxtaposed euchromatic regions, which results in stably inherited gene silencing in a clonal subpopulation of cells, is known as position-effect variegation (PEV; for review see Karpen, 1994). Evidence that PEV indeed results from heterochromatinization has come from genetic studies of specific genes that encode proteins involved in heterochromatin formation and have been shown to enhance or suppress these position effects (Pirrotta, 1995; Orlando and Paro, 1995). Although the dynamics of formation of the open and closed states of a locus are still obscure, it is clear that some genomic elements and nuclear factors can influence the chromatin configuration. In mammals, position effects that result in the stable silencing of genes in a clonal subpopulation of cells have also been observed (Elliot et al., 1995; Festenstein et al., 1996).

Another type of position effect frequently observed in transgenic experiments is thought to be due to the action of regulatory elements at the site of integration. This can result in both positive and negative effects on the transgene. Alternatively, the regulatory elements on the transgenic construct may interact with a gene located at the site of integration and result in a lower level of transgene expression (see Milot et al., 1996). These non-PEV effects result in a change in the level of expression in all of the expressing cells.

An increasing number of gene loci and gene clusters have been shown to contain *cis* elements known as locus control regions (LCRs), which are thought to be responsible for initiating or maintaining (or both) a cell type–specific open chromatin structure within a specified domain. The most thoroughly characterized LCR is that of the human β-globin locus (Grosveld et al., 1993). It consists of a series of five DNase I hypersensitive site (HS) regions located upstream of the globin cluster (Figure 1) each containing a number of binding sites for ubiquitous and erythroid-specific DNA-binding factors. Expression of β-globin transgenes without an LCR were found to be low, variable, and dependent on the position of integration in the host genome. Inclusion of LCR sequences in β-globin transgene constructs led to reliable, high level, copy number–dependent expression irrespective of position in the host genome (Grosveld et al., 1987). This is in agreement with the study of patients with deletions of the LCR. The globin genes in the otherwise normal locus were found to be insensitive to DNase I, transcriptionally inactive (Kioussis et al., 1983; Forrester et al., 1990), and replicated late in S phase (Aladjem et al., 1995).

The human β-globin LCR is thought to overcome position effects through a dominant positive activity rather than through an insulator function (Schedl and Grosveld, 1995). Recent studies demonstrate that the individual globin genes compete for LCR function and that the LCR activates only one gene at a time (Wijgerde et al., 1995), probably through direct interaction between the LCR and the gene via DNA looping (Dillon et al., submitted). These results suggested that the relative levels of

‡Present address: Laboratory of Molecular Embryology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892.
§Present address: School of Professional and Continuing Education, University of Hong Kong, Hong Kong.
‖Present address: Department of Microbiology and Biochemistry, University of Orange Free State, Bloemfontein 9300, South Africa.

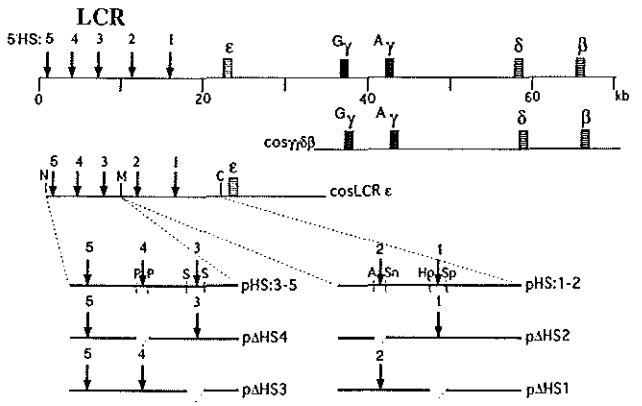


Figure 1. Construction of 70 kb Globin Locus LCR Deletion Mutants

The top line shows the 70 kb human globin locus construct. The individual HS regions of the LCR (5'HS:1–5) are designated with arrows and the globin genes with boxes. The two cosmids (cos LCR ε and cos γγδβ; Strouboulis et al., 1992a, 1992b) used to create the 70 kb construct are shown, as are the positions of restriction sites used to make the LCR plasmid subclones shown at the bottom. Individual HS deletions were created in the plasmids pHS:3–5 and pHS:1–2 using the restriction sites shown. Individual HS deletion plasmids were incorporated into cos LCR ε, which was then linked to cos γγδβ to create the injection constructs as described by Strouboulis et al. (1992b). Restriction sites are as follows: N, NotI; M, MluI; C, ClaI; P, PmtI; S, SauI; A, ApaLI; Sn, SnaBI; Hp, HpaI; Sp, SpeI.

gene expression in an active locus is (at least in part) determined by the duration of the interaction of the LCR with each of the genes. In this paper, we report the results obtained with transgenic mice containing the entire globin locus with LCR HS deletions. The results show that deletion of an individual HS from the LCR leads to severely reduced expression of the globin genes in some mouse lines. In situ hybridization analysis reveals that this loss of position independence is caused by integration at or near the centromere, resulting in pancellular and heterocellular position effects. We conclude that the LCR must be complete to overcome heterochromatin silencing and that it does so by ensuring that the locus is active all of the time in all of the (red) cells.

Results

Generation of LCR Deletion Transgenic Mice

The preparation of 70 kb LCR deletion DNA fragments for microinjection was carried out according to the cosmid linking method of Strouboulis et al. (1992a, 1992b). Four different locus constructs were produced with small deletions of HS:1, HS:2, HS:3, or HS:4 (Figure 1). The constructs were microinjected into the pronuclei of fertilized mouse eggs to produce transgenic founder mice. Tail DNA from founder mice was analyzed on Southern blots by probing with the entire 70 kb locus construct (data not shown). Those founder mice, which contained all of the EcoRI fragments characteristic of the entire locus, were selected for further breeding. F1 and F2 mice were similarly analyzed, and, in addition, end fragment, junction fragment, and internal probes were used to verify the structure and estimate the copy number. Transgenic line 72 was used as the control. This line contains a single copy of the human globin locus and expresses the human β-globin transgene at a level similar to that of the endogenous mouse globin genes (Strouboulis et al., 1992a). The mouse Thy-1 gene was used as an internal loading control. Copy numbers ranged from 1 to approximately 12 copies.

We prepared RNA from F2 transgenic 10.5 day whole embryos, 12.5, 14.5, and 16.5 day fetal livers, and adult blood and determined the level of mRNA by S1 nuclease protection assays. Each sample was assayed with the same mixture of six probes for mouse, εγ, βh1, β-major, and human, ε, γ, and β (see Experimental Procedures). S1 analyses are shown in Figure 2 for one transgenic line from each construct as well as transgenic line 72, which contains the full locus as a single integrated copy. The intensity of each S1 protected band was corrected for probe-specific activity and copy number and is presented in Figure 3 as a percentage of endogenous globin gene expression.

HS Deletions Affect Globin Gene Expression

The two HS:1 deletion lines (Δ1A, containing 1 copy of the human locus, and Δ1B, containing 2 copies) show widely varying levels of expression of the globin genes per copy, indicating a loss of copy number dependence. Δ1A globin gene expression is very low, ranging from 2%–10% per copy, whereas Δ1B expresses near normal levels when compared with line 72 (Figures 2 and 3). These results indicate that deletion of HS:1 may result in expression that is sensitive to the position of integration in the mouse genome. Δ1B shows the level of expression that would be expected on the basis of the previous data obtained for HS:1. HS:1 itself is known to have very little if any transcriptional activity (Collis et al., 1990; Fraser et al., 1990), and a patient with a deletion of HS:1 from the LCR has normal levels of β-globin expression (Kulozik et al., 1991). More interesting is the result obtained for the Δ1A line, which suggests that 5'HS:1 is nevertheless an essential component of the LCR to provide position-independent expression of the globin locus when taken out of the context of its normal chromosomal position (see below).

Based on previous data with the individual HS:2 (Talbot et al., 1989, 1990; Curtin et al., 1989; Ryan et al., 1989; Fraser et al., 1990, 1993; Morley et al., 1992) and the deletion of the mouse β-major globin HS:2 by homologous recombination (Fiering et al., 1995), we expected that the deletion of HS:2 would result in a mild (20%–30%) reduction in transcription of the genes in the locus. However, the four lines containing the HS:2 deletion construct (Δ2A, 8 copies; Δ2B, 3 copies; Δ2C, 1 copy;

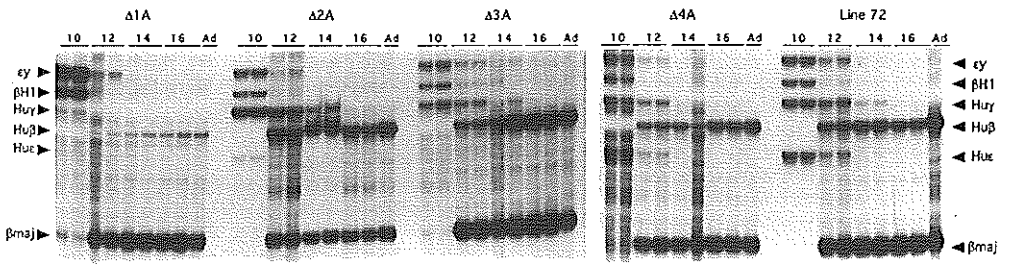


Figure 2. S1 Nuclease Protection Assays

S1 nuclease protection assays are shown for one transgenic line for each of the deletion constructs (Δ1A, 1 copy; Δ2A, 8 copies; Δ3A, 4 copies; Δ4A, 1 copy) used and line 72 (1 copy). RNA was prepared from transgenic 10 day whole embryos, 12, 14, and 16 day fetal livers, and adult blood as described in Experimental Procedures. Two samples for each timepoint except adult blood are shown. All samples were assayed with the same mixture of six radiolabeled probes for mouse εy, βH1, and β-major and human ε, γ, and β; the ratio of probe-specific activities was 0.8:0.8:0.8:3.3:1:1, respectively. The position of the protected fragments for each probe are indicated.

and Δ2D, 1 copy) show varying levels of expression per copy, indicating a loss of copy number dependence and hence position dependence (Figures 2 and 3). The most striking examples of position sensitivity of the Δ2 lines are Δ2C and Δ2D because they show a different effect for different genes in the locus. Δ2D shows appreciable levels of γ-globin expression in the embryonic period, but very low levels of human β-globin in the fetal liver and adult stages. In contrast, transgene expression in the Δ2C line is detectable only in the adult, where human β-globin expression is approximately 6% of the mouse β-major globin. The other two Δ2 lines (Δ2A and Δ2B) show a reduction in expression of all the human globin genes at all stages. Hence, HS:2 also appears to be crucial for the function of the LCR when integrated at another position in the genome (see below).

The transgenic mice containing the HS:3 deletion construct (Δ3A, 4 copies; Δ3B, 4 copies; Δ3C, 3 copies; and

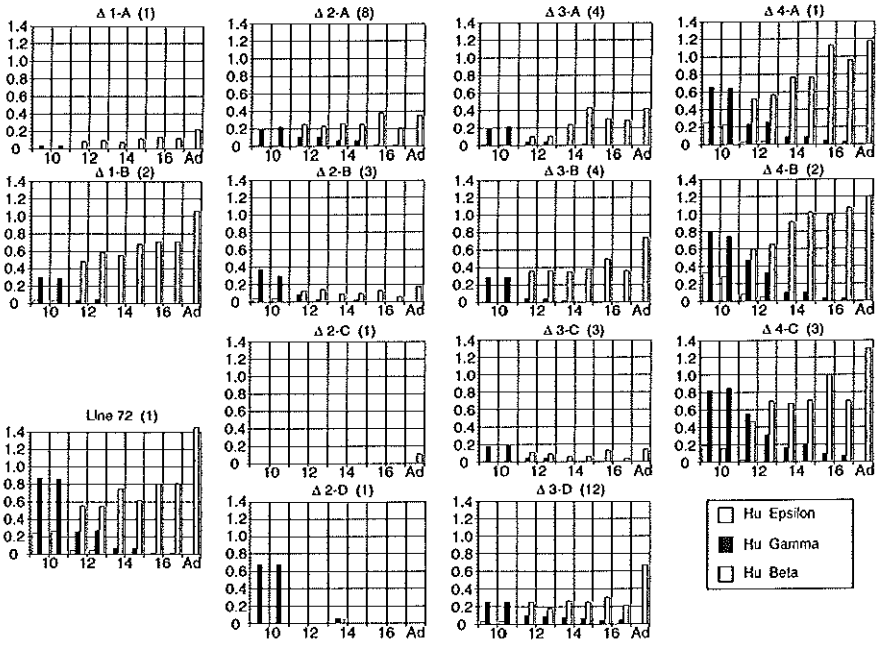


Figure 3. Quantitation of S1 Assays for All Transgenic Lines

S1 assays were quantitated by phosphorimage analysis and corrected for probe-specific activity and copy number of the transgenic line. Copy numbers are indicated in parentheses. The values shown are the per copy levels of expression of the human genes as a ratio of the endogenous mouse globin genes (human gene/copy number/[εy + βH1 + β-major]).

Δ3D, 12 copies) were expected to show at least the same or greater reduction in expression, in particular of the embryonic and fetal genes (Fraser et al., 1990, 1993; Philipsen et al., 1990, 1993; Hug et al., 1996). Indeed, the results show a general reduction in expression of all the human genes, but also a lack of copy number dependence (Figures 2 and 3). Δ3C in particular is expressing at less than half the level per copy of the other Δ3 lines. This suggests that deletion of HS:3 from the LCR also causes susceptibility to position effects (see below).

Three lines contain the HS:4 deletion construct (Δ4A, Δ4B, and Δ4C) with copy numbers of 1, 2, and 3, respectively. These lines appear to express the globin genes in a position-independent and copy number–dependent manner, as the level of expression per copy is similar to line 72.

Chromosomal Position of Transgenes

The results of the S1 analysis showed that the LCR deletion transgenic lines express the human globin genes at lower levels than expected. In addition, the levels of expression do not correlate with the number of copies integrated in the mouse genome. In particular, transgenic lines Δ1A, Δ2B, Δ2C, Δ2D, and Δ3C show extremely low levels of expression per copy when compared with line 72 or the other deletion lines. Studies in Drosophila and yeast have shown that chromosomal position can influence gene expression. To determine whether reduced gene expression in these mice can be correlated with chromosomal position, we performed fluorescence in situ hybridization (FISH) analysis on metaphase chromosomes from ten LCR deletion transgenic lines (lines Δ1B and Δ2D were lost before FISH analysis; Δ3D was not done). All four of the very low expressing lines (Δ1A, Δ2B, Δ2C, and Δ3C; Figure 3) were integrated at or near the centromere (Figure 4A). This suggests that pericentromeric integration of an incomplete LCR construct results in severely reduced gene expression. Such a position effect may be caused by heterochromatin spreading as is thought to occur in PEV. Since FISH on metaphase chromosomes is a crude method for determining the precise chromosomal location of a transgene, we refined these results by using a second probe for the mouse α-satellite repeat (Figure 4B), which is located primarily in centromeres. All four lines Δ1A, Δ2B, Δ2C, and Δ3C are integrated very close to or contacting the α-satellite DNA, whereas the next closest integrant (Δ2A) is clearly separate (Figure 4B).

The accumulated data from a number of laboratories present a large number of complete LCR transgenics that express their linked transgenes in a copy number–dependent, position-independent manner, suggesting that the full LCR is able to overcome centromeric position effects. This would not be expected on a statistical basis if centromeric silencing of the complete LCR were to occur. However, the chromosomal position of full LCR transgenic lines had not previously been reported, and hence we analyzed the chromosomal position of the eight full LCR transgenic lines available in our laboratory (Strouboulis et al., 1992a; Dillon et al., submitted; J. Gribnau et al., unpublished data). All were integrated in noncentromeric locations with the exception of line 2 (Strouboulis et al., 1992a), which is integrated close to a centromere (Figures 4A and 4B). This line, which contains a single copy of the 70 kb locus, expresses the human globin genes normally.

In Situ mRNA Analysis

In view of the well-characterized PEV observed in Drosophila and yeast, we next addressed the question of whether the position effect we observed in the LCR deletion lines was due to a reduction of the number of cells that express the transgene (as expected for PEV) or whether the expression level in each cell had been affected. To this end, in situ hybridization and indirect immunofluorescence were performed on disrupted fetal liver cells from the four position-effect lines as well as on other LCR-deletion and complete LCR lines with specific probes for human and mouse β-globin mRNA. The results show that in all lines tested, with the exception of Δ2B and Δ2C (Figure 5), the human transgene is expressed in all erythroid cells, since every cell that contains mouse β-major globin mRNA also contains human β-globin mRNA (Figure 5). In the two position-effect LCR deletion lines, Δ2B and Δ2C, the transgene shows variegated expression patterns, as human β-globin mRNA is present in only 25% and 4% of the cells, respectively, of every erythroid fetal liver examined. In situ hybridization of individual colonies from methyl cellulose cultures of line Δ2B fetal liver cells indicates that expression status is clonally inherited (data not shown). We conclude that the reduced level of transgene expression seen in these two lines results from clonally inherited PEV caused by centromeric heterochromatinization.

DNase I and Restriction Enzyme Sensitivity

Since the FISH and in situ mRNA analyses revealed that Δ2B and Δ2C were in heterochromatic regions and subject to PEV, whereas all other lines expressed the genes in all erythroid cells, it could be expected that a difference in DNase I sensitivity would be observed between the different lines. Expression of the locus in erythroid cells normally correlates with the presence of DNase I HS in the LCR and an HS in the promoter of the β-globin gene (Tuan et al., 1985; Forrester et al., 1987; Grosveld et al., 1987). We therefore tested for the presence of the erythroid-specific HS:5 (Zafarana et al., 1995) in the LCR (present in all constructs) and the HS found in the promoter of the β-globin gene using the mouse β-major gene promoter as a control (Table 1). Only lines Δ2B and Δ2C (the PEV lines) showed severely reduced or an absence of sensitivity, whereas all other lines, including Δ1A and Δ3C (severe PE), showed the LCR and promoter sites. We also assayed restriction enzyme sensitivity using ApaLI digestion in isolated nuclei from lines Δ1A and Δ2B and line 72. ApaLI cuts at +50 in exon 1 of the human β-globin gene. As expected from the DNase I results, ApaLI sensitivity in line 72 and Δ1A appeared similar, with approximately 60% digested, whereas the PEV line Δ2B showed a reduction in ApaLI sensitivity, with less than 20% digested (data not shown). These results suggest that the position effect observed in Δ2B and Δ2C is indeed caused by a

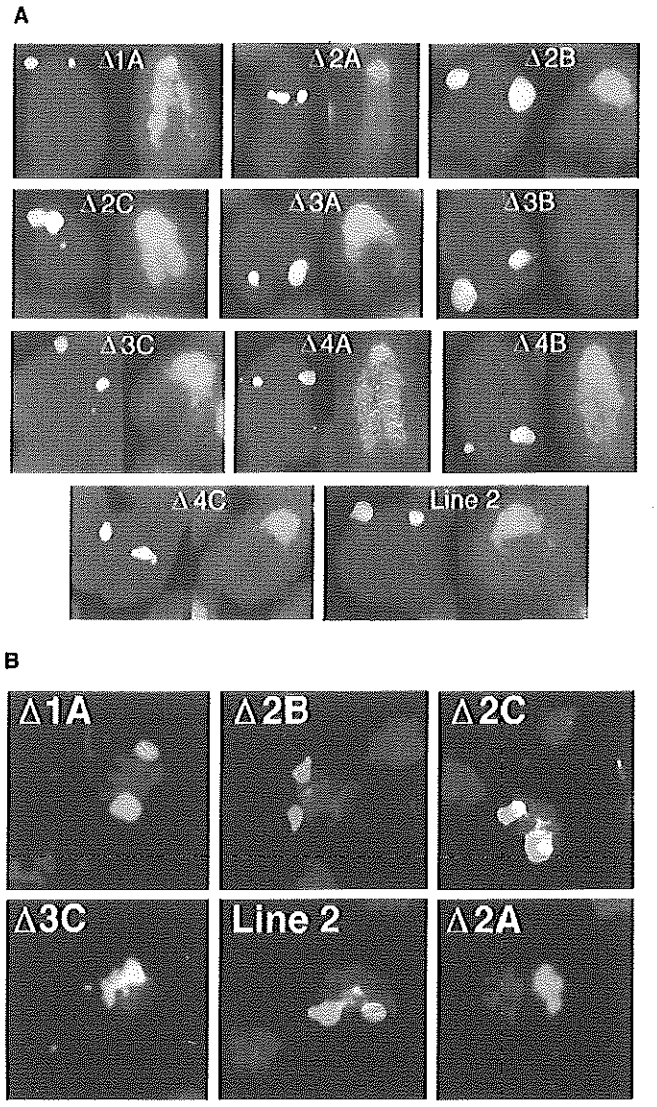


Figure 4. FISH Analysis of Transgenic Lines

(A) Left panels (red) show the human β-globin transgene locus detected by in situ hybridization of metaphase chromosomes of adult bone marrow cells with the complete insert as a probe (see Experimental Procedures). Right panels show DAPI staining (blue); centromeres correspond to the light blue areas. Transgenic line numbers are indicated on top of each sample.

(B) Double-label FISH of bone marrow metaphase chromosomes from the pericentromeric lines using probes for the human globin locus (green) and mouse α-satellite DNA (red).

closed chromatin conformation such as seen in heterochromatin, but they do not give an indication about the nature of the position effect observed in Δ1A and Δ3C.

In Situ Primary Transcript Analysis

Similar to the full LCR line 72, the two low expressing pericentromeric lines Δ1A and Δ3C express the human transgene in all erythroid cells. We decided to investigate the reduced expression levels of these two lines further by comparing the transcriptional status from several lines. We performed primary transcript in situ hybridizations with gene-specific intron probes for human β-globin and mouse β-major globin in 13.5 day fetal liver cells (Wijgerde et al., 1995). Lines Δ1A and Δ3C have significantly lower percentages (25% and 75%, respectively) of erythroid cells with human β-globin primary transcript signals (Figure 6; Table 2). The fact that all erythroid cells in these lines have human β-globin mRNA in their cytoplasm (Figure 5) demonstrates that the transgenes are active in each cell. The low percentage of transcriptionally active cells indicates that each transgene is transcribed for only a fraction of the normal time.

As expected, the PEV line Δ2B had human primary transcript signals in only 25% of erythroid fetal liver cells (Table 2), consistent with the percentage of erythroid

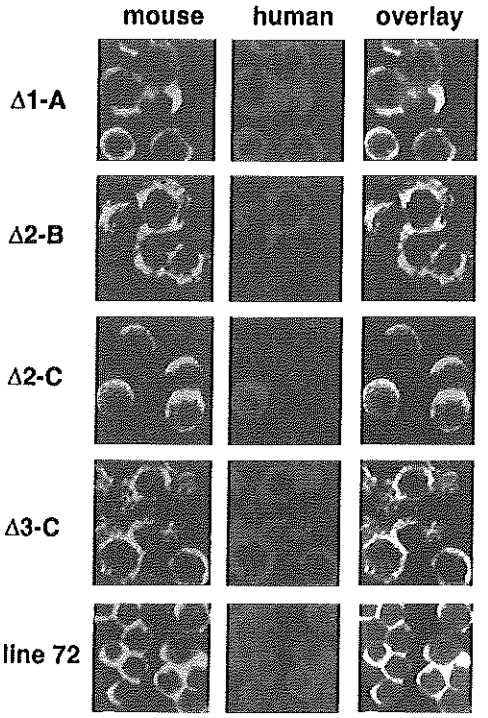


Figure 5. mRNA In Situ Hybridization

Left panels (green) show the in situ hybridization of 13.5 day fetal liver cells using a mouse β-major mRNA digoxigenin-labeled oligonucleotide from the third exon as the probe (Wijgerde et al., 1995). Middle panels show detection in the same cells of a deoxyribonucleoprotein (DNP)-labeled human β-globin third exon oligonucleotide probe (Wijgerde et al., 1995). Right panels show the overlay of the two images. Transgenic mouse line numbers are shown on the left.

cells that contained human β-globin mRNA (Figure 5). The full LCR control line 72 had β gene transcription signals in 87% of the erythroid cells as expected (10%–15% of loci transcribe the γ- or δ-globin genes at this stage [M. W. et al., unpublished data]). The highly expressing LCR deletion lines (Δ4A, Δ4B, and Δ4C) had positive transgene transcription signals in a high percentage of erythroid cells (Table 2). In the single-copy line Δ4A, the percentage of cells positive for β-globin transcription is similar to the single-copy line 72. The percentages approach 100% in the multicopy Δ4 lines (Δ4B and Δ4C) as expected, suggesting that the β-globin loci in these lines are continually active. Of potential significance is the fact that the multicopy lines Δ2A and Δ3B do not approach 100%, as would be expected if they were transcribing normally as in the Δ4 lines. This result suggests that they too may be transcribed for only part of the time.

Stochastic versus Timing Phenomenon

In an attempt to characterize further the novel position effect observed in Δ1A and Δ3C, we bred the PEV line

Table 1. DNase I Sensitivity

Transgenic Line	Human LCR HS:5	Human β-Globin Promoter	Mouse β-Major Globin Promoter
Line 72[a]	+	+	+
Δ1A[b]	+	+	+
Δ2A	+	+	+
Δ2B[b,c]	–	–	+
Δ2C[b,c]	–	–	+
Δ3B	+	+	+
Δ3C[b]	+	+	+
Δ4B	+	+	+
Δ4C	+	+	+

The DNase I hypersensitivity analysis was carried out on isolated nuclei from 13.5 day fetal liver cells. The presence or absence of LCR HS:5 or the HS in the human or mouse β-globin promoters is indicated by a plus or minus sign. The left column shows the transgenic mouse line number. The probes used are described in Experimental Procedures.

[a] Complete LCR.

[b] Pericentromeric location.

[c] PEV lines.

Δ2B and the timing effect line Δ1A to homozygosity. These two lines were chosen because the percentage of erythroid cells that had human globin transcription foci were the same in both lines (25%). If transgene activation is random, as is expected for PEV (Elliot et al., 1995), and the individual loci in a homozygous cell are stochastically activated or silenced, then the number of transcriptionally positive erythroid cells should increase from 25% to 44%. Most of these cells should have only one active locus and approximately 6% should have both transgene loci active. Transcription was analyzed via primary transcript in situ hybridizations for human and mouse β-globin. As expected for the PEV line (Δ2B), the percentage of transgene-expressing cells increased significantly in the homozygous fetal liver, from 25% to 43%. Most of these cells (74%) had only one chromosome actively transcribing the human genes. A quarter of these cells (26%) showed active transcription originating from both transgene chromosomes. These results are what would be expected from a stochastic activation of the globin transgene. In contrast, the percentage of transcription-positive cells in the Δ1A line did not increase significantly in the homozygotes (25%–31%), and surprisingly 80% of these cells show active transcription from both transgene loci. This result is the complete opposite of the PEV result. Therefore, locus activation in Δ1A appears to be a cellular timing phenomenon, since both chromosomes are activated simultaneously in a single cell and not stochastically at the level of the individual locus as in PEV. These results suggest that differentiation stage or cell cycle phase may be important in controlling heterochromatin-mediated silencing of the Δ1A transgene.

Discussion

We have analyzed the developmental expression pattern of the 70 kb human β-globin locus constructs in transgenic mice after deletion of individual HS from the LCR. Of 13 transgenic lines presented, at least 5 exhibit

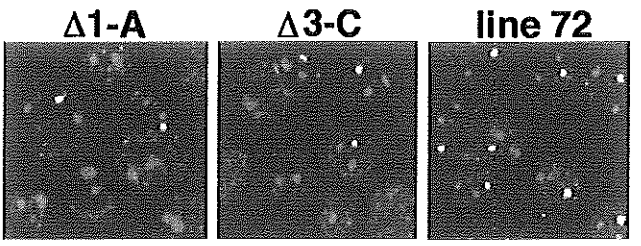


Figure 6. Primary Transcript In Situ Hybridization Analysis

In situ hybridization with gene-specific intron probes as described by Wijgerde et al. (1995). Red foci represent mouse β-major globin primary transcript (DNP labeled); green foci represent human β-globin primary transcript (digoxigenin). Transgenic mouse line numbers are indicated at the top.

extremely low levels of expression per copy. This demonstrates that deletion of a single HS can result in loss of position independence and copy number dependence, which are the major distinguishing features in the functional definition of an LCR (Grosveld et al., 1987), and suggests that expression of these transgenes is being influenced by the chromatin environment of the integration locus. The exceptions to this loss of function are the HS:4 deletion lines, which appear to show copy number–dependent and position-independent expression. However, with only three lines, it is possible that the transgenes in these mice have fortuitously landed in favorable chromosomal environments, and as a result the expression patterns observed appear position insensitive (Figure 4). However, others have shown that deletion of HS:4 from the locus can also result in loss of function (Bungert et al., 1995).

FISH analysis revealed that the transgenes in the lowest expressing lines were integrated at or near the centromere in each case. This type of pericentromeric localization of genes has been shown to lead to heterochromatin-induced silencing in Drosophila and yeast and, more recently, in mammals in the case of the CD2 LCR (Festenstein et al., 1996). We show through in situ analyses that reduced expression of pericentromeric transgenes can be accomplished by at least two functionally distinct mechanisms. The first is PEV, which results in a clonal subpopulation of erythroid cells that express the human globin genes. As a consequence, the levels of mRNA, DNase sensitivity, and restriction enzyme sensitivity have been severely decreased in the total population of cells. The second type of position effect exhibits low levels of transgene expression in all erythroid cells as evidenced by the S1 and mRNA in situ analysis. However, this low level of expression appears not to be caused by a general decrease in the rate of transcription of the individual genes. Rather, precursor RNA analysis reveals that only a fraction of the erythroid cells are transcribing the transgene at any moment, suggesting that low expression in these lines is caused by a decrease in the amount of time that the transgene is transcriptionally active in a particular cell. DNase I hypersensitivity and restriction enzyme sensitivity are maintained in these lines, suggesting that this position effect is caused by a chromatin conformation that differs from PEV. This result also suggests that DNase I hypersensitivity is not proportionally linked to transcription. Analysis of homozygous animals confirms that transcriptional activation is a cell-timing phenomenon and is not stochastic as in PEV. This presents two intriguing possibilities. First, we may be observing a window of reorganization of the genome during the cell cycle. Thus, the transgenes may be switching on briefly after disruption of higher order structure and then off again as heterochromatin spreads into the locus, or vice versa. Second, the transgene locus may be silenced prior to terminal differentiation of the red cells, which could account for the difference in transcriptional frequency observed between the transgene and endogenous loci.

Table 2. Percentage of Erythroid Cells Positive for Human Globin Transcription Foci

Transgenic Line	Copy Number	$\frac{\text{Human β Foci}}{\text{Mouse β Foci}} \times 100$
Line 72[a]	1	87
Δ1A[b]	1	25
Δ2A	8	89
Δ2B[b,c]	3	25
Δ3B	4	83
Δ3C[b]	3	75
Δ4A	1	86
Δ4B	2	97
Δ4C	3	95

Transcription foci were detected for mouse and human β-globin genes in heterozygous 13.5 day fetal liver cells as in Figure 6. The percentage of erythroid cells (i.e., cells positive for β-major transcription) that show transcription foci for the human β-globin gene are shown for various transgenic lines.

[a] Complete LCR.

[b] Pericentromeric location.

[c] PEV line.

The cell-timing effect may also be seen to a lesser extent in the noncentromeric LCR deletion transgenes (Δ2A [8 copies] and Δ3B [4 copies]). For example, transcription in situ analysis of line Δ3B shows that even though this line contains 4 copies of the transgenic locus, the percentage of erythroid cells that are transcriptionally positive for human β-globin is slightly lower than the single-copy full LCR line 72. This suggests that the LCR of each locus in these multicopy deletion lines spends a considerable amount of time uncoupled from the genes. The length of time that the LCR remains uncoupled appears to be exaggerated in the pericentromeric lines, which suggests that heterochromatin formation is somehow interfering with LCR activation, leading to the severe reduction in expression. Clearly this process is very different from PEV.

The results to date have shown that the individual HS of the LCR appear additive, as most or all were necessary for full expression (see Grosveld et al., 1993). The data presented here suggest that this additive effect is not accomplished through an increase in the density of

polymerases on an individual gene, but (at least in part) through an increase in the stability of the LCR-gene complex leading to increased frequency and duration of transcription periods. These results may seem to be in contradiction with the recently published analysis of deletions of the mouse β-major globin HS:2 and HS:3 in their native positions in the mouse genome (Fiering et al., 1995; Hug et al., 1996), both of which result in a 30% reduction in the expression of the mouse adult β genes only. Those experiments cannot measure position independence, since the mouse β-globin locus is maintained at its original position in the genome. Our observations suggest that the reduced expression seen in those knockout lines may be caused by a decrease in the stability of the LCR-gene complex leading to a reduction in frequency or duration of transcription. In this respect, it is worth noting that the individual deletions have essentially the same effect and that the effect is greatest on the most distal genes.

The constructs described here are subject to an inherent selection since the constructs are capable of very high levels of expression. As a result, the integration of multiple copies in areas of the genome that permit expression is lethal (Hanscombe et al., 1989). Thus, there is a bias toward survival of transgenic lines in which integration occurs in areas of the genome that reduce the level of expression. This is supported by the inherent difficulty in obtaining mouse lines that contain intact multiple copies of the complete locus (see also Stroubolis et al., 1992a; Dillon et al., submitted) and may explain why a higher than expected number of our lines have insertions in or near the centromere.

The results have important implications for the proposed mechanisms of LCR activation of gene expression. We have previously proposed that the different HS regions of the LCR may act together as a single functional unit that interacts with a single gene in the locus at any given time. This may be through formation of a holocomplex in which the individual HS must first interact among themselves to form a larger complex, which becomes the limiting element in gene competition. Clearly, deletion of one of the HS from the LCR results in position sensitivity and loss of copy number dependence, which suggests that the deletions may disrupt the normal interactions between the HS and the genes or formation of a holocomplex. This disruption appears to prevent the normal dominant positive effect of the LCR from overcoming negative effects that are thought to occur in heterochromatic regions of the genome.

Hence, we suggest that initiation of transcription from a competent gene in the "open" chromatin domain of the globin locus originates only through direct complex formation with the LCR and maintenance of that transcription requires continued association (see also Wijgerde et al., 1995). Deletions that decrease the stability of LCR–gene interactions decrease the frequency or shorten the duration (or both) of such associations, thereby reducing the transcriptional output from a given gene by decreasing the period of activity. This effect is most obvious in heterochromatic regions of the genome where the LCR spends a significant amount of time uncoupled (i.e., not activating any globin gene). Therefore, the complete LCR determines the level of gene expression by determining the frequency and the duration of transcription periods, rather than only controlling the rate of transcription.

Experimental Procedures

Construction of Human β-Globin Locus LCR Deletion Constructs

The human β-globin LCR was subcloned in two parts in pBR322. HS:4 was removed from the plasmid containing the 5′ half of the LCR by deletion of an 875 bp PmlI fragment that included the entire HS:4 core fragment as defined by Pruzina et al. (1991) to give plasmid ΔHS:4. HS:3 was deleted by removal of a 1.38 kb SauI fragment that includes the HS:3 core as defined by Philipsen et al. (1990) to give plasmid ΔHS:3. The HS:2 core (Talbot et al., 1990) was removed by deletion of a 742 bp ApaLI–SnaBI fragment to give plasmid ΔHS:2. The ΔHS:1 plasmid was created by deletion of a 1.7 kb HpaI–SpeI fragment that includes 919 bp from the 1.02 kb HS:1 fragment used by Talbot et al. (1989). The four 70 kb LCR deletion constructs were prepared by incorporating the deletion plasmids into cosmid (cos) LCR ϵ, which was linked to the 3′ cosmid (cos γγδβ) and purified for microinjection as described by Strouboulis et al. (1992).

Transgenic Mice

Purified fragments were microinjected into the pronuclei of FVB fertilized mouse eggs and transferred into the oviducts of pseudopregnant (CAB × C57B1) F1 female mice. Transgenic founders were identified via Southern blotting of tail DNA. Transgenic DNA was probed with the 70 kb injection fragment on Southern blots. Those that contained all the EcoRI fragments characteristic of the entire locus were selected for further breeding and analysis. Copy numbers were determined by quantitation of Southern blots via phosphorimage analysis. End fragment, junction fragment, and internal probes were used to determine the copy numbers and verify the structure of the transgenes.

Preparation of RNA and S1 Nuclease Protection Assays

RNA was prepared from whole, frozen, transgenic 10.5 day embryos and fetal livers from 12.5, 14.5, and 16.5 day and adult blood and subjected to S1 nuclease protection assays as previously described (Fraser et al., 1990). The probes used were those described by Lindenbaum and Grosveld (1990) with the exception of the human γ-globin probe. The human γ-globin probe was a 315 bp AvaII fragment taken from a plasmid containing the human A γ-globin cDNA ligated to the human A γ-globin 5′ flanking region. Human γ-globin RNA protects a 165 bp fragment of this probe from S1 nuclease cleavage. Protected bands on the S1 gels were quantitated via phosphorimage analysis (Molecular Dynamics).

DNA FISH Analysis

Extracted bone marrows were cultivated for 24 hr in RPMI 1640 medium (GIBCO BRL). Chromosome preparations were made according to standard procedures. FISH was carried out as described by Mulder et al. (1995). The probes used were the linked human β-globin locus (70 kb fragment) or a mouse minor satellite dimer (gift of A. Mitchell). The probes were labeled with biotin and digoxygenin and immunochemically detected with fluorescein or Texas red. The DNA was counterstained with DAPI.

HS Assay

Ten transgenic fetal livers (13.5 days) for each line were treated as described by Forrester et al. (1990). Suspension of nuclei was performed by 20 strokes of a Dounce pestle (type B) and 100 ml aliquots were digested for 3 min at 37°C with an increasing amount of DNase I. Reactions were stopped and treated with proteinase K and extracted with phenol–chloroform. After ethanol precipitation, the pellet was resuspended in 100 μl of water. We then digested 30 μl with EcoRI, followed by Southern blot analysis. Hybridizations were performed with the EcoRI fragment covering 5′HS:5, BamHI–EcoRI human βIVS II, or with an XbaI–Sau3A fragment covering the mouse β-major promoter.

mRNA and Primary Transcript In Situ Hybridization

Primary transcript and mRNA in situ hybridizations to detect transcriptional activity and expression of the human and mouse β-globin genes in fetal liver cells were performed as described by Wijgerde et al. (1995). Quantitation of transcription signals was done by counting at least 1000 cells from each line with an epifluorescence microscope. The figures presented were created with a laser scanning confocal microscope.

Acknowledgments

We are grateful to Lien Braam, Joost Gribnau, David Whyatt, David Michalovich, and Sara Pruzina. We thank A. Mitchel for the mouse satellite probe. E. M. was supported by a Medical Research Council (MRC) (Canada) postdoctoral fellowship and J. S. by the MRC (United Kingdom). This work was supported by the NWO (Netherlands), the Howard Hughes Foundation, and the European Union (Biotech).

Chapter IV

The effect of distant on long-range chromatin interactions

Molecular Cell 1997 1: 131-139

The Effect of Distance on Long-Range Chromatin Interactions

Niall Dillon,*‡ Tolleiv Trimborn,†
John Strouboulis,†§ Peter Fraser,†
and Frank Grosveld†‖
*MRC Clinical Sciences Centre
Imperial College School of Medicine
Hammersmith Hospital
London W12 ONN
United Kingdom
†MGC Department of Cell Biology and Genetics
Erasmus University
3000 DR, Rotterdam
The Netherlands
‡Gene Structure and Expression
National Institute for Medical Research
London NW7 1AA
United Kingdom

Summary

We have used gene competition to distinguish between possible mechanisms of transcriptional activation of the genes of the human β-globin locus. The insertion of a second β-globin gene at different points in the locus shows that the more proximal β gene competes more effectively for activation by the locus control region (LCR). Reducing the relative distance between the genes and the LCR reduces the competitive advantage of the proximal gene, a result that supports activation by direct interaction between the LCR and the genes. Visualization of the primary transcripts shows that the level of transcription is proportional to the frequency of transcriptional periods and that such periods last approximately 8 min in vivo. We also find that the position of the β-globin gene in the locus is important for correct developmental regulation.

Introduction

A variety of phenotypic effects in eukaryotes are known to be the result of long-range action by regions of DNA on one or more target sequences. Such effects include transcriptional activation by distal control elements, heterochromatinization of translocated genes to give position-effect variegation, and phenomena such as silencing and insulation. The phenomenon of transcriptional regulation by sequences that are located at long distances from the promoter has been known for many years, but there is still intense debate about how such regulation takes place. Looping models that bring distally located sequences into direct contact with promoters have been widely favored (Ptashne, 1988; Muller et al., 1989; Bickel and Pirotta, 1990). However, arguments have also been put forward in favor of a binary model that proposes that the sole function of distal elements is to generate a favorable chromatin structure that allows the promoter to function effectively (Weintraub, 1988; Martin et al., 1996; Walters et al., 1996).

The study of long-range effects presents formidable technical difficulties. Although it has been possible to observe looping between protein binding sites in vitro (Li et al., 1991; Mastrangelo et al., 1991; Su et al., 1991), such systems are necessarily extremely simplified compared with the in vivo situation. Direct visualization of contacts between transcriptional control elements in vivo has not yet been achieved, and most of our information about such contacts comes from genetic studies in Drosophila (e.g., Bickel and Pirotta, 1990). Methods for probing active and inactive chromatin configurations are limited and provide little direct information about the role played in transcriptional activation by changes in chromatin structure. An alternative approach that we have pursued is to use detailed functional analysis of a multigene locus to make inferences about chromatin dynamics in vivo.

The system that we use, the human β-globin locus, has long been a prototypic system for the study of transcription in vertebrates (reviewed in Grosveld et al., 1993). The locus consists of five developmentally regulated genes that are activated at different stages of erythroid development (Figure 1). The genes are arranged in the order in which they are activated during development with ε expressed first in the embryonic yolk sac. Between 6 and 10 weeks gestation, there is a gradual switch to expression of the γ genes, which predominate during the fetal liver stage. In the later fetal liver and neonatal stages, there is a further transition to expression of the β gene, and the γ genes are almost completely silenced during adult life.

The locus has been characterized in detail over a number of years, and this has resulted in the identification of the principal functional elements involved in its regulation. The entire locus is activated by the locus control region (LCR) contained within a 15 kb region located 5′ of the ε gene (Grosveld et al., 1987). It has long been known from genetic studies that expression of one gene in the locus can reduce that of the others (Giglioni et al., 1984). This effect has been studied extensively in transgenic mice (Enver et al., 1990; Hanscombe et al., 1991; Peterson and Stamatoyannopoulos, 1993), and there is clear evidence that it operates in a polar manner with the genes located proximally to the LCR having a stronger suppressive effect on the more distally located genes (Hanscombe et al., 1991). The simplest explanation for the observed down-regulation would be that activation is achieved through direct interaction between the genes and the LCR and that only one gene can interact with the LCR at any one time. In contrast, the binary model (Walters et al., 1996) excludes direct interaction and proposes transcriptional interference to explain the down-regulation of distal genes (Martin et al., 1996).

§Present Address: National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892.
‖To whom correspondence should be addressed.

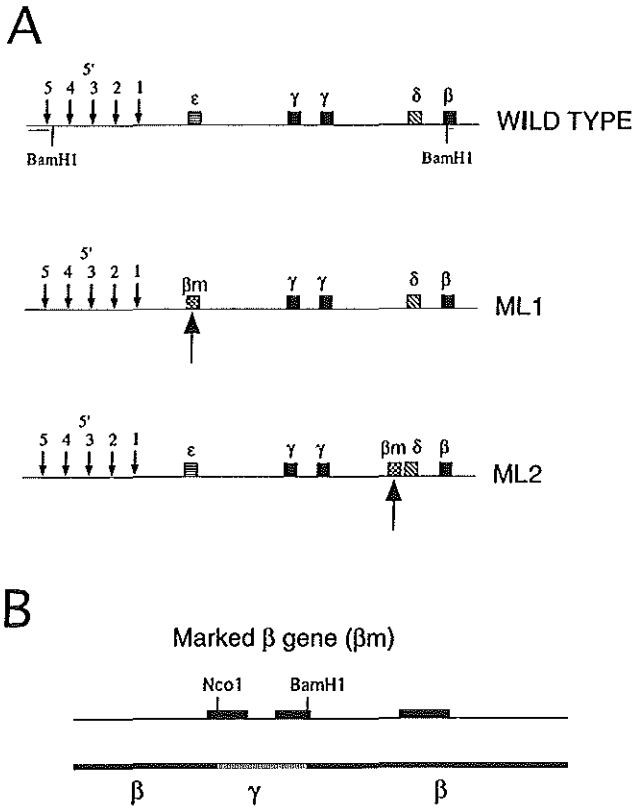


Figure 1. Structure of Normal and Modified β-Globin Loci

(A) Structure of the mutant β-globin loci used to generate transgenic mice. The LCR is indicated by the vertical arrows (1)–(5), and the genes as boxes. The thin lines and the BamHI sites indicate the positions of the probes that were used to analyze the ends of the transgenic loci (see Figure 2). Wild-type and mutant loci span a distance of 70 kb. In the ML1 locus, the β and βm genes are 50 kb and 6.8 kb, respectively, from the 3' end (HS1) of the LCR (56.7 kb and 13.6 kb from the middle of the LCR), while the equivalent distance for the Gγ gene is 28 kb. In ML2, the distances of the β and βm genes from the 3' end of the LCR are 53.7 kb and 42.8 kb, respectively (60.5 kb and 49.4 kb from the middle of the LCR).

(B) Structure of the marked β-globin gene (βm). A 410 bp fragment extending from the NcoI site at the ATG codon to a BamHI site located at the end of exon II was replaced with the equivalent fragment from the Aγ gene.

The study of interactions within the locus has recently been taken a step further by the use of in situ hybridization to directly visualize primary transcripts at individual alleles in single cells (Wijgerde et al., 1995). In transgenic mice homozygous for a human β-globin locus, cells from 12.5 day fetal liver were found to contain both γ and β primary transcripts, but almost all (87%) of the loci showed transcripts from only one gene. The result was interpreted as indicating that only one gene in the locus could be activated by the LCR at any one time. The presence of a minority of loci that gave double signals was interpreted as being due to the dynamics of the process. After a switch in gene activation, the newly synthesized RNA would colocalize with the decaying RNA of the previously transcribed gene. On the basis of these results and data obtained by manipulation of regulatory elements or transcription factors (Milot et al., 1996; Wijgerde et al., 1996), we postulated that transcriptional activation by distal elements is an all-or-nothing effect with transcription only taking place when the distal element is complexed with the gene. From this, it follows that the level of RNA produced would be determined by the frequency with which such transcriptional complexes are formed and the duration of the interactions.

However, it has never been possible to examine frequency and duration as independent variables because γ and β are different genes whose levels of expression change during the process of switching from γ to β. These factors introduce unknown variables into the analysis that limit the precision with which the kinetics of transcriptional competition can be measured (Wijgerde et al., 1995). In this study, we have eliminated these variables by generating mutant loci that contain a second functionally equivalent β gene at two different positions in the locus to allow the study of only the frequency of interaction.

We find that the steady-state RNA levels of the two β-globin genes are dependent on their relative distance from the LCR. The measurement of primary transcripts in individual cells shows that the number of transcriptional periods of each gene correlates with the levels of steady-state RNA. This provides independent evidence for direct interaction between the LCR and individual promoters and excludes transcriptional interference as a mechanism for the polarity of gene activation and silencing observed in the locus.

Results

Generation of Transgenic Mice Carrying Mutant β-Globin Loci

To test the effect of gene position on expression within the locus, we made use of a technique that involves

ligating two cosmid inserts to generate a 70 kb fragment containing the complete locus (Strouboulis et al., 1992b). Conventional cloning methodology carried out on the individual cosmids was used to generate two modified loci, each containing a second β-globin gene (βm) that was marked so that its transcript could be distinguished from that of the wild-type gene. In mutant locus 1 (ML1), the βm replaced the ε gene, while in mutant locus 2 (ML2), it was inserted close to the cap site of the δ gene (Figure 1A). Marking of the gene was achieved by replacing part of exon 1 and 2 and all of intron 1 with the equivalent sequences from the Aγ gene (Figure 1B). When marking the gene, it was important to reconstruct a fully functional globin gene to avoid generating a transcript that would be less stable in erythroid cells than the wild-type β-globin mRNA. Extensive functional testing has not detected any regulatory sequences within this region of the γ or β genes (Behringer et al., 1987; Bodine and Ley, 1987; Antoniou et al., 1988).

Transgenic mice were generated by microinjection of the mutant loci into oocytes. Since a multicopy tandem array would place an LCR close to the 3' end of the wild-type β-globin gene, it was necessary to analyze animals that carried the modified loci at single copy. To do this, we generated a large number of transgenic founder animals and then mapped the transgenes extensively with particular emphasis on the ends of the injected fragment, to identify mice that carried a complete single copy of the locus.

Analysis of Transgene Structure

A total of 57 founder transgenics were generated (23 for ML1 and 34 for ML2). Tail DNA from these founders was digested with BamHI and probed with fragments from either end of the locus (Figure 2). Founders that showed a joining fragment indicative of a multicopy tandem repeat were discarded, while putative single-copy animals were bred to generate transgenic lines. This approach resulted in the generation of 2 single-copy lines for ML1 and 3 for ML2 (see below and Figure 2 with accompanying legend). The integrity of the locus in each of these lines was tested by probing blots of EcoRI-digested DNA with the complete cosmids used to generate the locus (Strouboulis et al., 1992a). Five of the lines had the locus fully intact, while a small rearrangement was detected in the middle of the locus in one of the ML1 lines (data not shown). This line was not analyzed further. Line 610 had 2 copies of the locus integrated as a tandem head-to-tail repeat, while line 217 had 2 copies of the locus linked together but separated by an unknown amount of mouse DNA (Figure 2). Such events appear to occur frequently during transgene integration (Singh et al., 1991; Strouboulis et al., 1992a). The structure of the locus in line 335 is not clear. The EcoRI blot failed to show a 3'-end fragment although end fragments were observed with other digests (data not shown). This indicates that there is at least one intact copy of the locus in this line. However, it is also clear that the intensity of the 5'-end fragment (Figure 2, left panel) is greater than that observed for the other lines when compared to the weakly hybridizing mouse band observed in all lanes. This suggests that line 335 contains a second truncated copy of the locus (with a segment of unknown size missing from the 3' end of the second copy).

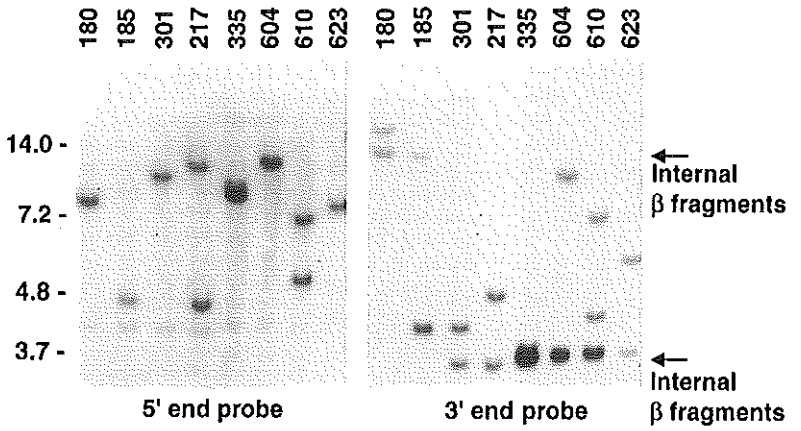
Effect of Position in the Locus on Transcriptional Function

Three different analytic approaches were used to obtain a comprehensive picture of the effect of position on the regulation of the two β genes in the mutant loci. Steady-state levels of transcripts from the two genes were measured by S1 analysis of adult blood RNA from each transgenic line. RNA FISH using intronic probes (Wijgerde et al., 1995) was used to detect the presence of primary transcripts at the site of transcription of the individual genes in 16.5 day fetal liver cells. This gives a measure of the relative frequency of transcription of the two β genes in erythroblasts. S1 analysis was also used to measure transcription of the genes at different stages of mouse development and to assess whether position in the locus affects developmental regulation. The results of these analyses are summarized in Figures 3 and 4 and described below.

Effect of Placing βm in a Distal Position in the Locus

The mutant locus construct ML2 places the βm gene in a position 10 kb upstream from the wild-type β gene. Placing the marked gene relatively close to the wild-type gene should show whether gene order affects transcription, either through competition for the LCR or transcriptional interference. Analysis of steady-state levels of mRNA from both genes in adult blood from four transgenic lines carrying the construct showed that the combined output of the two β genes is the same as that of endogenous mouse β (per copy of the mouse β locus). The data in Figure 3 show that expression of the two human β-globin genes is not equal. The presence of the βm gene at this position results in an expression level of approximately 75% (lines 301, 604, and 623) of the total and a reduction of expression of the wild-type human β gene to around 25% as measured by S1 protection analysis of steady-state RNA (Figure 3).

FISH analysis of primary transcripts was used to determine the number of actively transcribing genes in erythroblasts from 16.5 day fetal liver (when only the adult β genes are transcribed). βm and β primary transcription signals were detected by in situ hybridization with probes specific for the first intron of the γ genes (which detect primary transcripts from the βm gene) and the first intron of the β gene (which detects wild-type β transcripts). For each of the ML2 lines (301 and 623), approximately 500 expressing loci were counted for primary transcription signals. 50% (51% and 49% for lines 301 and 623, respectively) of the alleles show a single transcription signal for βm (Figure 3, green), whereas 13% (in 301 and 623) have a single β signal (Figure 3, red). A combined signal (red + green = yellow) is the result of simultaneous transcription of one gene and decay of primary transcripts of the other gene (Wijgerde et al., 1995; Gribnau and de Boer, personal communication). This is observed in 37% of the cells (36% and 38% in lines 301 and 623, respectively). Since the βm and β



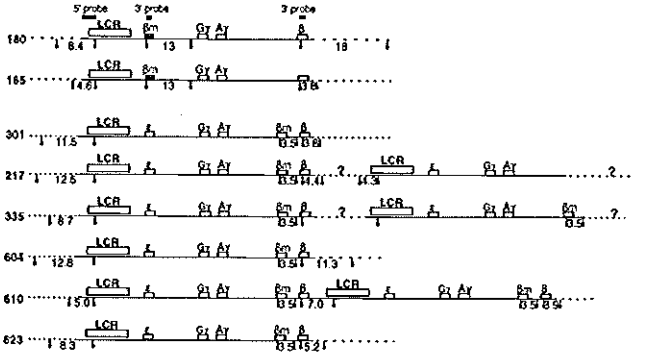


Figure 2. Analysis of Transgene Structure

End-blot of five single-copy lines together with two additional two-copy lines for ML2. The right-hand panel shows a BamHI digest probed with a 3.3 kb EcoRI fragment from the 5' end of the locus (see Figure 1). The left-hand panel shows the same blot probed with a fragment from the second intron of the human β-globin gene. The size of the internal fragment differs in the ML1 and ML2 lines because of the different location of the gene in the locus. Each of the five single-copy lines gives a differently sized end fragment when hybridized with probes specific for the 5' and 3' ends of the locus (the additional internal fragment observed with the 3' probe comes from the βm gene). Line 335 shows a higher signal than expected for a single-copy integration, but we have not detected a joining fragment as would be expected for a multicopy integration. It also fails to show a 3'-end fragment when digested with BamHI; possibly the BamHI end fragment is larger than the average size of the DNA in these preparations. Further blots did not clarify the situation, and hence, the exact structure of line 335 is not clear. Line 610 gave two end fragments and a fragment that corresponds to the size expected for a head-to-tail joining fragment, indicating that the line carries two copies of the transgene arranged as a head-to-tail repeat. Line 217 also gives two different end fragments when probed with the 5'-end probe. Although one end fragment is visible for the 3'-end probe, NcoI and HindIII digests each give two 3'-end fragments (not shown), indicating that this line contains two copies of the locus integrated in the same region but separated by an unknown amount of mouse DNA. The end fragments segregated together when the line was bred, indicating that two integration events have occurred close together.

genes are the same, the number of double spots (red + green = yellow) is only dependent on their frequency of transcription and signal decay. Using the average numbers (50% green, 13% red, and 37% yellow), it can be calculated (see Experimental Procedures and Discussion) that the frequency of β transcription is 31% and that of βm is 69%. This is in good agreement with the S1 protection analysis.

We conclude that the presence of the additional β gene down-regulates the distal β gene. Since the two β genes have identical promoters and flanking sequences, we also conclude that the down-regulation is caused

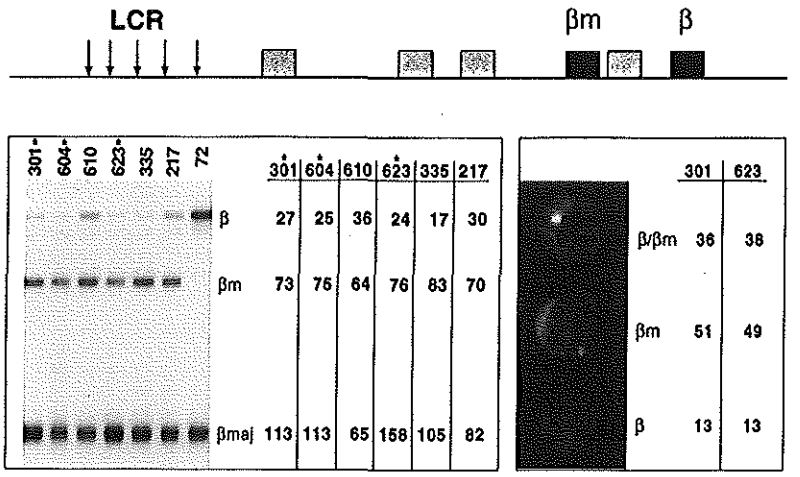


Figure 3. Transcriptional Analysis of the ML2 Locus

(Top Line) The ML2 construct with the βm and β genes indicated in black.

(Left Bottom) S1 analysis of adult blood RNA from lines transgenic for ML2 and the wild-type β-globin locus (Strouboulis et al., 1992a). The protected fragments are indicated on the right. The bands were quantitated by phosphorimage analysis and corrected for copy number. The numbers were normalized to 100% total human β and are shown on the right.

(Right Bottom) In situ hybridization of two of the ML2 lines with the percentages of cells showing βm/β double, βm single, and β single signals.

by a decrease in the frequency of transcription of the distal β gene. Possible explanations for this decrease would be competition for the LCR (Enver et al., 1990; Hanscombe et al., 1991) or transcriptional interference (Proudfoot, 1986; Martin et al., 1996).

Effect of Placing βm in a More Proximal Position

If transcriptional interference plays a role in down-regulating distal genes, then the effect would be expected to be most pronounced when the βm gene is closest to the β gene and to be reduced when the βm gene is placed in a more proximal position (i.e., closer to the LCR). If relative distance from the LCR is the main determinant, then the opposite effect would be expected. These predictions were directly tested using the construct ML1 in which βm is in the position normally occupied by the ε gene (Figure 1). The results of S1 analysis of adult blood from the two transgenic lines (180 and 185) obtained for ML1 are shown in Figure 4. Again, the total output of the locus is very similar to that of the endogenous mouse loci. In line 180, expression of the level of β RNA was 12% of the total human β RNA, while in line 185, expression was reduced to less than 1% (Figure 4). These results are confirmed by in situ analysis. Line 185 shows only 0.4% of the expressing loci with a single β (red) and 0.4% with a double βm/β (yellow) signal, while 99.2% express βm only (Figure 4). Line 180 shows 1.2% single β, 14.4% double βm/β, and 84.4% single βm signals. From these data, it can be calculated (see Discussion) that the frequency of transcription of βm in line 185 is 99% of the total, and in line 180, 91% of the total. As in the ML2 mice, these in situ data are in agreement with the S1 protection analysis. However, the variation between the ML1 mice indicates that one of these lines is subject to a position effect (see Discussion).

These data clearly show that moving the proximal gene closer to the LCR and much farther away from the distal gene results in a more effective suppression of the distal gene. In addition, since the proximal βm gene is now in a position where transcriptional interference on the distal β gene would be expected to be reduced (and would certainly not be increased), we conclude that down-regulation of the distal gene is the result of competition between the genes for direct interaction with the LCR.

Developmental Regulation

It has been suggested that the β-globin gene is normally at least in part suppressed in the early stages of embryonic development due to its position in the locus and, as a consequence, the competition from the ε and the γ genes (Enver et al ., 1990; Hanscombe et al., 1991; Dillon and Grosveld, 1993; Peterson and Stammatoyannopoulos, 1993). Since the ML1 mice carry a β-globin gene in the position of the ε gene, they provide an excellent opportunity to test the competitive silencing early in development. Timed matings were carried out for four of the transgenic lines (two for each construct), and the RNA was analyzed from 10.5 day yolk sac, 12.5 day fetal liver, and adult blood for the presence of transcripts from the human β, βm, and γ genes and the mouse embryonic βh1 and adult βmaj genes (Figure 5). Embryonic yolk sac from the transgenic line 72 carrying the wild-type locus shows the expected pattern of human

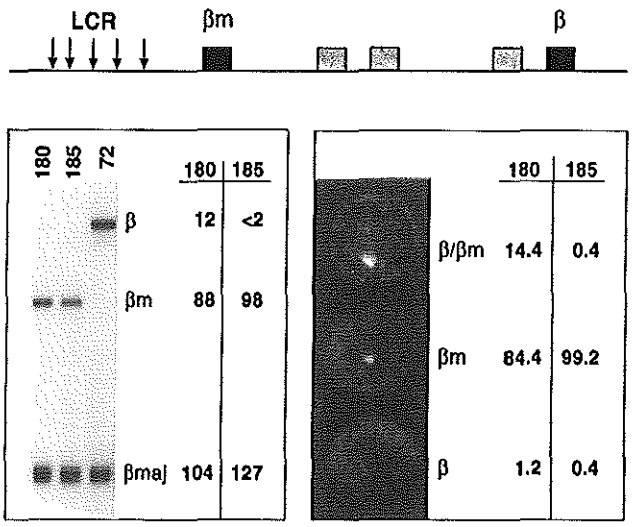


Figure 4. Transcriptional Analysis of the ML1 Locus

(Top Line) The ML1 construct with the βm and β genes indicated in black.

(Bottom Left) S1 analysis of adult blood RNA from lines transgenic for ML1 and the wild-type β-globin locus (Strouboulis et al., 1992a). The protected fragments are indicated on the right. The bands were quantitated by phosphorimage analysis, and the normalized numbers are shown on the right.

(Bottom Right) In situ hybridization on day 16.5 of development of two of the ML1 lines with the percentages of cells showing βm/β double, βm single, and β single signals.

globin expression with the γ genes expressed at high levels and the β gene completely suppressed. A similar pattern is observed for the ML2 lines (the βm gene in the distal position) with both βm and β completely suppressed. In contrast, in the two ML1 lines (the βm in the proximal position), βm is expressed at levels that are approximately equivalent to those of the γ genes. This result shows that the β-globin gene does not contain any flanking sequences (2 kb in either the 5′ or 3′ direction) that actively suppress its transcription in early development.

In fetal liver, the control line 72 shows the expected levels of human γ and β expression (Strouboulis et al., 1992a; Peterson et al., 1993, Figure 5). In contrast to both ML1 lines (βm located upstream from the γ genes), expression of γ in 12.5 day fetal liver is completely suppressed. In the ML2 lines (βm located downstream from γ), expression is observed in the early fetal liver. We therefore conclude that the position of the genes relative to the LCR profoundly affects their developmental regulation and that this is the effect of transcriptional competition. Our results indicate that the β-globin gene is normally supressed in embryonic and fetal tissues because its distal location results in a competitive disadvantage compared with the ε and γ genes.

Discussion

In this study, we have used a combination of technologies to examine chromatin interactions over large DNA regions. The LCR approach (i.e., using all of the elements required for physiological expression) allows quantitative conclusions to be drawn from measurements of steady-state levels of mRNA. The in situ analysis of primary transcripts provides qualitative information about transcription in each cell in a population. Techniques for manipulating large fragments and introducing them into mice permits the application of reverse genetics at the level of a complete locus. Using these approaches, we have been able to address a number of different questions.

Competition versus Transcriptional Interference

It has long been known that the transcription of one gene can affect the expression of other genes in the β-globin locus. Two different models have been suggested as explanations for this phenomenon. The competition model proposes that transcription depends on direct interactions between the genes and the LCR (Hanscombe et al., 1991; Wijgerde et al., 1995) and that such interactions are monogenic. The binary model proposes that the effect of the LCR on transcription is due to an alteration in chromatin structure and that the level of expression is determined entirely by the promoter without direct interaction with distal elements. As this model does not explain the fact that up-regulation of one gene results in a reduction in the expression of the others, it was necessary to invoke transcriptional interference as an additional mechanism to explain this effect (Martin et al., 1996). We distinguished between the competition and binary model by placing an extra β-globin gene (βm) at two different positions in the β locus. It would be expected that if down-regulation is caused by transcriptional interference, then the effect would decrease when the extra β gene is moved away from the gene with which it would interfere (ML1 versus ML2, Figure 1). In fact, suppression is substantially increased in ML1, and we conclude that interference cannot explain our result. By inference, transcriptional interference plays at most a minor role in the regulation of the β-globin locus. Of course, our results do not exclude the possibility that transcriptional interference plays a role in other situations.

The LCR contains five hypersensitive sites spread over a distance of 15 kb. Functional analysis in transgenic mice has shown that four of the sites are required

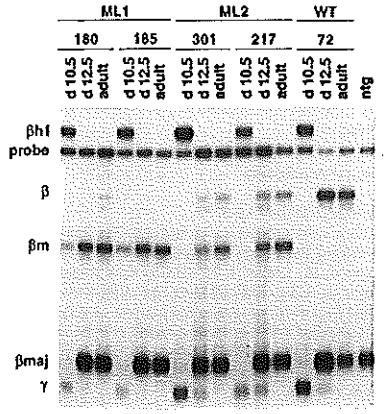


Figure 5. Developmental Regulation of the ML1 and ML2 Loci

S1 analysis of RNA from 10.5 day yolk sac, 12.5 day fetal liver, and adult blood from lines transgenic for ML1, ML2, and the wild-type β-globin locus. The protected fragments of the various probes are indicated on the left.

for full activation of a linked β-globin gene (Bungert et al., 1996; Milot et al., 1996; Peterson et al., 1996). The fact that the genes compete for the LCR and that a proximal gene can almost completely suppress transcription of a distal gene implies that the LCR functions as a single holocomplex as proposed by Ellis et al. (1996) rather than individual sites contacting the genes independently (Engel, 1993).

Kinetics of Complex Formation between the LCR and the Genes

The finding that there is competition between the genes in the locus and the observation of a predominance of single signals by in situ analysis demonstrates that the LCR can activate only one gene at a time (Wijgerde et al., 1995). The proportion of single and double signals depends on the frequency of switching, the duration of the interaction between the LCR and each of the genes, and the lifetime of the double signal after a switch. The two identical competing genes, βm and β, will have the same duration of interaction with the LCR. Since the lifetime of the signal (7 min, Gribnau and de Boer, personal communication) and the number of single and double signals are known, the frequency and duration of a transcriptional period can be estimated. The data obtained with the ML2 mice show that the number of single βm- and β-gene transcription signals for every 100 cells is 50 and 13, respectively, whereas 37 cells have a double signal (Figure 3). From these numbers, it is calculated (see Experimental Procedures) that there is a 69% probability of transcribing the βm gene versus a 31% probability of transcribing the β gene when both are in a distal position, as in the ML2 mice (301 and 623). Thus, the βm gene is transcribed 2.3 times more frequently than the β gene. It can also be calculated that the time of a transcriptional period lasts 8.2 min on average (see Experimental Procedures).

When the same calculation is used on the two ML1 lines (180 and 185) that were analyzed by in situ analysis, it results in two different values, namely a βm-to-β frequency of 160-fold and a transcriptional period of 20.5 min for line 185 (the very low β-globin expressor) and a βm-to-β frequency of 9-fold and a period of 7.5 min for line 180. The latter is very similar to the 8.2 min observed for the ML2 lines, and we conclude that the distal gene in line 185 is influenced by a negative position effect. In general, distal genes in single-copy integrations show some sensitivity to the position of integration in the mouse genome (Strouboulis et al., 1992a), in particular when the LCR interacts predominantly with a proximal gene. In that situation, the distal gene is not engaged by the LCR for most of the time and would therefore be available for an interaction with neighboring mouse sequences.

Effect of Position in the Locus on Competition

Our results show that placing a gene at different positions in the locus affects the frequency of its interaction with the LCR. A more distal gene is always at a disadvantage, but the effect is accentuated when the proximal gene is much closer to the LCR than the distal gene. What is the mechanism for this effect of relative distance from the LCR on competition? The more efficient competition by a proximal gene could be explained by a tracking model, where the LCR traverses along the DNA until it finds a gene and forms a complex with it. However, a simple tracking model predicts that the effect of the proximal gene would be the same whatever its position relative to the distal gene and does not explain the effect of relative distance. An alternative model (Dillon and Grosveld, 1993) proposed that LCR and genes can move freely in solution and that random contact between the genes and the LCR results in the formation of stable complexes. A gene that is closer to the LCR would come into contact with it more frequently and would therefore compete more effectively. As the proximal gene is moved away from the LCR and toward the distal gene, the difference in frequency of interaction would become less and the competitive advantage of the proximal gene would be reduced. The kinetics of free movement of two tethered elements predict that frequency of contact will be a function of relative distance to the power of 3/2 (Rippe et al., 1995). Calculation of the distance from the genes to the LCR is complicated by the fact that there are four hypersensitive sites spread out over a distance of 15 kb. Since the structure of the LCR in vivo is not known, only a range of 15 kb can be used by measuring the distance from the genes to either the 5′ end (HS4) or the 3′ end (HS1) of the LCR. This would result in a distance of β-LCR/βm-LCR to the power 3/2 of 6–20 for ML1 and 1.4–1.5 for ML2. Thus, if the genes were moving freely in solution, the predicted expression ratios of βm/β would be in the case of ML1 of 6- to 20-fold and ML2 of 1.4- to 1.5-fold.

However, the measured effect of competition by the proximal gene on the distal gene does not quite fit these predicted numbers (9-fold for ML1 line 180 [no position effect] and 2.3-fold for ML2). In particular, the deviation of the more reproducible numbers for ML2 implies that movement of two points on a fragment of DNA in the

nucleus is subject to additional constraints over and above those imposed by the fact that they are tethered together. The nature of these constraints is unclear, but the most likely constraint is probably the fact that the locus is part of a bigger loop.

Effect of Position on Developmental Regulation

Our results also demonstrate that the effect of position on competition between the genes of the β-globin locus for the activating function of the LCR has a substantial effect on the developmental regulation of the locus. Specifically, we have shown that embryonic silencing of the β-globin gene is dependent on its location in a distal position in the locus. We cannot say for certain which sequences in the proximal region of the locus are responsible for this embryonic activation, but the 2.7 kb fragment containing the ε gene that was removed and replaced with the β gene contains sufficient information to direct high-level expression when linked to the LCR (Raich et al., 1990; Shih et al., 1990), suggesting that any ε-specific enhancers have been removed from ML1. Embryonic expression of the γ genes in ML1 is significantly reduced compared with the wild-type locus, suggesting that the more proximal β gene actually competes more effectively with γ than with the ε gene that normally occupies this position. However, a similar reduction is observed when an extra β gene is located downstream from the γ genes, even though this gene is completely silent. Taken together with our other results, this suggests that the locus has evolved as a functional system in which the number of genes and the spacing between the genes are important parameters that act in concert with all of the regulatory elements to give the final level of expression. This has important implications for our understanding of the regulation of other complex multigene loci such as the highly conserved *Hox* loci and the imprinted murine *H19* region. In the latter case, an enhancer competition model has been proposed to explain the fact that deletion of the *H19* gene eliminates imprinting of the *Ins2* and *Igf2* genes located 95 and 80 kb away, respectively (Leighton et al., 1995). The study described here provides the first direct evidence that competition for shared elements does occur over distances of this magnitude in chromatin.

Polar competition clearly modulates switching between the γ and β genes in the early fetal liver and gives the γ genes a strong advantage by virtue of their more proximal position. However, the γ genes are expressed at similar levels to that of β in the early fetal liver, which means that other parameters must be acting to counteract this positional advantage. For example, destabilization of the interaction between the LCR and γ genes may lead to a reduced frequency of productive interaction following contact between the LCR and the γ genes.

Our results demonstrate the power of this type of functional analysis for making inferences about transcriptional behavior in chromatin. A particular advantage of this approach of functional probing is the fact that it uses the modulation of gene function in the native context to draw inferences about the relationship between structure and transcriptional behavior. The most interesting implication of this analysis is that the initiation of transcription of a gene can only take place while the distant regulatory sequence (LCR) is in direct contact with the gene and that the process of initiation stops as soon as the LCR/gene complex dissociates. Further use of functional probing should allow direct testing of a variety of different models for gene function in chromatin in vivo.

Experimental Procedures

Construction of Mutant Loci

Modifications were carried out on the cosmids Cos-LCRε and Cos γγδβ containing the 5′ and 3′ segments of the locus, respectively (Strouboulis et al., 1992a). A 2.7 kb ClaI/KpnI fragment containing the human ε gene was excised from cos-LCRε and replaced with a 4.9 kb BglII fragment containing the marked β-globin gene (Figure 1B). The same BglII fragment was inserted into the SalI site located at the cap site of the δ gene in Cos γγδβ. The two cosmids were joined together using the oligo-tailing procedure described by Strouboulis et al. (1992b).

Generation and Analysis of Transgenic Mice

The 70 kb fragments containing the complete mutant loci were purified by preparative agarose gel electrophoresis and injected into mouse oocytes. Southern blotting of tail DNA was carried out using probes specific for the ends of the locus. The intactness of the locus was confirmed by probing with the complete Cos-LCRε and Cos γγδβ. All methods used have been previously described by Strouboulis et al. (1992a).

S1 Analysis of RNA

RNA was isolated from 10.5 day yolk sac, 12.5 day fetal liver, and adult blood from transgenic lines and subjected to S1 analysis using 5′-end-labeled probes. The S1 procedure and the probes used for the mapping were identical to those used by Strouboulis et al. (1992a). The probe for the βm gene was a 190 bp fragment extending from −67 to +123 in the fusion gene. Quantitation was carried out by scanning.

In Situ Primary Transcription Analysis

The in situ analysis was carried out essentially as described by Wijgerde et al. (1995), using oligonucleotide probes that detect the first intron of the γ gene (visualizing βm as green signals) and the first intron of the β gene (visualizing the β gene as red signals). The equilibrium between βm and β transcription is then divided up in transcriptional time intervals T, while the lifetime of the double signal is 7 min (Gribnau and deBoer, personal communication). At the start, there is a possibility of each time interval to transcribe βm or β with a probability P or 1-P, respectively. There are four possible states per interval: (1) switching from β to βm (with probability P) results in 7 min double signal followed by T-7 min single βm signal; (2) switching from βm to β (with probability 1-P) results in 7 min double signal followed by T-7 min single β signal; (3) going from βm back to βm (with probability P) results in T min βm signal; and (4) going from β back to β (with probability P) results in T min β signal. This gives a Markov chain that can be solved in three equations per time interval T: (1) expected time that βm is observed, $P(1 - P)(T - 7) + P^2$ = (%βm signals)T; (2) expected time that β is observed, $P(1 - P)(T - 7) + (1 - P)^2$ = (%β signals)T; and (3) expected time that βm/β double is observed, $2P(1 - P)7$ = (%βm/β signals)T. For example, these equations can be solved to result for line ML2 in P = 69% and T = 8.2 min.

Acknowledgments

We are grateful to Dr. T. Stijnen and J. Gribnau for their advice on the statistical evaluation of the in situ transcription data, to M. Kuit for photography, and to L. Braam for animal care. This work was supported by the Medical Research Council (United Kingdom), the Howard Hughes Medical Foundation, and the NWO (The Netherlands).

Received June 2, 1997; revised September 25, 1997.

Chapter V

Chromatin interaction mechanism of transcriptional control *in vivo.*

in press

Chromatin interaction mechanism of transcriptional control *in vivo*

Joost Gribnau, Ernie de Boer, Tolleiv Trimborn, Mark Wijgerde, Eric Milot, Frank Grosveld and Peter Fraser

MGC Department of Cell Biology Erasmus University, PO Box 1738, 3000 DR Rotterdam, Netherlands.

We have used a kinetic analysis to distinguish possible mechanisms of activation of transcription of the different genes in the human β globin locus. Based on *in situ* studies at the single cell level we have previously suggested a dynamic mechanism of single genes alternately interacting with the Locus Control Region (LCR) to activate transcription. However those steady state experiments did not allow a direct measurement of the dynamics of the mechanism and the presence of loci with *in situ* primary transcript signals from two β-like genes *in cis* has left open the possibility that multiple genes in the locus could initiate transcription simultaneously. Kinetic assays involving removal of a block to transcription elongation in conjunction with RNA FISH show that multiple β gene primary transcript signals *in cis* represent a transition between alternating transcriptional periods of single genes, supporting a dynamic interaction mechanism.

Several models have been proposed for the activation of gene transcription after chromatin activation of a locus. In one model, the genes would simply be accessible to binding of transcription factors and be transcribed in a stochastic fashion (Groudine and Weintraub, 1982; Martin et al., 1996). In such a model the dimensional aspects of the locus would not play an important role. In a second model regulatory sequences could be the entry site of (part of) the transcriptional machinery which would scan the DNA for genes to be transcribed (Herendeen et al., 1992; Tuan et al., 1992). Such a model is basically linear and predicts that the order of the genes relative to the regulator is an important parameter. In a third model it has been proposed that gene regulatory elements participate in direct chromatin interactions with regulatory elements at a large distance as a prerequisite to transcriptional activation (Ptashne, 1988; Bickel and Pirotta, 1990; Mueller-Storm et al., 1989; Foley and Engel, 1992; Wijgerde et al., 1995; Dillon et al., 1997). Such a looping model is three dimensional and predicts that the relative distance of the genes to the regulator is important. All these models have been put forward to explain the role of the human β globin locus control region (LCR) in the developmental regulation of transcription of the β gene cluster (Martin et al., 1996; Tuan et al., 1992; Wijgerde et al., 1995; Dillon et al., 1997).

The β globin system has long been a prototypic system for the study of transcription in vertebrates (reviewed in Grosveld et al., 1993). The locus consists of five active genes that are activated and silenced at different stages of erythroid development (Figure 1A). The expression of all of these genes is dependent on the presence of the LCR which is located 15kb upstream of the ε gene (Grosveld et al., 1987). The ε gene is expressed first in the embryonic yolk sac followed by a gradual switch to expression of the γ genes between six and ten weeks of gestation. Expression of the γ genes predominates during the fetal liver stage. In the later fetal liver and neonatal stages, there is a second transition to expression of the β gene and the γ genes are almost completely silenced during adult life. When the entire human β locus is incorporated in transgenic mice a similar expression pattern is observed, although the γ genes are expressed early in the embryo and are switched off at day 16 of development in the fetal liver (Strouboulis et al., 1992; Peterson et al., 1993). The analysis of mutated loci found in patients and the use of single ε, γ and β genes in transgenic mice have shown that the ε and γ genes are suppressed autonomously through sequences directly flanking the genes (Dillon et al., 1991; Raich et al., 1990). However the β globin gene when present in the whole locus is (at least in large part) silenced during early development in a non autonomous manner. This β globin gene suppression can be explained by a scanning mechanism because genes closer to the regulatory sequences would have a natural advantage over distal genes due to proximity. This would also be the case in a looping mechanism, proximal genes would have a higher frequency of interaction with the regulatory sequences and thus have

a competitive advantage over distal genes (Giglioni et al., 1984; Enver et al., 1990; Hanscombe et al., 1991; Peterson and Stamatoyannopoulos, 1993; Dillon et al., 1997). However this would not be the case in an accessibility model and hence an extra parameter was postulated to explain the silencing of the distal β gene in early development, namely a process of interference of the proximal genes with the distal genes via some topological constraint (Martin et al., 1996).

Recent analysis of primary transcription in single cells (Wijgerde et al., 1995, 1996) and the results obtained by placing a second β globin gene at different positions in the locus (Dillon et al., 1997) support a dynamic looping mechanism with single genes alternately interacting with the LCR. However, the presence of a minority of loci which display two gene signals *in cis* (Wijgerde et al., 1995) could be interpreted as evidence in support of the scanning or the accessibility model of transcription initiation (Martin et al 1996). Thus a crucial difference between the looping model and the others is single vs multiple gene activation at any moment in a single locus. We have therefore used a novel kinetic analysis utilizing inhibition of transcription elongation and release in conjunction with RNA FISH to show that multiple β gene primary transcript signals *in cis* represent a transition between alternating transcriptional periods of single genes, rather than the co-initiation of transcription of multiple genes in the locus.

Results

The lifetime of primary transcript in situ hybridization signals

A key parameter in a kinetic analysis of the transcription process *in vivo* using *in situ* hybridization, is the time required for a signal to decay below the level of detection. To enable the detection of short lived events we probed for the presence of intronic RNA as these sequences are rapidly cleaved from the primary transcript and degraded. We used actinomycin-D to measure the detection lifetime of the γ and β primary transcript signals at days 11.5 (Figure 1C) and 12.5 (not shown) and ζ and α primary transcript signals at day 10.5 (Figure 1D) of development. On both day 11.5 and 12.5 the intron signals of the γ and β genes as well as ζ and α intron signals at day 10.5 disappear below detection level 7.5 minutes after the addition of actinomycin-D. Interestingly the intron signals of the ζ and α genes do not decrease immediately when compared to the β-like genes. This could indicate a possible difference between the two loci (see below), but could also be due to experimental parameters such as a higher sensitivity of the α-like probes.

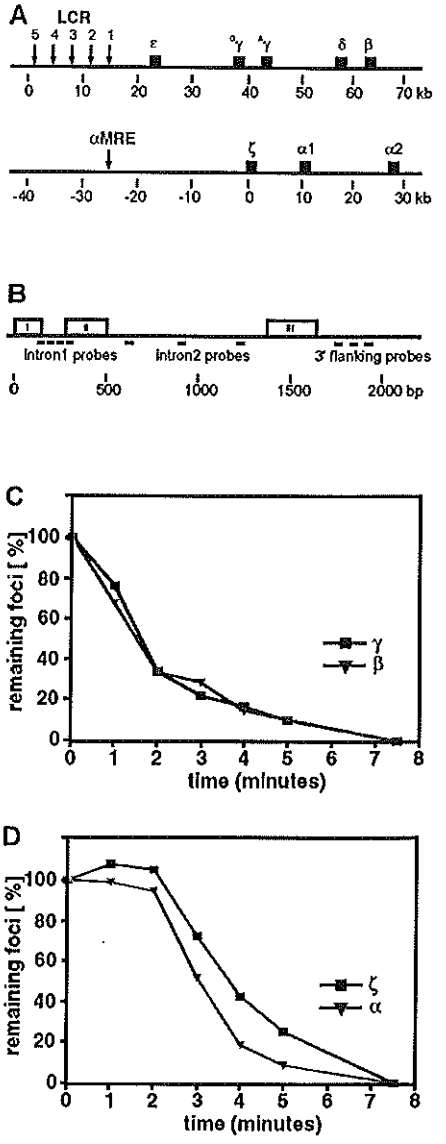


Figure 1. Lifetime of the human β and mouse α globin primary transcript signals *in situ*.. (A) Schematic diagrams of the human β globin locus and mouse α globin locus. Black boxes represent genes and vertical arrows represent the hypersensitive sites of the β globin LCR and the α globin major regulatory element (αMRE). (B) Intron/exon structure of a β like globin gene with localization of probes used for *in situ* hybridizations. (C) Decay of the γ and β primary transcripts *in situ* signals in 11.5 day fetal liver cells treated with 5 μg/ml actinomycin-D. The percent of remaining intron 2 signals are shown as a function of time. (D) same as panel C for the mouse ζ and α genes in 10.5 day embryonic blood cells.

Unfortunately actinomycin-D inhibition is irreversible and hence cannot be used in reactivation experiments. We therefore used 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole (DRB) to reversibly inhibit pol-II transcriptional elongation in single copy human β globin transgenic mouse (Strouboulis et al., 1992) fetal liver cells. Previous studies have shown that DRB does not effect initiation of transcription (Fraser et al., 1978; Marshall et al., 1992) but prematurely aborts elongating transcripts approximately 400-600 bp from the initiation site (Chodosh et al., 1989; Marshall et al., 1996) by inhibiting the activity of the P-TEFb kinase which phosphorylates the carboxy terminal domain (CTD) of pol II (Marshall et al., 1996; Peng et al., 1998). The effect of DRB treatment on globin gene transcription in mouse erythroleukemia cells has been reported previously (Tweeten and Molloy 1981). The results demonstrate that DRB causes premature termination without affecting initiation of transcription. We tested this in transgenic mouse fetal liver cells using *in situ* hybridization with probes that hybridize at different distances relative to the site of initiation of the β globin primary transcript (Figure 1B and 2). *In situ* signals with probes that hybridize to intron 1, located in the first 300 bases of the β globin primary transcript, are still visible in 85% of the erythroid cells after 15 minutes of DRB treatment (Figure 2A, C) when compared to the untreated control. Probes that hybridize to intron 2, 600-1200 bases 3' of the initiation site (Figure 1B) have completely disappeared after 7.5 minutes of DRB treatment (Figure 2B, C). The fact that intron 1 signals are not affected by DRB confirms earlier reports which indicated that the process of transcription initiation is not disturbed but only elongation is affected resulting in short, prematurely aborted transcripts. Since initiation continues and the balance between γ and β is maintained we conclude that whichever mechanism (scanning, accessibility or looping) is responsible for the activation of the genes, it is not disturbed by the addition of DRB. The results also show that after inhibition of transcription by DRB the time required to decay the existing β globin intron 2 primary transcript signal via splicing, to levels beyond the limit of detection is in good agreement with the Actinomycin-D results (Figure 1C). The same is found for the α-like genes (not shown).

Kinetic analysis of single and double primary transcript signals

Primary transcript *in situ* hybridization with gene-specific intron probes for human γ and β globin in transgenic 11.5 day fetal liver cells containing a single copy of the complete human β globin locus show single gene transcription signals in approximately 85% of the human globin loci (Wijgerde et al., 1995). A small percentage of loci (~15%, e.g. Figure 3A, bottom right) contain signals for both γ and β globin genes *in cis* and it is these signals that make a distinction between the different mechanisms difficult. If double signals are due to simultaneous initiation of the γ and β genes according to the scanning and (on a random basis) accessibility models then release of the DRB block should result in the reappearance of double signals (using intron 2 probes) at the same rate as single signals. A lag in the reappearance of double signals compared to single gene signals would be indicative of alternating single gene initiation. This model predicts that double signals would result from the overlap between decaying primary transcripts (7.5 minutes, see above) from a recently active gene and the nascent transcription of an active gene.

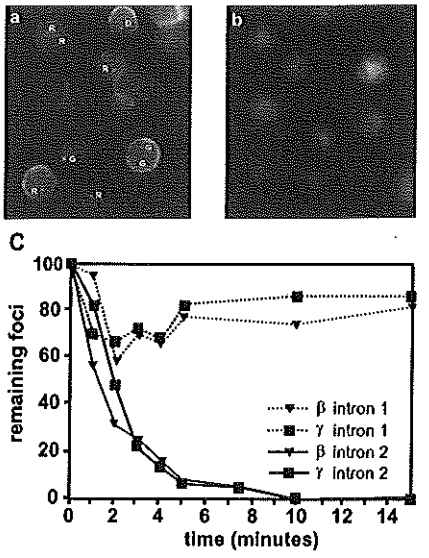


Figure 2. The effect of DRB on transcriptional elongation in vivo (A) *In situ* hybridization with intron 1 specific probes on homozygous transgenic 11.5 day fetal liver cells treated for 15 minutes with 100 μM DRB. γ globin signals are shown in red (Texas red), and β globin signals are green (FITC). (B) *In situ* hybridization with intron 2 specific probes on homozygous transgenic 11.5 day fetal liver cells treated for 7 minutes with 100 μM DRB. γ globin signals are red (Texas red), and β globin signals are green (FITC). (C) Decay of γ and β intron 1 and intron 2 *in situ* signals plotted as a function of time after the addition of DRB.

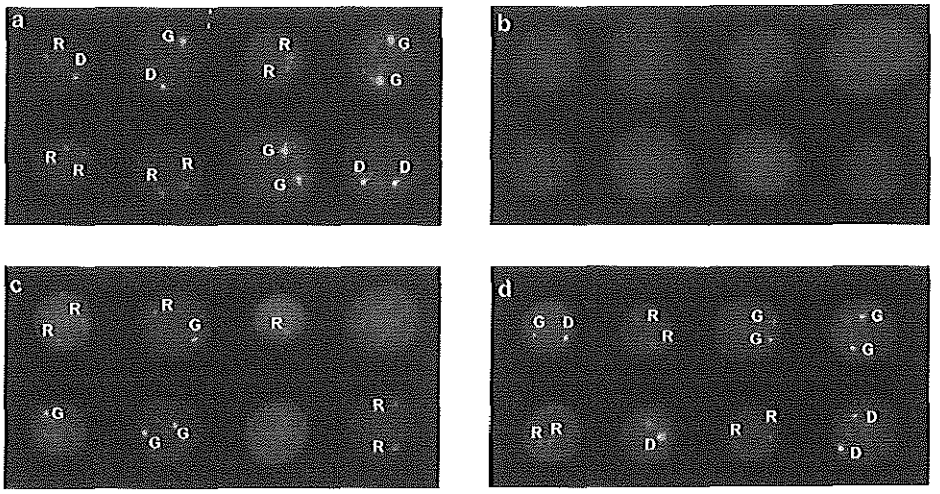


Figure 3. Reversible inhibition of transcription elongation on homozygous 11.5 day transgenic fetal liver cells. (A-D) Primary transcript *in situ* hybridizations using γ and β intron 2 probes, γ signals are red (Texas red) and β signals are green (FITC). (A) γ and β primary transcript signals prior to DRB treatment. (B) γ and β primary transcript signals after 15 minutes treatment with 100 μM DRB. (C) 5 minutes after release of the transcriptional elongation block by washing out DRB. Note most cells show single γ or β transcription signals *in cis*, while many have more than one signal *in trans*. (D) 20 minutes after release from the elongation block, the distribution of loci having single γ or β and double signals (γ and β *in cis*) are back to control levels. Representative cells are shown for each time point.

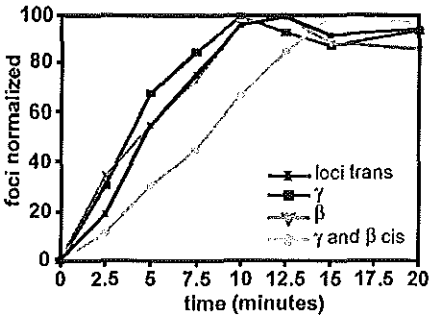


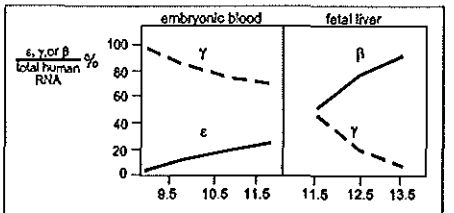
Figure 4. Kinetics of reappearance of single versus double gene signals. Primary transcript *in situ* signals for γ and β (see Figure 3) were scored after release of the DRB block to transcription elongation in 11.5 day fetal liver cells and plotted versus time. The curves were normalized to their maximum values. The results show the early reappearance of single γ or β signals and a clear lag in reappearance of double (γ and β *in cis*) signals. Cells in which both homologues display signals for either γ or β are plotted as loci *trans* showing the independent reappearance of two signals in the same cell.

We treated transgenic 11.5 day homozygous fetal liver cells with DRB for 15 minutes to block elongation of globin primary transcripts (Figure 3). The cells were then released from the DRB block by washing with PBS and aliquots of cells were fixed onto slides at various intervals and prepared for *in situ* hybridization as previously described (Wijgerde et al., 1995). The cells were probed with intron 2-specific probes for human γ (probes detect both $^{G}\gamma$ and $^{A}\gamma$ primary transcripts) and β globin primary transcripts. No transcription signals are visible in the zero time point slide immediately after the wash (Figure 3B). Five minutes after washing out the DRB (Figure 3C) primary transcript signals are again detectable in a high proportion of cells and continue to rise to the level observed before addition of DRB. Single (γ or β) and double (γ and β *in cis*) gene signals were counted and the averaged results of three separate experiments are presented in Figure 4. The results show that the reappearance of double gene signals significantly lags behind the reappearance of single gene signals in the population. A lag in the reappearance of double signals is also consistently observed when the two γ genes are compared with gene-specific probes for $^{G}\gamma$ and $^{A}\gamma$ (data not shown).

To determine whether there is an intrinsic bias against the reappearance of double signals we counted the reappearance of two signals *in trans* (i.e. a signal on each chromosome) as an internal control. Their reappearance (Figure 4; loci trans) closely approximates the single signal curves. We also calculated a theoretical curve that would be expected when the two loci are activated independently *in trans* (not shown). This curve coincides with the curve found for the appearance of two signals *in trans*. We therefore conclude that the two allelic globin loci behave independently from each other in terms of transcription (showing a stochastic reappearance of two signals without a lag) whereas the genes within a locus on one chromosome do not behave independently (double signal reappearance is non random with a lag).

Weak competition versus strong competition

The γ genes are being expressed in embryonic cells in conjunction with ε and in the early fetal liver cells in competition with β. This provides a unique opportunity to investigate the interdependence of genes within a locus by examining the effects of weak versus strong gene competition during development independent of a DRB treatment. The ε-gene, which is a weak competitor (Wijgerde et al., 1995), increases in expression from 9.5 to 11.5 days in the embryonic blood and γ-gene expression decreases reciprocally (Strouboulis et al., 1992; see Table 1). In the early fetal liver erythroid cells γ expression is further decreased due to strong competition from the β gene (Wijgerde et al., 1995, 1996). If co-initiation of transcription occurs one would predict that the percentage of G_{γ}-A_{γ} double signals in the γ expressing cells would change very little. In contrast if alternating transcription of the γ genes and the other globin genes (ε or β) occurs then addition of a third gene into the competition should affect the percentage of double γ signals due to the resulting three-way alternation. When the percentage of G_{γ}-A_{γ} double signals is measured in the γ expressing cells during development, it changes from 83% at day 9.5 when ε expression is low to 56% at day 11.5 in embryonic blood when ε expression is at its maximum (Table 1). In the fetal liver where the γ genes are expressed with high levels of β expression, the percentage of double G_{γ}-A_{γ} signals decreases further to 29% (Table 1). Thus when γ gene expression is accompanied by relatively low level expression of the ε gene in embryonic cells or high level expression of the β gene in fetal liver cells, the percentage of G_{γ}-A_{γ} double signals decreases accordingly. This result is most easily explained by an alternating single gene mechanism and is difficult to explain if one assumes that the γ genes are co-initiated.



	% in situ primary γ RNA signals			% total γ RNA	
	Gγ	Gγ/Aγ	Aγ	Gγ	Aγ
day 9.5 blood	5	83	12	39	61
day 10.5 blood	8	78	14	31	69
day 11.5 blood	7	66	37	27	73
day 11.5 liver	23	29	48	26	74

Table 1. Developmental expression and transcription of the human γ genes.
The top panel shows the relative expression of the ε, γ and β genes during development in transgenic mice (Strouboulis et al., 1992a). Shown below are the percent of loci with single Gγ or Aγ signals and double primary transcript *in situ* signals during development. The bottom right shows the percent of Gγ and Aγ m RNA as determined by primer extension (data not shown). Note that changes in steady state RNA levels lag behind changes in transcription as detected by primary transcript *in situ* analysis.

Transcription initiation of the α-like globin genes

The DRB analysis above does not exclude the possibility that treatment with the drug could somehow artifactually cause a delay in the reappearance of double signals *in cis*, even though this is clearly not the case for the reappearance of double signals *in trans* (Figure 4). There is indirect evidence that the α-globin genes may be regulated differently than the β-like genes (Craddock et al., 1995). Instead of the five hypersensitive regions present in the β globin LCR, only a single hypersensitive site has been identified 40kb or 26kb upstream of the human and mouse α genes respectively (Gourdon et al., 1994, 1995). When this site is present as part of the human α locus in transgenic mice, the expression of the α genes is suppressed as development proceeds (Sharpe et al., 1993; Gourdon et al., 1994) indicating that additional sequences are required for full expression.
We examined the primary transcription of the mouse α-like genes in embryonic red cells where all three α genes are expressed. Primary transcript *in situ* analysis shows a high percentage of double ζ and α signals *in*

cis (80%) versus single ζ or α signals (Figure 5A), while the lifetime of the signal is similar to that observed for the β signals (Figure 1D) in the presence of actinomycin-D as well as DRB (data not shown). This could mean that the genes are co-initiated in many loci or that frequently alternating initiation takes place. If alternating initiation occurs well within the lifetime of the signal a large proportion of double signals *in cis* would result. We then measured the reappearance of single and double signals after the addition and removal of DRB (Figure 5B, C and D). Plotting the reappearance of the signals as a function of time (Figure 6) shows that the double signals appear at the same rate as the single signals and that there is no measurable conversion of single signals into double signals. On the basis of these data it is tempting to suggest that the α-like genes are co-initiated, but rapidly alternating initiation of transcription would give a similar result and hence can not be excluded. Importantly the result shows that the treatment with DRB does not artificially result in a lag in the appearance of double signals *in cis*.

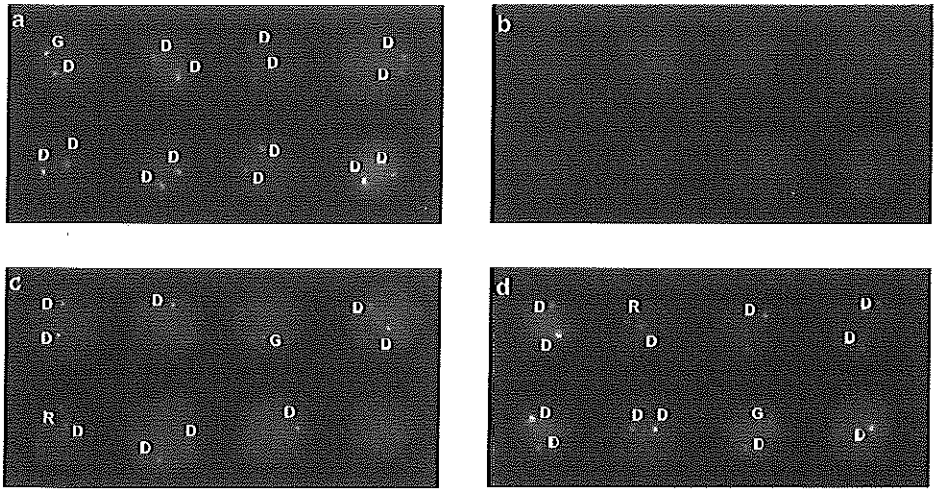


Figure 5. Reversible inhibition of transcription elongation on 10.5 day blood cells. (A-D) Primary transcript *in situ* hybridizations using mouse ζ and α intron probes, ζ signals are shown in red (Texas red) and α signals in green (FITC). (A) ζ and α signals prior to DRB treatment. (B) 15 minutes after treatment with 100 μM DRB. (C) 5 minutes after release of the transcriptional elongation block by washing out DRB. Note that many loci show double ζ- and α-globin transcription signals *in cis*. (D) 20 minutes after release of elongation block, the distribution of loci having single ζ or α and double signals (ζ and α *in cis*) are back to control levels. Representative cells are shown for each time point.

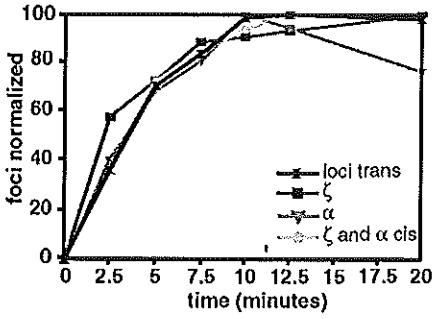


Figure 6. Kinetics of reappearance of single versus double ζ and α primary transcript signals. *In situ* signals as shown in Figure 5 were scored after release of the transcriptional elongation block in 10.5 day embryonic blood cells and normalized to their maximum values for single (ζ or α) and double (ζ and α *in cis*) signals. Cells in which both homologues display signals for either ζ or α are plotted as loci *trans* showing the independent reappearance of two signals in the same cell.

DISCUSSION

The results obtained previously with primary transcript *in situ* analysis suggested that the human β globin genes are regulated via a dynamic process of alternating initiation of transcription of the different genes. That proposal appeared to contradict existing data. It had been known for many years that γ and β mRNA and proteins are found in the same cell during the switchover from γ to β, even in heterozygous, and this appeared to be confirmed by single cell PCR data (Furukawa et al., 1994). These data, taken together with the primary transcript *in situ* results which revealed the presence of a minority of loci with double primary transcription signals *in cis*, were interpreted to mean that the γ and β globin genes within a single locus were or could be co-initiated (Martin et al., 1996; Bresnick et al., 1994; Furukawa et al., 1994).

Three models have been proposed to explain transcriptional regulation by distant regulatory sequences. The accessibility model proposes that after the activation of the chromatin, the genes bind transcription factors and are activated in a stochastic fashion (Groudine and Weintraub, 1982; Martin et al., 1996). The scanning model suggests that regulatory sequences act as a nucleation site for (part of) the transcription machinery, which subsequently scans the DNA for transcriptionally competent genes (Herendeen et al., 1992; Tuan et al., 1992). Finally the looping model postulates that the distant regulatory sequences interact directly with the gene to initiate transcription (see Ptashne, 1988 and Wijgerde et al., 1995 and ref. therein). Both the accessibility and the scanning models allow co-initiation of multiple genes *in cis* whereas the looping model predicts initiation of only one gene at a time. It is therefore important in understanding the process of transcriptional initiation *in vivo* involving distant regulatory sequences to distinguish whether alternating single or multiple co-initiation takes place.

We have shown that in the presence of DRB *in vivo*, initiation of transcription is unaffected whereas transcriptional elongation is reversibly blocked. Inhibition studies with both DRB and actinomycin-D show that the globin primary transcript signals have a maximum lifetime of about 7 minutes confirming that we are detecting actively transcribed genes or genes transcribed within the last 7 minutes. By reversibly inhibiting elongation with DRB for 15 minutes we have allowed intron 2 primary transcript signals to decay via splicing and degradation. Removal of the block theoretically permits the immediate detection of only actively transcribed genes. Obviously not all loci cross the detection threshold simultaneously, introducing a small but measurable degree of asynchrony to the experiment, and therefore allowing the possibility of the reappearance of some double signals in the earliest timepoint. By analyzing large populations of cells at various time points after removal of the elongation block we have been able to demonstrate a clear and consistent difference in the amount of time required to detect single versus double signals. If we compare the $t_{1/2max}$ (the time required to re-establish 50% of the maximum number of single or double foci) for single (γ or β) and double (γ and β) signals we find 4 minutes and 8 minutes respectively. This shows that on average double signals take twice as long to reappear as single signals indicating that double signals result from sequential periods of initiation of single genes.

In contrast, in the α locus the $t_{1/2max}$ of the single and double signals are essentially the same (approx. 3 minutes) suggesting that the α genes are either co-initiated or that they alternate at a frequency which is shorter than the time required to reach the detection threshold. There is indirect evidence for competition in the mouse α locus which may be indicative of an alternating transcriptional mechanism. Insertional mutation of the ζ gene with a PGK-Neo cassette results in decreased expression of the α genes in definitive erythroid cells when ζ would normally be silenced (Leder et al., 1997). Although the DRB results alone do not allow us to make firm conclusions regarding the mechanism of multiple α-gene expression they do show a clear difference compared to the γ and β genes and exclude an artificial lag of reappearing double signals *in cis* as a consequence of the DRB treatment.

Restriction digestion of sites in the $^{G}\gamma$ and $^{A}\gamma$ promoter regions in isolated K562 nuclei indicated that both promoters were accessible to digestion in cis in approximately 50% of the loci (Bresnick et al., 1994). Co-accessibility of the promoters was inferred to be synonymous with simultaneous nuclease hypersensitivity of the two gene promoters. Although these links have not formally been established, it was suggested that this was an indication of co-initiation of transcription of the genes. Our DRB results in conjunction with the developmental transcription analysis indicate that the two γ genes are alternately transcribed. The high percentage of double $^{G}\gamma$-$^{A}\gamma$ signals in comparison to the low percent of double γ-β signals suggests that alternations occur at a higher frequency between the two γ genes than between the γ genes and the β gene.

Thus the data presented show that the double signals in the β locus are not generated as predicted by a co-initiation mechanism, but represent a transition between alternating transcriptional periods of single genes. A number of additional *in vivo* observations support this conclusion. A single regulatory region (the LCR) is required by all of the genes in the β globin

locus for activation (reviewed in Fraser et al., 1998) and the genes compete with each other for this function (Giglioni et al., 1984; Enver et al., 1990; Behringer et al., 1990; Hanscombe et al., 1991; Dillon et al., 1997) with LCR proximal genes having a competitive advantage over distal genes (Hanscombe et al., 1991; Peterson and Stamatoyannopoulos, 1993; Dillon et al., 1997). During the period of switching from γ to β expression nearly all erythroid cells have both γ and β mRNA in the cytoplasm yet the overwhelming majority of loci have only γ or β transcription signals (Wijgerde et al., 1995). In addition, individual loci within the same cell can respond differently to the same trans-acting factor environment with γ transcription on one homologue and β on the other. The balance of expression between the γ and β genes can be tipped in either direction by mutations in the γ promoter which prevent normal γ gene silencing or alterations in the level of EKLF, which is required for β gene transcription (Wijgerde et al., 1996, Tewari et al., 1998). In each case modulation of the expression of one gene leads to reciprocal changes in the expression of the other. A dynamic interaction between the LCR and the genes via looping explains all of the basic properties. It explains competition, because it predicts that the time taken up by LCR driven transcription of one gene takes time away from another gene. In this model the competitive advantage of a gene is the result of increased frequency of LCR-gene interactions which are dependent on distance (see Dillon et al., 1997).
Looping and direct contact between regulatory regions is therefore the simplest mechanistic explanation for the observed results and is supported by previous experiments with other systems (Mueller-Storm et al., 1989; Bickel and Pirotta, 1990; Dunaway and Dröge, 1989). It implies that direct chromatin interactions between the LCR and a single gene are required for initiation of transcription. Continued loading of polymerases or re-initiation of that gene would require continuous LCR contact (Wijgerde et al., 1995; Milot et al., 1996; Dillon et al., 1997). In the context of this mechanism the data suggests that chromatin *in vivo* is highly dynamic or diffusible, allowing the LCR-gene complex to change rapidly to bring about co-expression of multiple genes.

Materials and Methods

Reversible inhibition of transcription elongation with DRB.

Homozygous transgenic mice containing a single integrated copy of the complete human β globin locus were bred to obtain embryos. Peripheral blood and fetal livers from the indicated developmental time points were dissected out in PBS. Fetal livers were gently disrupted by repeated pipetting. Actinomycin- D was used at a final concentration of 5 μg/ml. DRB (Sigma) was added to cell suspensions to a final concentration of 100 μM and incubated at 37°C. for 15 minutes. Five volumes of ice cold PBS were added and the cells were immediately pelleted by centrifugation for 2 minutes at 1500 rpm in an eppendorf centrifuge. Cells were washed two more times with 1.5 ml ice cold PBS and resuspended in 250 μl of PBS at 25°C. Aliquots were taken at the designated intervals and fixed onto poly-L-lysine coated slides (Sigma) for *in situ* hybridization.

Probe sequences and in situ *hybridization analysis*

The following probes were used for the *in situ* hybridization analysis:
Human β intron 1 probes:
CTGTCTCCACATGCCCAGTTTCTATGGTCTCCTT
AAACCTGTCTTGTAA
GGGTGGGAAAATAGACCAAAGGCAGAGAGAGT
CAGTGCCTATCAGAAACAGGGCAGTAACGGCA
GACTCTCCTCAGGAGTCAGGT
ATAACAGCATCAGGAGGGACAGATCCCCAAAG
GACTCA
Human β intron 2 probes:
TTCCACACTGATGCAATCATTCGTCGTTTCCCA
TTCTAAACTGTACCCT
CTGATTTGGTCAATATGTGTACACAATTAAAAC
ATTACACTTTAACCCA
5'GGTAGCTGGATTGTAGCTGCTATAGCAATATG
AAACCTCTTACATCAGT
Human γ intron1 probes:
AGGCACAGGGTCCTTCCTTCCCTCCCTTGTCCT
GGTCAC
TGACAAGAACAGTTTGACAGTCAGAAGGTGCC
ACAATCCTGAGAAGCGA
AGGCTTGTGATAGTAGCCTTGTCCTCCTCTGTG
AAATGACCCA
AGAGCCTACCTTCCCAGGGTTTCTCCTCCAGCA
TCTTCCACATT
Human γ intron2 probes:
GCAGTTTCTTCACTCCCAACCCCAGATCTTCAA
ACAGCTCACACCCGC
CCTTCTGCCTGCATCTTTTTAACGACCAACTTG
TCCTGCCTCCAGAAG
ACAGAGCTGACTTTCAAAATCTACCCAGCCCAA
ATGTTTCAATTGTCC
Human $^{A}\gamma$ 3' flanking region probes:
TCATATAAAAATAAATGAGGAGCATGCACACA
CCACAAACACAAACAGGC
CAGAACTCCCGTGTACAAGTGTCTTTACTGCTT
TTAT
TTCATTAAGAACCATCCTTGCTACTAGCTGCAA
TCAATCCAGCCCCCA

ATTTCACTTTCTTAGGCATCCACAAGGCTGTGA
AAAGCTAAGTGCCAT

Human $^{G}\gamma$ 3' flanking region probes:
AAAAAAGTGTGGAGTGTGCACATGACACAAAC
ACACATAGCCATGTATAA
TGCAGACGCTCCCATGTATAAGTTTCTTTATTG
CCTAGTTCTTTTATTTG
ACGTAAACAAAAAAGTGTGGAGTGGCACATGA
CACAAACACACATAG
GCAGACGCTCCCATGTATAAGTTCTTTATTGCC
TAGTTCTTTTATTT
Mouse α intron probes:
CACAGAAAAGCATAGTTAGAAGCGCCCACTGA
GCGAGTGCCAGGTCC
AGCCCTTCCTAGGGGCCCAGATGCCGCCTGCCA
GGTCCC
GCTCCCCTTCCTGGGACCACTATGTCCCTGCCT
TGGGCACGAGGACCC
Mouse ζ intron probes:
CCTTCTCAGTGGCTTCTCCTCACAACTGCTCTTT
GTCACTTCTGTCTC
ATGGAAGACTCTGGTGAGCTCTGGAATGCCAG
CCCACCTCCTTTAGTA
ACAACCCCAAGAGTGATGTTACTATTGCTGTTG
CACAAGGGTCTACA
AAGGGGATTTGATGCCTCCAGCCCCAATGGCA
CCCATGCCTGCGCTCG

The two γ genes are highly homologous in both intron and exon sequence making the use of gene-specific intron probes extremely difficult. Transcriptional termination of γ gene transcription is known to occur 1-2 kb downstream of the poly-adenylation site (Ashe et al., 1997). We therefore used probes (3' flanking region probes) which hybridize to regions 300-500 bp downstream of the poly-adenylation sites of the γ genes to detect the $^{G}\gamma$ and $^{A}\gamma$ gene primary transcripts separately.

Cells were fixed onto poly-L-lysine coated slides in 4% formaldehyde/5% acetic acid for 18 min. at room temp. The cells were subsequently washed 3 times for 5 min in PBS and stored in 70% ethanol at -20°C. The slides were pretreated for hybridization by a 0.01% pepsin digestion (5 min., 37°C) in 0.01M HCl, followed by a short wash in water and a 5 min. fixation in 3.7% formaldehyde at room temp. The slides were washed in PBS, dehydrated in 70%, 90% and 100% ethanol steps and air dried. The hybridization mixture was applied (12 µl per 24x24mm coverslip) and incubated at 37°C in a humidified chamber for 12hr. The hybridization mixture contained 1ng/µl of each oligonucleotide probe haptenized with either digoxygenin or biotin side chains in the middle and on the 5' and 3' ends of the oligonucleotide (Eurogentec, Belgium) in 25% formamide, 2x SSC, 200µg/ml salmon sperm DNA, 5x Denhardts, 1mM EDTA and 50mM sodium phosphate pH 7.0. The coverslip was removed by dipping in 2x SSC and the cells were washed three times for 10 min. in 2x SSC at 37°C, followed by a 5 min wash in 0.1M Tris, 0.15M NaCl, 0.05% Tween 20 at room temp. Antibody detection of the labels was essentially as described by Dirks et al. (1993), with three or four amplification steps. Mounting was in DAPI/DABCO:Vectashield (1:1) in glycerol (90%) and stored at 4°C in the dark. Fluorescence was detected by epifluorenscence microscopy and photographs recorded with an CCD camera.

The graphical results represent the average of three separate experiments. In all cases more than 1000 cells were counted per data point using a dictaphone to record the results.

Acknowledgements: We wish to thank Dr. M. Antoniou for advice on DRB and Act.D and Drs. T. McMorrow and S. Philipsen for helpful discussions and L. Braam for animal care. J.G. was supported by the Jeantet Foundation and E.M. by the MRC Canada. This work was supported by the NWO Netherlands.

Chapter VI

Mechanisms of developmental control of transcription in the murine α and β globin loci

Submitted for publication

Mechanisms of developmental control of transcription in the murine α and β globin loci.

Tolleiv Trimborn, Joost Gribnau, Frank Grosveld and Peter Fraser

MGC Department of Cell Biology, Erasmus University, P.O.Box 1738, 3000 DR, Rotterdam, the Netherlands.

We have characterized mRNA expression and transcription of the mouse α and β globin loci during development. S1 nuclease and primary transcript *in situ* hybridization analyses demonstrates that all seven murine globin genes (ζ, α1, α2, εy, βH1, βmaj and βmin) are transcribed during primitive erythropoiesis, however transcription of the ζ, εy and βH1 genes is restricted to the primitive erythroid lineage. Transcription of the βmaj and βmin genes in primitive cells is EKLF-dependent demonstrating EKLF activity in embryonic red cells. Novel kinetic analyses suggest that multi-gene expression in the β-locus occurs via alternating single gene transcription whereas co-initiation cannot be ruled out in the α locus. Transcriptional activation of the individual murine β genes in primitive cells correlates inversely with their distance from the locus control region which contrasts with the human β-locus in which the adult genes are only activated in definitive erythroid cells. The results suggest that the multi-gene expression mechanism of alternating transcription is evolutionarily conserved between mouse and human β globin loci but that the timing of activation of the adult genes is altered, indicating important fundamental differences in globin gene switching.

The murine α and β-globin loci are multi-gene clusters located on chromosomes 11 and 7 respectively (Fig. 1). They are highly homologous to their human counterparts in organizational structure and function and represent paradigms for the study of developmental gene regulation. The alpha locus consists of three genes ζ, α1 and α2 which are dependent for expression on the major regulatory element (αMRE) which appears as an erythroid-specific DNase I hypersensitive site approximately 26 kb upstream of the ζ gene (Gourdon et al., 1995). The murine β locus consists of four functional genes εy, βH1, β-major (βmaj) and β-minor (βmin) which are controlled by the locus control region (LCR), a series of five DNase I hypersensitive sites in erythroid chromatin located 5-25 kb upstream of the εy gene (Moon and Ley, 1990; Hug et al., 1992; Jimenez et al., 1992; Fiering et al., 1995). The genes are arranged in the order of their developmental expression, as are their human homologues. Embryonic yolk sac derived erythroid cells co-express high levels of both ζ and α globin mRNA (Leder et al., 1992) and primarily εy and βH1 beta-like globin mRNA with small amounts of βmaj and βmin (Brotherton et al., 1979; Chui et al., 1979; Wawrzyniak and Popp, 1987; Whitelaw et al, 1990). At 11.5 days of gestation the major site of erythropoiesis in the developing embryo switches from the yolk sac to the fetal liver. This switch in site is coincident with a change to definitive gene expression in both the α and β clusters leading to predominant expression of the α1 and α2 genes and the βmaj and βmin genes. Although the small amount of βmaj and βmin expression in embryonic cells appears to be genuine and not due to maternal contamination (Wawrzyniak and Popp, 1987) it is unclear whether the embryonic genes are expressed in early fetal liver cells (Wong et al., 1983; Whitelaw et al., 1990).

The human globin loci have been more thoroughly studied, facilitated by the use of transgenic mice. The β LCR has been shown to be required for the initial activation of the locus and provides erythroid-specific, high level, copy number dependent, position independent expression to linked genes (Grosveld et al., 1987). Studies with the α locus have shown that HS-40 is required for expression of the α-like genes (Bernet et al., 1995) but transgenic results have suggested that additional sequences are required for developmental position independent expression (Higgs et al., 1990; Sharpe et al., 1992).

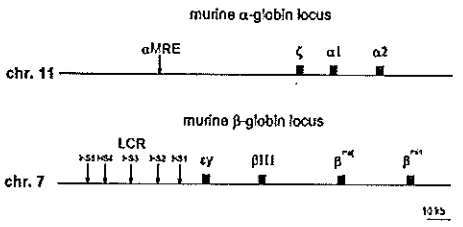


Figure 1. Schematic diagram of the murine α globin and β globin loci. Genes are represented by solid boxes and vertical arrows represent DNase I hypersensitive sites of the αMRE and β globin LCR.

A detailed model of the mechanism of developmental regulation of the β-like genes has been proposed based on the observations of gene competition for LCR function (Enver et al., 1990, Hanscombe et al., 1991, Peterson and Stamatoyannopoulos, 1993, Wijgerde et al., 1995, 1997; Dillon et al., 1997) and single gene activation by the LCR (Wijgerde et al. 1995; Gribnau et al., submitted). These data strongly argue in favor of a looping model in which the LCR forms direct, mutually exclusive chromatin interactions with gene-local regulatory elements to activate transcription of a single gene. Multiple genes are co-expressed in *cis* through alternating transcriptional periods of single genes (Wijgerde et al., 1995; Gribnau et al., submitted) suggesting that LCR/gene interactions are dynamic but also semi-stable, persisting on the order of several minutes (Wijgerde et al., 1995; Dillon et al., 1997).

It has therefore been suggested that two parameters determine the transcriptional output and hence expression level of a given gene during development. The first is the frequency with which the LCR productively contacts a particular gene and the second is the stability of that interaction. The frequency of LCR-gene contact has been proposed to be dependent on distance from the LCR (Dillon et al., 1997). The relative distance between two competing genes and the LCR has been shown to be important in controlling both the level and timing of expression (Enver et al., 1990, Hanscombe et al., 1991, Peterson and Stamatoyannopoulos, 1993). Dillon et al., (1997) measured the effects of distance on the frequency of LCR-gene interaction by comparing genes of equal stability at varying positions in the locus in combination with primary transcript *in situ* hybridization. Insertion of a β gene into more LCR-proximal positions resulted in that gene being transcriptionally activated more often and at the expense of the equivalent downstream gene in relation to the difference in distance.

The stability of the LCR-gene interaction has been proposed to be determined by the transcription factor environment. Targeted disruption of the erythroid Kruppel-like factor (EKLF) has shown that it is required for transcription of only the adult-type β globin genes (Nuez et al., 1996; Perkins et al., 1996). EKLF binds selectively with high affinity to the CCACCC element present in the promoters of the mouse and human adult-type β globin genes (Donze et al., 1995). Studies with compound human globin locus transgenic/EKLF knockout mice have shown that reductions in EKLF expression in heterozygous and homozygous knockout mice lead to decreased expression of β globin and reciprocally increased expression of γ globin mRNA (Wijgerde et al., 1996; Perkins et al., 1996). We have shown that these changes are due to reductions in the number of transcriptionally active β-genes in the fetal liver population with reciprocally increased numbers of active γ-genes (Wijgerde et. al., 1996). These studies have been interpreted to suggest that reduced EKLF levels lead to a decrease in the stability of the LCR/β gene complex. Reduction in the amount of time that the LCR complexes with the β gene allows more frequent interaction with the γ genes.

Here we present detailed characterization of the developmental expression and transcriptional regulation of the murine α and β globin loci at the single cell level. The results of *in situ* hybridization and novel kinetic analyses suggest that transcriptional regulation of the mouse β genes is mediated by a similar dynamic chromatin interaction mechanism as has been proposed for the human β locus. However, unlike the human locus transcription of the adult β-like genes occurs in embryonic cells and hence is only partially suppressed and not silenced through competition for the LCR. Similar analyses of the α-globin locus indicate that a co-activational mechanism of multiple genes in *cis* cannot be ruled out.

RESULTS

Developmental expression.

RNA samples were collected from embryonic and fetal erythroid tissues at different stages of development and subjected to S1 nuclease protection assay to determine the expression pattern for the murine α and β globin genes (Fig. 2). Previous studies have suggested that small amounts of the adult β-like globin genes are expressed in embryonic cells (Brotherton et al., 1979; Chui et al., 1979; Wawrzyniak and Popp, 1987; Whitelaw et al, 1990). Quantitative phosphorimage analysis of S1 assays (Fig. 2A) shows that in 10.5 day embryonic blood the level of εy is 60 % and βH1 34 % of total β-like globin (Table 1). βmaj-βmin (the β S1 probe does not distinguish between βmaj and βmin) is detected at levels of approximately 6%.

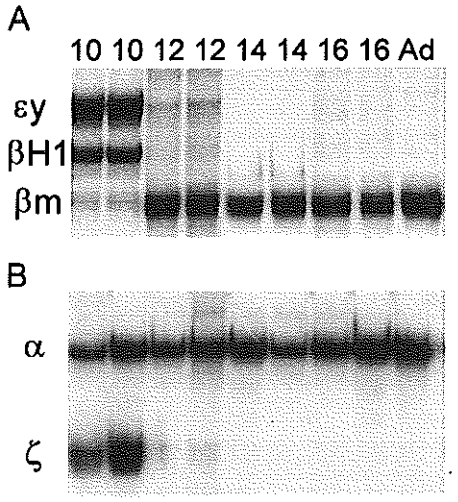


Figure 2. S1 nuclease protection assays. Total RNA from the indicated developmental timepoints (10.5 day whole embryo and 12.5, 14.5, 16.5 day fetal liver and adult blood) was subjected to S1 nuclease protections assay as described in materials and methods with probes of equal specific activity. A. β locus probes εy, βH1 and βm (βmaj and βmin). B. α locus probes ζ and α (α1 and α2). The position of protected fragments is indicated on the left.

Expression of the α genes at 10.5 days is comprised of 46% ζ and 54 % α mRNA (the α S1 probe does not distinguish between α1 and α2) (Fig. 2B and table 1). This is in agreement with previously published results (Whitelaw et al, 1990). At day 11.5 of gestation the main site of erythropoiesis changes from the embryonic yolk sac to the fetal liver which is the site of definitive erythropoiesis. Analysis of fetal liver RNA at 12.5 day shows that 91% of the β-like mRNA is βmaj-βmin, and εy and βH1 are still detectable at 7% and 2% respectively. The appearance of small amounts of βH1 and especially εy mRNA in the early fetal liver as late as 14.5 days has been suggested to indicate that expression of these genes is not restricted to yolk sac derived cells (Whitelaw et al., 1990). However this residual mRNA may also be due to contamination of fetal livers with circulating embryonic cells. A definitive conclusion is only possible through individual cell analyses with cell type markers and morphological analysis (see below). The levels of α globin mRNA in 12.5 day fetal liver cells also shows a switch from roughly equal ζ and α expression in the embryonic cells to 99% α and only 1% ζ mRNA. ζ mRNA is undetectable by 14.5 days suggesting but not proving that ζ transcription is restricted to embryonic cells.

Table 1: Quantitative S1-nuclease protection analysis.

	10.5 blood	12.5 fetal liver	14.5 fetal liver	16.5 fetal liver
εy	60	7	<1	0
βH1	34	2	<1	0
βmaj-βmin	6	91	99	100
ζ	46	1	0	0
α1-α2	54	99	100	100

Phosphorimager quantitation of S1 nuclease protection assays shown in Fig 2. The numbers represent the percentage of total β globin or α globin mRNA expression for the individual mouse β and α genes.

Primary transcript in situ hybridization

We have shown previously that co-expression of the human γ and β-globin genes in transgenic mouse fetal erythroid cells containing a single, complete human β-globin locus (Strouboulis et al., 1992) is achieved through alternating transcriptional periods of individual genes (Wijgerde et al., 1995). Primary transcript *in situ* hybridization with gene-specific intron probes showed that the vast majority of loci (88%) had only single gene signals. Approximately 12% of loci displayed two gene signals in *cis*, which we proposed represented a recent switch in gene activation. To ensure significance in this type of analysis the hybridization efficiency must be extremely high to guarantee the detection of nearly all transcriptionally active genes.

We determined the hybridization efficiency of the mouse ζ, α, εy and βH1 globin intron probes by hybridizing them to 10.5 day primitive erythroid cells (Fig. 3A,B and Table 2). The ζ probes detect primary transcript signals at 96% of the loci. The α globin intron probes which recognize the primary transcripts of both highly homologous α1 and α2 globin genes demonstrate detectable signals at 90% of the loci. Double label experiments with both the ζ and α globin probes shows that a signal (α, ζ, or both) is detected at greater than 95% of the loci. If all three α globin genes are constitutively transcribed in primitive cells then our hybridization efficiencies are very high, at least greater than 90%. It is possible that some of the genes are off at certain times, in this case we would conclude that our efficiency is even higher.

We performed similar quantitations for the εy and βH1 globin primary transcript probes (Fig. 4B, Table 2). Approximately 90% of the loci have a εy signal and 63% have a βH1 signal at 10.5 dpc. Calculation of the relative percent of εy versus βH1 primary transcript signals yields the same percentages as the S1 analysis of εy and βH1 mRNA (59 and 38 percent of total β-like globin respectively, Table 1). This precise correlation between the relative percentages of transcriptionally active genes and mRNA expression suggests that the number of fully active genes in the population determines the level of mRNA expression. This was also found to be the case in the human β locus where the results indicated that a gene is either fully on or off (Wijgerde et al., 1995; 1996; Dillon et al., 1997).

Table 2. Quantitation of transcriptional cell types and percenatges of loci with transcription signals for the individual globin genes.

% Transcriptional cell types [a]						% loci with specific transcription signal [b]					
	○ εy βH1 ●	●ζ αO	○βmaj βmin ●			εy	βH1	ζ	α	βmaj	βmin
●	1	<1	1			0	0.5	0.5	0	0	0.5
● ●	<1	3	<1			0	1	3	0	0	1
● ●*	4	13	3			2	4	13	6.5	1.5	3
● ●	37	76	69			37	37	76	76	69	69
● ○*	31	4	14			31	15.5	2	4	14	7
● ○*	3	2	<1			1.5	1.5	1	1	0.5	0.5
●	6	<1	8			3	3	0.5	0.5	4	4
○ ○	14	1	2			14	0	0	1	2	0
○	3	<1	2			1.5	0	0	0.5	1	0
				Total		90	63	96	90	92	85

a. Double label in situ hybridizations as shown in Fig 3 were counted and the percentages of the nine different transcriptional cell types are shown. The open and closed circles (left) represent the different primary transcript signals for the individual globin genes as indicated at the top of each column. The signals present on each of the homologous chromosomes are shown. Overlapping open and closed circles represent loci with two different primary transcript signals in cis. Results were taken from 10.5 day embryonic blood cells for εy, βH1, ζ and α. βmaj and βmin results are from 13.5 day fetal liver cells.
b. Shown are the percentages of loci with transcription signals for the indicated genes.

*Cells with different transcription signals or combinations of signals on each locus.

Globin gene transcription in primitive erythroid cells

Gene competition in the mouse β-globin locus has been suggested by a naturally occurring deletion of the βmaj gene and its gene-local regulatory elements (Skow et al., 1983) which leads to increased expression of the βmin gene in homozygotes (Curcio et al., 1986). In contrast, targeted disruption of the βmaj gene via insertion of a selectable marker leads to perinatal lethality with no increase in β-minor expression (Shehee et al., 1993;). In the α locus, marker gene insertion into the ζ-gene leads to reductions in the level of α-gene expression in the fetal and adult erythroid cells suggesting that the marker gene may be competing for activation with the α-genes (Leder et al., 1997). To investigate the mechanism of co-expression in the mouse α and β-globin loci and characterize the pattern of transcriptional regulation we performed primary transcript *in situ* hybridization, on 9.5, 10.5 and 11.5 day peripheral blood. Oligonucleotide probes specific for the introns of εy, βH1, βmaj, βmin, ζ and α were used in double label experiments. Primary transcript *in situ* hybridization shows that embryonic blood cells from day 9.5 through day 11.5 transcribe all four β-like globin genes to varying degrees (Fig. 4A-C, Table 2). Each of the loci in a single cell can have a different primary transcript signal or combination of signals (Table 2) demonstrating that, like the human β genes, each locus responds independently to the same *trans*-acting factor environment. Almost all loci (approx. 90%) have signals for the εy gene, which remains fairly constant in embryonic blood cells throughout the period from 9.5 to 11.5 days (Fig. 4A). The εy gene is closest to the LCR and is the most highly expressed globin gene during embryonic erythropoiesis (Fig. 2A). The εy and βH1 foci frequently appear in *cis* (i.e. double signals on one allele) decreasing from 85% at 9.5 days to 45% at 11.5 days of gestation (Fig. 3B, 4A) as a result of the decrease in βH1 gene transcription from 95% at 9.5 days to 53% at 11.5 days. Thus fewer βH1 genes are being transcribed as development proceeds.

Adult β-gene transcription is already observed in a small percentage of βmaj loci in 9.5 day embryonic blood, increasing to 25% in 11.5 day blood (Fig. 3C, 4B). Three percent of the βmin genes had primary transcript signals at 10.5 days (Fig. 3D) and increased to around 10% in 11.5 day blood. The βmaj and βmin foci nearly always appear in *cis* with an εy signal in the early primitive cells, but by 11.5 days (blood) approximately 20% of the βmaj signals are single signals (Fig. 4B).

These results of β gene transcription in the mouse globin locus are markedly different from those of the human transgene locus. The early transcription of the "adult" mouse β-genes (βmaj and βmin) is in contrast to the human locus in which transcription of the adult β globin gene is not detected in embryonic cells but is restricted to fetal derived erythroid cells. These results show that the entire murine β globin locus is activatable in embryonic erythroid cells.

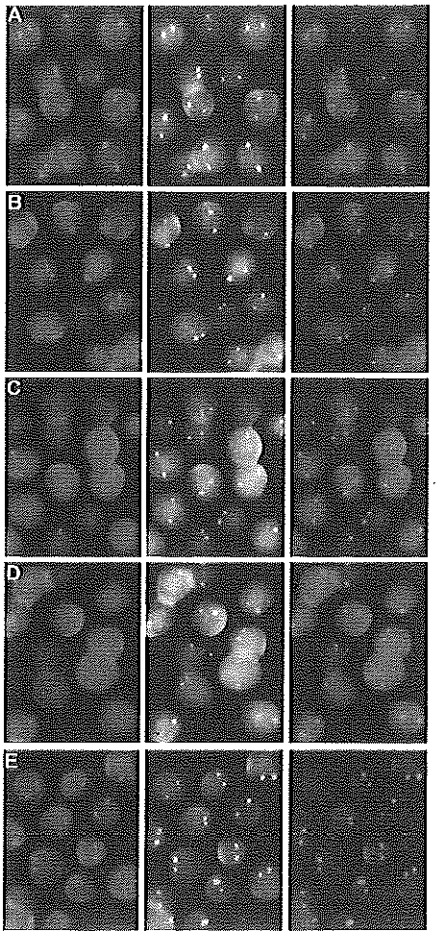


Figure 3. Primary transcript *in situ* hybridization on embryonic and fetal erythroid tissues. Double label *in situ* hybridizations were performed with the indicated probe sets on 10.5 day blood (A-D) and 13.5 day fetal liver cells (E) as described in material and methods. Three separate images are shown for each hybridization; left panels, red signal; right panels, green signal; middle panels overlay of red and green. A. ζ in red and α in green. B. βH1 in red and εy in green. C. βmaj in red and εy in green. D. βmin in red and εy in green. E. βmin in red and βmaj in green. The figures are composites of CCD images which represent a single focal plane and therefore may not show both loci from an individual cell in focus.

In situ transcription analysis of the ζ and α genes in embryonic cells demonstrates that both genes are transcriptionally active as expected but also reveals a high percentage of double signals in *cis* starting at 90% in 9.5 day red cells and decreasing to around 70% at 11.5 days (Fig. 4C). The percentage of α genes with a signal remains fairly constant around 90% throughout the embryonic period but the percentage of ζ gene signals decreases from 99% of all ζ genes at 9.5 days to 80% at 11.5 days (blood), resulting in fewer double signals and more single α gene signals. As in the β locus the pattern of transcriptional activity of the genes in the α locus does not remain static from 9.5 to 11.5 days.

εy, βH1 and ζ gene transcription is restricted to primitive erythroid cells

The same probes were used to analyze gene transcription in fetal liver erythroid cells at 11.5 and 12.5 days of gestation (Fig. 4A-C). Livers were isolated from fetal mice at the time points indicated and gently disrupted and prepared for *in situ* hybridization as described in materials and methods. *In situ* analysis of 11.5 and 12.5 day fetal liver cells shows that the βmaj and βmin genes are now the most highly expressed genes which is in agreement with the S1 analysis (Fig. 2A). At 11.5 days there is a small percentage of erythroid cells which continue to transcribe the εy, βH1 and ζ genes in the fetal liver preparations (Fig. 4A-C). It varies from one preparation to the next (average 15%) but by 12.5 days it is reduced to near zero. *In situ* analysis in 12.5 day peripheral blood in which greater than 75% of the cells are nucleated primitive erythrocytes shows that transcription of the globin genes is dramatically switched off (not shown). The near complete disappearance of εy, βH1 and ζ gene signals in the 12.5 day fetal liver preparations (Fig. 4A-C) suggested that those signals in the 11.5 day fetal liver preparations arose from contaminating embryonic erythrocytes. In addition the cells which display foci for the εy, βH1 and ζ genes in the 11.5 day fetal liver slides are microscopically distinct from the bulk of fetal liver cells since they display a high degree of autofluoresence, a characteristic of embryonic cells.

We confirmed that these cells were indeed embryonic and not fetal derived by three separate experiments. Hybridization of 11.5 day fetal liver slides with probes for εy, ζ and βmaj show that a small percentage of autofluorescent erythroid cells (~15%) transcribe both the εy and βH1 genes indicating that their expression is restricted to a sub-population of cells in the fetal liver preparations (Fig. 5A-C). Transcription signals for mouse ζ and εy in cells from 11.5 day fetal liver preparations from a homozygous transgenic line which contains a single copy of the human β globin, were completely separate from the cells which had transcription signals for the fetal restricted human β gene (Wijgerde et al., 1995) (not shown).

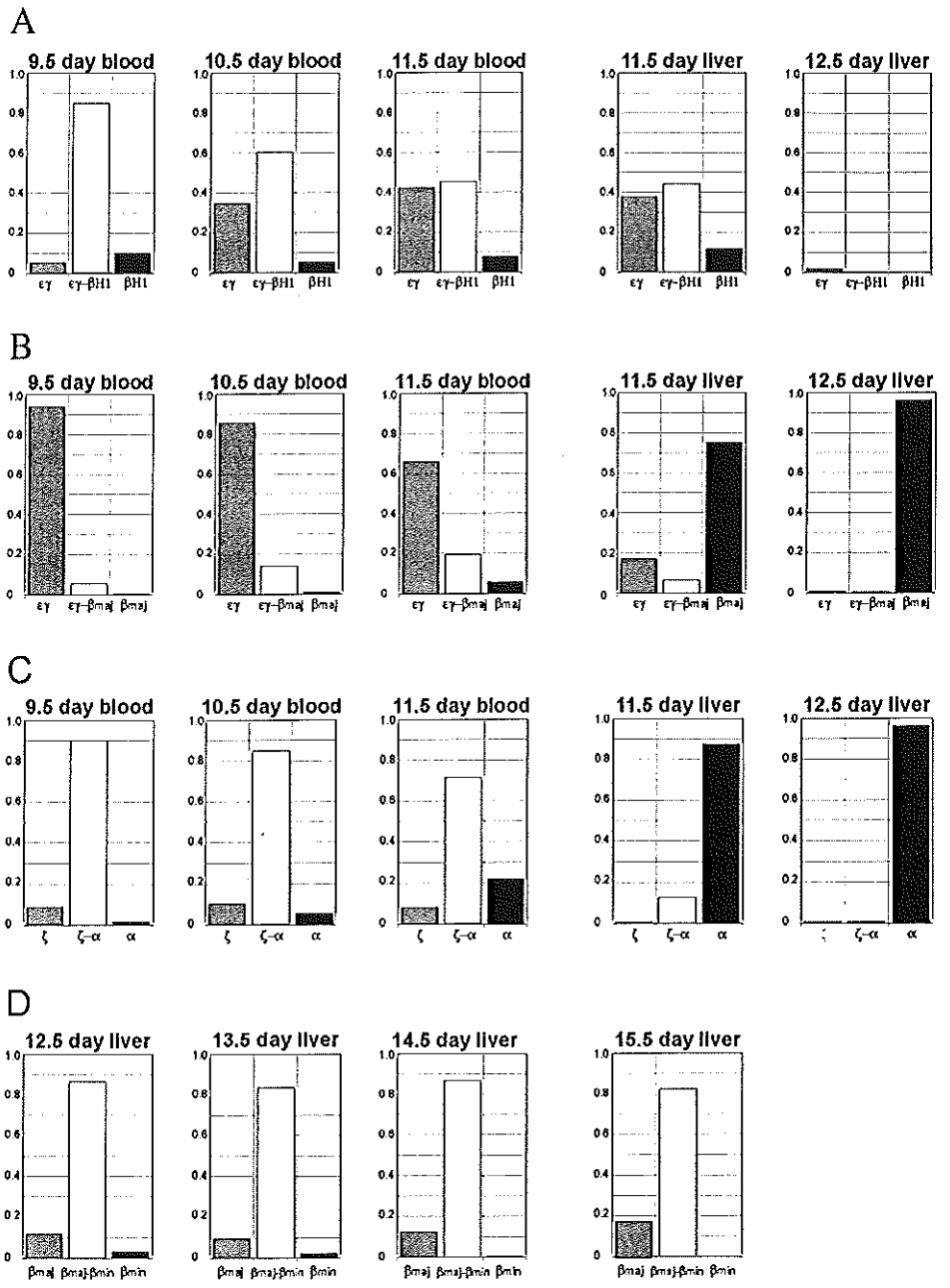


Figure 4. Quantitation of loci with single versus double *in situ* primary transcript signals during embryonic and fetal development. Double label *in situ* hybridizations as shown in figure 3 were quantitated and plotted as the percentage of loci with single versus double signals at the timepoints indicated. A. εy versus βH1. B. εy versus βmaj. C. ζ versus α. D. βmaj versus βmin. The values represent the percent of individual loci with detectable single or double signals.

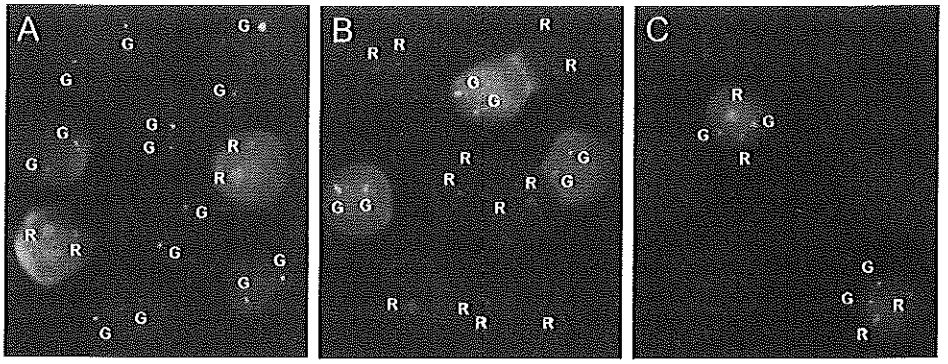


Figure 5. Primary transcript *in situ* analysis of 11.5 day fetal liver preparations. Cells were hybridized with; A. ζ in red and βmaj in green. B. εy in green and βmaj in red. C. ζ in red and εy in green. Note the relatively high autofluorescence in cells transcribing εy and/or ζ genes compared to the cells transcribing the βmaj gene.

Finally, *in situ* analysis of 11.5 day blood and fetal liver cells from an EKLF$^{-/-}$ mouse with εy and βmaj probes showed that the lack of βmaj transcription due to the EKLF knockout (Nuez et al., 1995; Perkins et al., 1995) did not lead to the appearance of εy transcription foci in the fetal liver cells. This indicates that the εy gene is silenced and not in competition with the βmaj gene in fetal cells (data not shown). Interestingly, βmaj transcription foci were not present in the EKLF$^{-/-}$ embryonic blood cells indicating that transcription of the β maj gene is dependent on EKLF activity in primitive cells. We conclude from these experiments that the cells with mouse ζ and/or εy and by inference those with βH1 signals, are embryonically derived erythroid cells and that transcription of these genes is restricted to the primitive lineage and does not occur in true definitive cells.Globin gene transcription in definitive erythroid cells

Presumably the mouse ζ, εy and βH1 genes are autonomously silenced during erythroid development as are the human ε (Raich et al., 1990) and ζ genes (Liebhaber et al., 1996). Definitive cells derived from the fetal liver express only the α1 and α2 genes from the α locus and the βmaj and βmin genes from the β locus. As mentioned previously the α1 and α2 genes are highly homologous even in their intervening sequences precluding the use of gene-specific intron probes to distinguish them.

The βmaj and βmin intron sequences are divergent and gene-specific probes were used to detect primary transcripts from these genes separately in double label *in situ* hybridizations (Fig. 3E). Quantitation of both βmaj and βmin signals together shows that signals are detected at 95% of the loci in 13.5 day fetal liver indicating that the probe efficiency is very high (Table 2). Separate quantitation shows that 92% of the βmaj alleles and 85% of the βmin alleles have primary transcript signals (Table 2). Approximately 80% of the loci have βmaj and βmin signals in *cis* (double signals), 10% have βmaj alone and 3% have βmin alone (Fig. 4D). A cell by cell analysis shows that in approximately 17% of the cells transcribe a different combination of genes while the *trans*-acting factor environment is the same (Table 2; indicated by *). The results from fetal liver cells from 12.5-15.5 days show that the percentage of βmin transcription foci is declining during development (Fig. 4D) which fits well with the previously reported changes in levels of beta-major and beta-minor proteins during fetal development (Whitney, 1977; Alter and Goff, 1980; Wawrzniak and Popp, 1987).

Co-activation or alternating single gene activation

The relatively high percentage of double signals in the mouse α and β loci suggest that the mechanism of multi-gene expression in these loci may be different from that proposed for the human β locus. In the human β locus the LCR is thought to flip-flop between genes to alternately activate transcription. Primary transcript *in situ* hybridization analysis of γ and β transcription in the early fetal liver cells showed that 85% of the loci have single gene signals (Wijgerde et al., 1995). The fact that nearly all cells had human γ and β mRNA in their cytoplasm indicated that alternation must occur and suggested that the small amount of double signals were due to a recent switch from γ to β or vice versa. The results from the mouse α and β loci could be interpreted to indicate that transcription is co-initiated from multiple globin genes in *cis* in most cells. However, there are two indications from the data that contest this conclusion in the case of the β locus. Firstly, a significant proportion of the cells has loci, which are responding differently to the same *trans*-acting factor environment (38% at 10.5 days and 18% at 13.5 days, Table 2). Secondly, there is a significant proportion of loci with only single gene signals (38% at 10.5 days and 14% at 13.5 days, Fig. 4D and Table 2). These results suggest the possibility that the individual genes in the mouse β locus may be alternating, however if true alternation must occur

more often in the mouse locus than in the human β locus.

We therefore designed an experiment to provide further evidence of co-initiation or alternating transcription of multiple genes in the mouse globin loci making use of the inhibitor of transcriptional elongation DRB (5,6-dichloro-1-β-D-ribofuranosyl-benzimidazole). Previous studies have shown that DRB does not affect initiation of transcription (Fraser et al., 1978,; Marshall and Price, 1992) but prematurely aborts elongating transcripts 400-600 bp from the initiation site (Chodosh et al., 1989; Marshall et al., 1996) by inhibiting the activity of the P-TEFb kinase which phosphorylates the carboxy terminal domain (CTD) of Pol II (Marshall et al., 1996). The effect of DRB treatment on globin gene transcription in mouse erythroleukemia cells has been reported previously (Tweeten and Molloy 1981). The results demonstrate that DRB causes premature termination without affecting initiation of transcription. This was tested in transgenic DRB treated mouse fetal liver cells using *in situ* hybridization with promoter-proximal and -distal probes to the human γ- and β-globin primary transcripts (Gribnau et al., submitted). *In situ* signals with probes that hybridize to intron 1, located in the first 300 bases of the primary transcript are still visible after 15 minutes of DRB treatment, whereas probes that hybridize to intron 2, 600-1200 bases 3' of the initiation site are no longer visible after 7 minutes of DRB treatment.

What makes DRB inhibition of elongation useful is the fact that it is reversible allowing elongation to proceed normally after removal of DRB. We reasoned that if co-initiation was occurring in the roughly 85% of loci that have double signals then restoration of elongation after DRB treatment should result in the near simultaneous reappearance of single and double gene signals using probes which hybridize to a distal region of the primary transcript (intron 2). If alternation of transcription is the cause of the double signals then single signals should appear first after washing out the DRB and be followed by the reappearance of double signals after a lag that reflects the extra time required to produce a second signal in *cis*.

We isolated 10.5 day blood cells and pulse treated them with DRB as described in materials and methods. We assayed the reappearance of intron 2 primary transcript signals by taking aliquots at the indicated times after washing out the DRB and quantitated single and double signals. The cells in the untreated panel (Fig. 6A) show the normal distribution of single versus double signals for the ζ and α genes. After fifteen minutes of DRB treatment no primary transcript signals are detectable (Fig. 6B). Five minutes after removal of the DRB primary transcript signals are again detectable at approximately 50% of the erythroid loci with single and double signals reappearing at the same rate (Fig. 6C,E). After 20 minutes greater than 90% of the loci in the DRB treated cells show detectable signals compared to untreated control cells indicating that the cells recover quickly and are not adversely affected by the elongation block (Fig. 6D,E). It is important to point out that the reappearing signals are therefore detected in a subset of recovered cells and are not spread randomly throughout the entire population of cells or loci. The fact that the amount of double (ζ/α) versus single (ζ or α) signals is the same as in the untreated cells and that this is maintained regardless of the number of cells that have recovered (Fig. 6K), suggests that a co-initiation mechanism may explain multi-gene expression in the mouse α locus. However a transcriptional mechanism which involves very frequent alternations (i.e. alternations occurring every minute for example) would give the same result. We have determined the lifetime of globin primary transcript signals in the presence of DRB or actinomycin D. The primary transcript signals fade to undetectable levels 7 minutes after addition of the inhibitor (Gribnau et al., submitted) which means that if alternations occur frequently, i.e. on a time scale of less than 7 minutes, then nearly all signals would appear as double signals.

The same DRB treated cell aliquots were hybridized with intron 2 probes for mouse εy and βH1 primary transcripts. Here again cells recovered quickly from the DRB treatment (40% after 2.5 min and 90% after 10 min) as measured by the number of cells with two actively transcribing β-globin chromosomes per cell. The results for the β genes were strikingly different from those of the α genes (Fig. 6K). Single εy and βH1 signals reappear and reach near maximal levels within five minutes after washing out DRB (Fig. 6H,J). In contrast, reappearance of the double signals lags significantly behind and reaches its maximum only after 7.5 minutes (Fig. 6J,K). In addition, the single signals initially increase and then decrease suggesting that they are being converted to double signals. After 10 to 20 minutes the relative amounts of double versus single signals are very close to control values (Fig. 6K). These results are consistent with an alternating mechanism of multi-gene expression in the mouse β locus. This type of mechanism suggests that the double signals on the same allele arise in loci, which have switched from one β gene to the other in the last 7 minutes (as indicated by the 7 minute lifetime of the primary transcript signal in the presence of inhibitors of transcription). Thus an individual gene would be on for little more than 7 minutes at a time on average, which is similar to that observed in the human β-globin locus (Dillon et al., 1997).

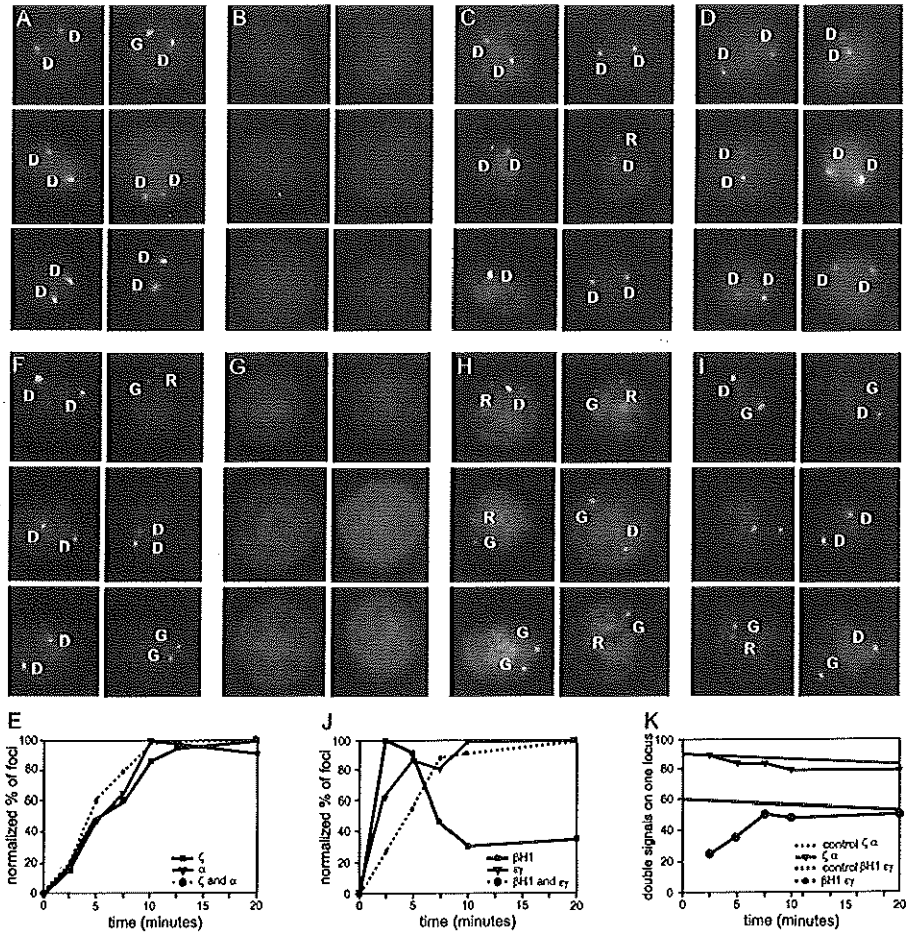


Figure 6. Kinetic analysis of reappearance of primary transcript signals after DRB inhibition of elongation. 10.5 day blood cells were pulse treated with DRB as described in materials and methods and hybridized with mouse ζ and α globin primary transcript probes (A-D) and mouse εγ and βH1 primary transcript probes (F-I). Quantitation of single and double signals are presented from the indicated timepoints in panels E (ζ and α) and J (εγ and βH1). Curves were normalized to the maximum values for single and double signals. A, F. Prior to DRB treatment. B, G. Cells treated with DRB for 15 minutes. C, H. 5 minutes after washing out DRB. D, I. 20 minutes after washing out DRB. K. These curves show the percentage of double signals on one allele with or without DRB treatment. Dotted green line: percentage of ζ/α double signals in untreated cells as a function of time. Solid green line: as above but for DRB treated cells after release of DRB. Dotted red line: percentage of εγ/βH1 double signals in untreated cells. Solid red line: as above but for DRB treated cells after release of DRB treatment. The data represent the results of a single experiment, which was repeated three times yielding essentially the same curve shapes with the same relative positions to each other, but differed slightly in their position along the x-axis, reflecting the inherent difficulties in obtaining short time points.

Discussion

We have used a combination of S1 nuclease protection assays on total RNA and primary transcript *in situ* hybridization in erythroid cells to analyze the gene expression and transcription patterns of the murine α and β globin genes. The results show that mRNAs from the embryonic εy, βH1 and ζ globin genes are expressed at high levels in primitive erythroid cells and are still detectable at low levels in the early fetal liver until 12.5-14.5 days of gestation. Primary transcript *in situ* hybridization indicates that the small amount of residual expression of these genes in early fetal liver preparations is due to the presence of a sub-population of cells that transcribe both the εy and ζ genes. These cells display a high degree of autofluoresence and do not transcribe the fetal restricted human β gene indicating that they are circulating embryonic cells and not true definitive cells derived from the fetal liver. We conclude from this data that transcription of the εy, βH1 and ζ globin genes is restricted to the primitive erythroid lineage.

In contrast, the α1, α2 and to a lesser extent the βmaj and βmin genes are expressed in the primitive lineage and are the only globin genes expressed in definitive cells. α1 and α2 mRNA makes up approximately 50% of the embryonic α globin at 10.5 days whereas βmaj and βmin expression is approximately 6%. In the case of the β locus where all four genes are distinguishable, the results show that they are transcribed at frequencies, which are inversely proportional to their distances from the LCR. These results are reminiscent of those of Dillon et al., (1997) in which a marked β gene was placed at different positions in the human globin locus and the effect of distance on transcription of the introduced gene and the other globin genes was measured. The results showed that altering the distance of a gene from the LCR had an effect on the frequency of transcription of that gene. An LCR proximal gene was transcribed more often than a distal gene with reciprocal consequences for the remaining downstream genes in the locus. Those results suggested that the LCR activates gene transcription by forming direct chromatin interactions with gene-local regulatory elements and that the frequency of contact between two or more competing genes and the LCR is controlled in part by differences in distance.

The results highlight an important difference in the pattern of developmental regulation between the mouse and human β loci. In the mouse locus βmaj and βmin gene transcription is activated in the primitive lineage, whereas human β gene transcription is restricted to definitive erythroid cells in transgenics. The human β gene is thought to be silenced in primitive cells in part through competition for the LCR by the more LCR-proximal ε and γ genes. A competition mechanism could also be operating in the mouse β globin locus, which is suggested by the DRB experiments, and could account for the observed inverse correlation between percentage of transcriptionally active genes and distance from the LCR in primitive cells. One might expect from these results that if the human β gene in its distal location in the locus is activatable in primitive erythroid cells then some transcription should be detected *in situ*. Clearly the human β gene is activatable in primitive cells when placed next to the LCR (Enver et al., 1990; Hanscombe et al., 1991) or in the position of the ε gene (Dillon et al., 1997) but not in its wild type location (Wijgerde et al., 1995) or when placed just 5' of the δ globin gene (Dillon et al., 1997). These results suggest the possibility that the distal part of the human locus which contains the δ and β genes is not accessible to LCR activation in primitive cells. This concept is supported by results from multi-copy transgenic mice which contain head-to-tail tandemly integrated copies of a complete 70 kb human globin locus in which the β gene is juxtaposed on the 3' side by an LCR (Milot et al., 1996). These mice do not express β mRNA in the primitive erythroid cells even though the β gene would be closer to the downstream LCR than the γ genes. It is important to note here that the LCR can activate the β gene in a 5 kb fragment when it is placed 5' of HS5 (Zafarana et al., 1995) i.e. it works in both orientations. Taken together with the results presented in this paper which indicate that the βmaj and βmin genes are partially suppressed but not silenced by gene competition in primitive cells, it suggests that the human β gene is silenced in primitive cells through another mechanism which may involve epigenetic chromatin modification (Fig. 7). It has been proposed long ago that the human β locus may be divided into distinct chromatin sub-domains (see Collins and Weissman and references therein). This idea gains support from experiments with somatic cell hybrids between primitive transgenic erythroid cells and MEL cells which show that γ gene expression is retained through several hybrid cell divisions before eventually switching to β expression (Papayannopoulou et al., 1986; Stanworth et al., 1995; see also Fraser et al., 1998). The above discussion is not intended to suggest that competition plays no role in transcriptional regulation of the β globin genes. Competition is clearly operating in the human locus when multiple genes (or domains) are activatable, such as in primitive cells between ε and γ and in early definitive cells between γ and β.

The DRB experiments presented here suggest that alternating transcriptional activation of the mouse β genes is occurring which may be indicative of gene

competition for the LCR. Thus the results presented are most easily explained by the direct chromatin interaction mechanism between the LCR and an individual β gene that has been proposed for the human β locus. We propose a model in which the entire mouse β locus is accessible to LCR activation in primitive cells and that the individual genes are transcribed at frequencies which are inversely proportional to their distance from the LCR due largely to gene competition (Fig. 7). In definitive cells the εy and βH1 genes are no longer activatable due to silencing which may involve chromatin modification allowing the adult genes to be transcribed more frequently. Thus we propose that switching in the mouse locus involves shutting off the embryonic genes. In contrast, the switch to definitive erythropoiesis in the human transgene β locus involves "opening up" of the adult δ-β domain through chromatin modification, making it accessible to activation by the LCR in competition with the γ genes.

The DRB results with the mouse α locus are more difficult to interpret but are clearly different from those of the β locus. The results of targeted disruption of the ζ and α1 genes by insertion of a PGK-Neo cassette (Leder et al., 1997) are strongly indicative of gene competition for the αMRE. Insertion into the ζ gene had a more profound effect on α1 and α2 gene expression than did insertion into the α1 gene, clearly indicating that relative position of the PGK-Neo cassette with respect to the αMRE is important. It has been firmly established that insertion of an activatable promoter (or gene) between a regulatory element and its normal target promoter leads to decreased activation of the downstream gene (Kim et al., 1992; Fiering et al., 1995; Hug et al., 1996; Pham et al., 1996; Dillon et al., 1997) indicating that gene order and/or relative distance from the regulatory element are key parameters in determining the competitive ability of a gene. We cannot conclude that the mouse α genes are alternately transcribed from our results, however the accumulated data suggest that they may be frequently alternating in a competitive mechanism for interaction with the αMRE. Obviously experiments which are able to show a direct interaction between the LCR or αMRE and an individual gene will allow firm conclusions regarding this type of mechanism.

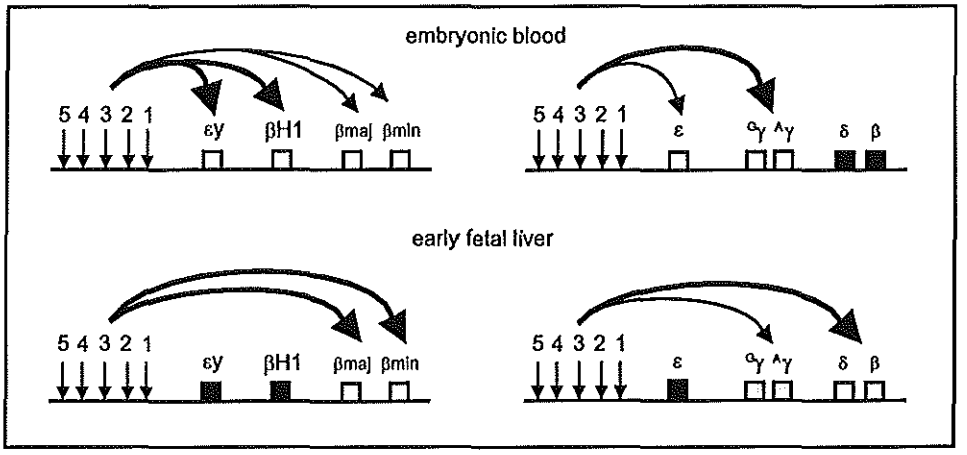


Figure 7. Models of developmental switching in the mouse and human β globin loci. Schematic diagrams of the mouse (left) and human (right) β globin loci are shown during embryonic (top) and early fetal (bottom) erythropoiesis. Vertical arrows denote DNase I hypersensitive sites of the LCRs. The curved arrows signify an interaction between the LCR and an individual globin gene. The relative thickness, of the arrows represents the percent of transcriptionally active genes in the population of erythroid cells. Open boxes represent activatable genes. Closed boxes represent genes which are silenced or in an inaccessible chromatin conformation (or both). Gray boxes (γ genes, fetal liver) reflect the possibility that the γ genes in some early fetal liver cells may be silenced (i.e. cells committed to β transcription; Wijgerde et al., 1996).

Materials and Methods

Preparation of RNA and S1 Nuclease protection assay.

RNA was prepared from 10.5 embryos, 12.5 and 14.5 day fetal liver and adult blood and subjected to S1 nuclease assay as previously described (Weaver and Weissman, 1979). The εy, βH1 and adult β probes used were those described by Lindenbaum and Grosveld (1990). The mouse α and ζ probes were a 300 bp BamHI fragment from plasmid GSE 1454 and a 300 bp ApaI/AvaII fragment subcloned from cosmid cML1 (Kielman et al., 1996) respectively. S1 nuclease protected bands were quantitated using phosphorimage analysis (Molecular Dynamics).

Primary transcript in situ hybridization

Primary transcript *in situ* hybridization to detect transcriptional activity of the mouse β globin genes in 9.5, 10.5, 11.5 day blood, 11.5, 12.5 and 13.5 day fetal liver cells was performed as described by Wijgerde et al., (1995). Embryonic blood cells were collected in PBS, fetal liver cells were disrupted in PBS (by pipetting up and down several times) and spotted on poly-L-lysine coated slides. Slides were placed in fixative (4% formaldehyde, 5% acetic acid in saline) for 20 minutes at room temperature then washed three times in PBS at room temperature and stored in 70% ethanol at -20^0. Slides were pre-treated for hybridization by rinsing in tris/saline (0,1M tris, pH7 and 0.85% saline) and incubated in 0.01% pepsin, 10mM HCl, 37°C. for 5 min. The slides were briefly rinsed in H_2O and fixed again for 5 min, in 3.7% formaldehyde in PBS at room temp. After a PBS wash slides were dehydrated in ethanol (70%, 90% and 100% respectively 3 min. each). Slides were hybridized in 25% formamide, 2xSSC, 1mM EDTA, 5x Denhardt, 50 mM $NaH_2PO_4/NaHPO_4$ (pH7.0) and 200 ng sheared salmon sperm DNA using 1-5 ng/µl of the appropriate oligonucleotide probes at 37°C. overnight. After hybridization slides were washed in PBS for 30 min at 37°C. and briefly rinsed in TST (tris/saline with 0.05% tween 20). The slides were then incubated in tris/saline containing 1% blocking reagent (TSB) (Boehringer) for 30 minutes at room temp. The slides were subsequently incubated with TSB containing the appropriate primary antibody (or Avidin D texas red for biotinylated oligos) for 30 minutes at room temp and washed 2 times for 5 min. in TST. The previous steps were repeated with fluorescently labeled secondary (biotinylated goat anti-avidin D for biotinylated oligos) and tertiary antibodies. After the final washing step the slides were dehydrated (70%, 90% and 100% ethanol), air dried and mounted with a 1:1 mixture of 1% DABCO, 0.4µM DAPI (sigma), 90% glycerol, 0.02% sodium azide 10mM tris (pH 8.0) with Vectashield (Vector labs).

Quantitation of primary transcript signals was done by counting at least 400 cells for each timepoint using an epiflourescence microscope (Leitz). The figures presented were created with a CCD camera (Hamamatzu).

Oligonucleotide probes

Antisense oligonucleotide probes which recognize the intronic sequences of the respective globin gene primary transcripts were labeled with digoxygenin, dinitrophenol or biotin haptens at both ends and in the middle as indicated.

εy 1-4 (digoxygenin)
CTCAGAATTCTTGATTTCCCTAGCTCTTTGTACA
CTAAAAAACAATTCTT
CAGCCATTCACTGTCACCCTTACTGGGACCAAT
TAATTAACTTTGACAGC
ACTTCTCTTCATATTACTCTCCATATAAATCCAT
GATAAATTTTATCACG
GTTGTCTTGCAAGACTTTTCTTCAACATCAATA
AATAGGACCGCGCAAAA

βH1 1-4 (biotin)
CAAAACCCTATAGAAACCCTGGAAATTTCTGCC
ATGCATAAGGATAATTT
TGGACCCATGGACTCTAACATCTGTCAAGGCAT
TGCCAATCACAGTCTCA
AAATGCTGGGCGCTCACTCAAATCTGCACCCAA
ATCATTGTTGCCCACAA
GCATAGATGTATTAATTTATAAAAACATACTCC
TTTTTAAAAAAGATCCA

βmaj 2-4 (dinitrophenol)
GAACTCTTGTCAACACTCCACACACAGTCATGG
AGACTGCTCCCTAGAAT
ATGGGAAGTAAATAACCAGAGCTTAATTAATTT
AGTAAAATGCAACTGGA
GACAAATTATTATAAGAATCCTATGTCAAACAG
AATTTATATGTAAAATA

βmin 1-4 (biotin)
TATGAAGTAGAGCAACAATACAAGATGCTGAA
GGCCGATTTCAAATGGAA
ACTGTGGAAAGGATCAGAGAATCATTTATCTTT
TTGTCCTCAGAGTAAGA
AAAAAAGAATTATTCTATGACACACAAAATTTA
GCCACAAAATATACTCT
GGTAAAATGGCAGCTGGGTTCTACTGGTCAATT
TTGATAAGAATTATTCT

ζ 1-5 (biotin)
CCTTCTCAGTGGCTTCTCCTCACTAACTGCTCTT
TGTCACTTCTGTCTC
ATGGAAGACTCTGGTGAGCTCTGTGAATGCCAG
CCCACCTCCTTTAGTA

ACAACCCCAAGAGTGATGTTACTATTTGCTGTT
GCACAAGGGTCTACA
AAGGGGATTTGATGCCTCCAGCCTCCAATGGCA
CCCATGCCTGCGCTCG
α 1-3 (digoxygenin)
TCACAGAAAAGCATAGTTAGAAGCTGCCCACT
GAGCGAGTGCCAGGTCCAT
TAGCCCTTCCTAGGGGTCCCAGATGCCGCCTGC
CAGGTCCCT
GCTCCCCTTCCTGGGACCACTATGTTCCCTGCC
TTGGGCACGAGGACCCT

Reversible inhibition of elongation with DRB

Blood was collected from 10.5 day embryos in PBS and diluted to approx. 10^6 cells/ml. Half of the cells were treated with 100 μM DRB (5,6 dichloro-1-B-D-ribofuranosylbenzimidazole, Sigma) for 15 minutes at 37^0 C., the other half were used as untreated controls. The cells were then diluted with several volumes of ice cold PBS, spun at 200 g and washed two more times with ice cold PBS and resuspended in PBS at approx. 10^6 cells/ml at 25^0C. Aliquots were removed at 0, 2.5, 5, 7.5, 10, and 20 minutes after washing, spotted onto poly-L-lysine slides (Sigma) and prepared for *in situ* hybridization as described.

Acknowledgments We thank Drs. S Philipsen and A. van den Wijngaard for critical reading of this manuscript. We are grateful to Drs. T McMorrow. E. Milot, M. Wijgerde, M. Bender, M. Groudine, and D.Higgs for helpful discussions and L. Braam for animal care. We are indebted to Dr. M. Kielman for the generous gift of the mouse α globin cosmid. T.T. is supported by the Nederlandse Organisatie voor Wetenschappelijk Onderzoek (N.W.O.). J.G. is supported by the Jeantet Foundation. This work is supported by the N.W.O.

Concluding remarks

Our understanding of chromatin structure in relation to the regulation of gene expression has become exceedingly enriched during the last few years. When this research was initiated much was known about the regulation of the genes in the human β-globin cluster; the LCR was identified as the dominant regulator of the locus and transgenic mice containing a single copy of the human locus had been generated (Strouboulis et al., 1992; Peterson et al., 1993: Gaensler et al., 1993). The use of small constructs in transgenic mice revealed that the individual genes are competing for activation by the LCR and that a gene closer to the LCR had an advantage over a gene that was located further away from the LCR (Hanscombe et al., 1991). From these experiments it was hypothesised that the LCR might loop out the intergenic DNA to interact directly with for instance the regulatory region of a gene, to activate it. In such a model (Fig. 9) the expression levels of a gene are controlled by two parameters; the frequency of LCR-gene interaction and its stability.

Within the last four years we have been able to study these parameters in greater detail; the effect of frequency of interaction was tested in transgenic mice carrying a single copy of the complete locus that contained duplicated β-genes. Two different lines were generated in which the position of the second β-gene was different (chapter 4). Since these duplicated β-genes have the same regulatory regions there is no effect of differences in stability, enabling us to study the effect of the relative distance of a gene to the LCR in which the distance of a gene to the LCR determines the frequency of that gene to becomes activated.
We have shown that a gene more proximal to the LCR is activated with a competitive advantage over a more distally located gene and that the level of expression is proportional to the frequency of transcription periods

The stability of the LCR/gene interaction was studied in mice containing a copy of the human β-globin locus together with a disruption in the EKLF gene, an erythroid transcription (Wijgerde et al., 1996; Perkins et al., 1996). A reduction in the levels of EKLF resulted in decreased levels of β-gene expression with a concomitant increase in γ-globin expression levels. In a looping model this could indicate that EKLF is involved in the stabilisation of the LCR-gene interaction.
These data together with the data from experiments described in chapter 5 and 6, show that the activation of genes in the globin locus is conserved during evolution and that the LCR activates only one gene at a time. They all support a looping model as a mechanism for gene

activation in the human β-globin locus. The best evidence for such a model has to await the actual visualisation of the LCR/gene interaction – an idea that is technically challenging at present.

Histone acetylation has not as yet been studied in the human or mouse β-globin loci. It will be interesting to investigate the effect of histone acetylation in these loci. From experiments on the chicken β-globin locus, it became clear that the general DNase I sensitivity in this locus correlates with histone acetylation (Hebbes et al., 1992, 1994). It will be important to clarify if histone acetylation in the human locus similarly accounts for the general DNase I sensitivity observed in erythrocytes and what role the LCR plays in establishing this[11], by analysing the mice with deletions in the individual hypersensitive sites (chapter 3). The identification of proteins involved in promoting histone acetylation in the β-globin locus could guide these experiments. An attractive candidate for such a protein would be the erythroid specific transcription factor EKLF.

In chapter 3 we describe the analysis of transgenic mice containing different deletions of individual hypersensitive sites of the human β-globin LCR. The result of such alteration to the LCR on the expression of the different globin genes is not dramatic. However, when such a transgene lands in or near heterochromatin[12] it no longer expresses the globin genes in a position-independent manner, the transgenes are subject to position effects (PE) and it appears that the LCR upon mutation loses its ability to overcome the repressive effects of heterochromatin. These transgenic mice allow us to study the organisation of heterochromatin and the organisation of the nucleus in mammals *in vivo*. Similar studies with the endogenous mouse β-globin locus may complement the data from transgenic mice with the human β-globin locus. These findings may help to elucidate the complex mechanism of heterochromatin formation with the added advantage that in this case the locus is being studied in its appropriate genomic context/location.
Using embryonic stem cell technology, deletions were made of individual hypersensitive sites of the murine LCR (Fiering et al., 1996; Hug et al., 1997; Trimborn et al., chapter 6). The effect of the different deletions on the level of globin transcription were minor (10-30% reduction in gene expression) and comparable to the deletions in the human locus. However, it will be important to show what causes the reduction in transcription levels.

[11] Note that in the absence of the LCR the locus is not sensitive to DNase I (Kioussis et al., 1983).

[12] The centromere is the only defined region in the genome, which is known for sure consists largely of heterochromatin.

Studies in which these LCR deletion mice [of the human locus as well as of the murine β globin locus] are crossed with mice that have increased or decreased levels of mammalian homologues of the Pc-G proteins will add to our understanding of the mechanism of heterochromatin formation. Since we can monitor transcription in individual cells for the different mutant LCR transgenes that are integrated in or near heterochromatin, the effect of the different levels Pc-G or Trx-G proteins can be studied *in vivo*.
Recently it was demonstrated that the Ikaros proteins, which were originally described as being transcriptional regulators, localise to discrete heterochromatin–containing foci in interphase nuclei (Brown et al., 1997). Ikaros is a DNA binding, zinc-finger protein with a Krüppel-like domain with target genes mainly in the lymphoid associated genes. Ikaros null mutants lack B cells and show aberrations in T lymphocyte development and a severe hematopoietic phenotype (Wang et al., 1996). The fact that Ikaros is localised with centromeric heterochromatin and furthermore that Ikaros is mainly co-localised with inactive genes suggested that Ikaros might recruit transcriptional active genes/domains to a repressive environment. It could be that an Ikaros (like) protein is involved in the silencing of the transgenes described in chapter 3 that suffer from position effects.

The β-globin LCR is a collection of five developmentally stable hypersensitive sites that have been conserved during evolution, as was shown by the fact that the mouse β-globin LCR is very similar in sequence, structure and also in its function to the human β-globin LCR (chapter 6). One cannot dissect the LCR without adverse effects, relatively small deletions result in the loss of the LCR's prominent feature: expression of linked genes, independent of the site of integration in the host genome. However, if the LCR is intact it seems to be capable of competing successfully with the repressive environment of heterochromatin. This hints at a highly selective pressure for conservation of the LCR elements during evolution.

References

Aladjem, M., M. Groudine, L.Brody, E. Dieken, R. Fournier, G. Wahl, E. Epner. 1995. Participation of the human β-globin LCR in initiation of DNA replication. Science 270: 815-818.
Alfrey V., R. Faulkner and a. Mirsky. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. PNAS 51: 786-794.
Alter, B.P. and S.C. Goff. 1980. A murine model for the switch from fetal to adult hemoglobin during ontogeny. Blood 56: 1100-1105.
Amrolia. P., L. Ramamuhurthy, D. Sauja, N. Tanese, S. Jane, and J. Cunningham. 1997. The activation of the enhancer binding protein Nf-E2 interacts with TAFII 130 and mediates long-range activation of the α- and β-globin gene loci in an erythroid cell line.
Andrews, N.C., H. Erdjument-Bromage, M.B. Davidson, P. Tempst and S.H. Orkin, 1993. Erythroid transcription factor NF-E2 is a hematopoietic-specific basic-leucine zipper protein. Nature 362: 722-728.
Antoniou, M., E. deBoer, G. Habets and F. Grosveld. 1988. The human beta-globin gene contains multiple regulatory regions: identification of one promoter and two downstream enhancers. EMBO J. 7: 377-384.
Antoniou, M. and Grosveld, F. 1990. The beta-globin gene dominant control region interacts differently with distal and proximal promoter elements. G&D 4:1007-1013.
Antoniou, M., deBoer, E., E. Spanopoulou, A. Imam and F. Grosveld. 1995. TBP binding and the rate of transcription initiation from the human β-globin gene. NAR 23: 3473-3480.
Armstrong, J.A. and B.M. Emerson. 1996. NF-E2 disrupts chromatin structure at human beta-globin locus control region hypersensitive site 2 in vitro. MCB 16: 5634-5644.
Ashe, H., J. Monks, M. Wijgerde, P. Fraser, N. Proudfoot. 1997. Intergenic transcription and trans-induction of the human β-globin locus. G&D 11: 2494-2509.
Bannister, A.J. and T. Kouzarides. 1996. The CBP co-activator is a histone acetyltransferase. Nature 384: 641-
Behringer, R.R., R.E. Hammer, R.L. Brinster, R.D. Palmiter and T.M. Townes. 1987. Two 3' sequences direct adult erythroid-specific expression of human beta-globin genes in transgenic mice. PNAS 94: 7056-7060.
Behringer, R.R., Ryan, T.M., Palmiter, R.D., Brinster, R.L. and Townes, T.M. (1990) Human gamma- to beta-globin gene switching in transgenic mice. G&D 4: 380-389.
Benyajati, C., and A. Worcel. 1976. Isolation, characterisation and structure of the folded interphase genome of Drosophila. Cell 9: 393-407.
Bernet, A., S. Sabatier, D. Picketts, R. Ouazana, F. Morle, D. Higgs, and J. Godet. 1995. Targeted inactivation of the major positive regulatory element (HS-40) of the human • -globin gene locus. Blood 86: 1202-1211.
Bickel, S. and V. Pirotta. 1990. Self association of the Drosophila zeste protein is responsible for transvection effects. EMBO J. 9: 2959-2967.
Bickmore, W., and K. Oghene. 1996. Visualising the spatial relationship between defined DNA sequences and the axial egion of extracted metaphase chromosomes. Cell 84: 95-104.
Blom van Assendelft, G., O. Hamscombe, F. Grosveld and D. Greaves. !989. The β-globin dominant controlregion activates homologous and heterologous promoters in a tissue-specific manner. Cell 56: 969-977.
Bodine, D.M. and T.J. Ley. 1987. An enhancer element lies 3' to the human $^{A}\gamma$ globin gene. EMBO J 6: 2997-3
Bonifer, C., M. Vidal, F. Grosveld and A.E Sippel. 1990. Tissue specific and position independent expression of the complete gene domain for chicken lysozyme in transgenic mice. EMBO J. 9: 2843-2848.
Boyes, J. and G. Felsenfeld. 1996. Tissue-specific factors additively increase the probability of the all-or-none formation of a hypersensitive site. EMBO J. 15: 2496-2507.
Bresnick, E.H. and G. Felsenfeld. 1994. Dual promoter activation by the human β-globin locus control region. PNAS 91: 1314-1317.
Briegel, K., K.C. Lim, C. Plank, H. Beug, J.D. Engel and M. Zenke. 1993. Ectopic expression of a conditional GATA-2/estrogen receptor chimera arrests erythroid differentiation in a hormone-dependent manner. G&D 7: 1097-1109.
Brotherton, T., D.H. Chui, J. Gauldie, and M. Patterson. 1979. Hemoglobin ontogeny during normal mouse fetal development. PNAS 76: 2853-2857.
Bungert, J., U. Dave, K. Lim, K. Lieuw, J. Shavit, Q Liu, and D. Engel. 1995. Synergistic regulation of human β-globin gene switching by LCR elements HS3 and HS4. G&D 9: 3083-3096.
Cairns, B.R., Y. Lorch, Y. Li, M. Zhang, L. Lacomis, H. Erdjument-Bromage, P. Tempst, J. Du, B. Laurent and R.D. Kornberg. 1996. RSC, an essential, abundant chromatin-remodelling complex. Cell 87: 1249-1260.
Candau R. and S Berger. 1996. Structural and functional analysis of yeast putative adapters. J.B.C. 271: 5237-

Cao, S., P.D. Gutman, , H.P.G. Dave and A.J. Schechter. 1989. Identification of a transcriptional silencer in the 5'-flanking region of the human ε-globin gene. PNAS 86: 5306-5309.
Carson , and Wiles. 1993. Far upstream regions of Class II MHC Ea are necessary for position independent, copynumber dependent expression of Ea transgenesis. NAR 21: 2065-2073.
Caterina, J.J., D.J. Ciavatta, , D. Donze, R.R. Behringer and T.M. Townes. 1994. Multiple elements in human β-globin locus control region 5' HS2 are involved in enhancer activity and position-independent, transgene expression. NAR 22: 1006-1011.
Cavalli G. and R. Paro. 1998. The Drosophilla Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. Cell 93: 505-518.
Chada, K., J. Magram, K. Raphael, G. Radice, E. Lacy and F. Constantini. 1985. Specific expression of a foreign beta-globin gene in erythroid cells of transgenic mice. Nature 314: 377-380.
Chada, K., J. Magram and F. Constantini,. 1986. An embryonic pattern of expression of a human foetal globin gene in transgenic mice. Nature 319: 685-689.
Chao, M.V., P. Mellon, P. Charnay, T. Maniatis and R. Axel. 1983. The regulated expression of ß-globin genes introduced into mouse erythroleukemia cells. Cell 32: 483-493.
Chan, C.S., L. Rastelli and V. Pirotta. 1994. A polycomb response element in the Ubx gene that determines an epigenetically inherited state of repression. EMBO J. 13: 2553-2564.
Chen W.Y., E. Bailey, S. McCune, J-Y. Dong and T. Townes. 1997. Reactivation of silenced, virally transduced genes by inhibitors of histone deacetylases. PNAS 94: 5798-5803.
Chodosh, L.A., A. Fire, M. Samuels, and P.A. Sharp. 1989. DRB inhibits transcription elongation by RNA polymerase II in vitro. J. Biol. Chem. 264, 2250-2257.
Choi, O., and J.D. Engel. 1988. Developmental regulation of beta-globin gene switching. Cell 55: 17-26.
Chui, D.H., T. Brotherton and B. Gauldie. 1979. Hemoglobin ontogeny in fetal mice. Adult hemoglobin in yolk sac derived erythrocytes. In Cellular and molecular regulation of hemoglobin switching. (G Stamatoyannopoulos and A. Nienhuis, ed.), pp. 213-225 Alan R. Liss, Inc., New York.
Chung, J., M. Whiteley, and G. Felsenfeld. 1993. A 5' element of the chicken β-globin domain serves as an insulator in human erythroid cells and protects against position effects in Drosophilla. Cell 74: 505-514.
Clarke, D.J., L.P. O'Neill and B.M. Turner. 1993. Selective use of H4 acetylation sites in the yeast Saccharomyces cerevisiae. Biochem. J. 294: 557-561.
Cockerill, P.N. and W.T. Garrard. 1986. Chromosomal loop anchorage of the kappa immunoglobulin gene occurs net to the enhancer in a region containing topoisomerase II sites. Cell 44: 273-282.
Collins, F.S. and S.M Weissman. 1984. The molecular genetics of human haemoglobin. Prog. Acid Res. Mol. Biol. 31: 315-462.
Collis, P., Antoniou, M. and F.Grosveld. 1990. Definition of the minimal requirements within the human beta-globin gene and the dominant control region for high level expression. EMBO J., 9, 233-240.
Cook, P., and I.Brazell. 1976. Conformational constraints in nuclear DNA. J. Cell Sc. 22: 287-302.
Cook, P., RNA polymerase : structural determinant of the chromatin loop and the chromosome. Bioessays 16: 425-430.
Coté, J., J. Quinn J. Workman and C. Peterson. 1994. Stimulation of GAL4 derivative binding sites to nucleosomal DNA by yeast SWI/SNF complex. Science 265: 53-60.
Coté, J., C. Peterson and J. Workman. 1998. Perturbation of nucleosomes core structure by the SWI/SNF complex persists after its detachment, enhancing subsequent transcription factor binding. PNAS 95: 4947-
Craddock, C.F., P. Vyas, J.A. Sharpe, H. Ayyub, W.G.Wood and D.R. Higgs. 1995. Contrasting effects of alpha and beta globin regulatory elements on chromatin structure may be related to their different chromosomal environments. EMBO J 14: 1718-1726.
Cunningham, J., M. Puruckner, S. Jane, B. Safer, C. Lowrey, and A. Nienhuis. 1994. The regulatory element 3' to the Aγ globin gene binds to the nuclear matrix and interacts with SATB1, a SAR?MAR associated region binding protein. Blood 84: 1298-1308.
Curcio, M.J., P. Kantoff, M.P. Schafer, W.F. Anderson and B. Safer. 1986 Compensatory increase in levels of β^{minor} globin in murine β thalassemia is under translational control. JBC 261: 16126-16132.
Curtin, P., D. Liu, W. Liu, J. Wang and Y.W. Khan. 1989. Human β-globin gene expression in transgenic mice is enhanced by a distant DNase I hypersensitive site. PNAS 86: 7082-7086.
Dasso M., S. Dimitrov, A Wolffe. 1994. Nuclear assembly is independent of linker histones. PNAS 91: 12477-
Dernburg, A., K. Bowman, J. Fung, W. Marshall, J. Philips, D. Agard and J. Sedat. 1996. Perturbation of nuclear architecture by long-distance chromosome interactions. Cell 85: 745-759.

Dickinson, L., T. Joh, Y. Kohwi, and T. Kohwi-Shigematsu. 1992. A tissue specific MAR/SAR DNA binding protein with unusual binding site recognition. Cell 70: 631-731.
Dierks, P., van Ooyen, A., Cochran, M.D., Dobkin, C., Reiser, J. and C. Weissmann. 1983. Three regions upstream from the cap site are required for efficient and accurate transcription of the rabbit beta-globin gene in mouse 3T6 cells. Cell, 32, 695-706.
Dillon, N. and F.Grosveld. 1991. Human γ-globin genes silenced independently of other genes in the β-globin locus. Nature 350: 252-254.
Dillon, N. and F.Grosveld. 1993. Transcriptional regulation of multi-gene loci: multilevel control. Trends genet. 9: 134-137.
Dillon, N., Trimborn, T., Strouboulis, J., Fraser, P. and F. Grosveld. 1997. The effect of distance on long-range chromatin interactions. Mol. Cell 1: 131-139.
Dirks, R., van de Rijke, F., Fujishita, S., van der Ploeg, M. and Raap, A. (1993) Methodoligies for specific intron and exon RNA localisation in cultured cells by haptenized cells and fluorochromized probes. J. of Cell Sci. 104, 1187-1197.
Donze, D., T.M. Townes and J.J. Bieker. 1995. Role of erythroid Kruppel-like factor in human γ- to β-globin gene switching. JBC 270: 1955-1959.
Donze, D., P.H. Jeancake and T.M. Townes. 1996. Activation of δ-globin gene expression by erythroid Krupple-like factor: A potential approach for gene therapy of sickle cell disease. Blood 88: 4051-4057.
Dorer D., and S. Henikhoff. 1994. Expansions of transgene repeats cause heterochromatin formation and gene silencing in Drosophilla. Cell 77: 993-1002.
Driscoll, M.C., Dobkin, C.S. and B.P. Alter. 1989. Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. PNAS 89: 7470-7474.
Dunaway, M. and Dröge, P. (1989) Transactivation of the Xenopus rRNA gene promoter by its enhancer. Nature 341, 657-659.
Eisenberg, J., S. Elgin and R. Paro. 1995. Epigenetic regulation in Drosophila : a conspiracy of silence. In: Chromatin structure and gene expression. S. Elgin, ed. Oxford university press. UK. pp 147-171.
Ellis, J., Talbot, D., Dillon, N. and F.Grosveld. 1993. Synthetic human β-globin 5'HS2 constructs function as locus control regions only in multicopy transgene concatamers. EMBO J., 12, 127-134.
Ellis, J., Tan-Un, K.C., Harper, A., Michalovich, D., Yannoutsos, N., Philipsen, S. and F.Grosveld. 1996. A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human β-globin locus control region. EMBO J., 15, 562-568.
Elliot, J., R. Festenstein, M. Toliani and D. Kioussis. 1995. Random activation of a transgene under the control of a hybrid hCD2 hLCR/Ig enhancer regulatory element. EMBO J. 14: 575-584.
Engel, J.D. 1993. Developmental regulation of humanβ-globin gene transcription: a switch of loyalties? Trends genet. 9: 304-309.
Enver, T., Raich, N., Ebens, A.J., Papayannopoulou, T., Costantini, F. and G. Stamatoyannopoulos. 1990. Developmental regulation of human foetal-to-adult globin gene switching in transgenic mice. Nature 344: 309-
Enver, T., Q. Li, K. Gale, M. Hu, G. May, J. Karlinsky, G. Jiminez, T. Papayannopoulou and F. Constantini. 1994. Analysis of the developmental and transcriptional potentiation functions of the 5'HS2 of the murine β-globin LCR in transgenic mice. Dev. Biol. 165: 574-584.
Epner E., R. Rifkind, P. Marks. 1981. Replication of α and β globin DNA sequences occurs during early S phase in MEL cells. PNAS 78: 3058-3063.
Felsenfeld, G. 1992. Chromatin as an essential part of the transcriptional mechanism. Nature 355: 219-224.
Felsenfeld, G. 1996. Chromatin unfolds. Cell, 86, 13-19.
Festenstein R., M. Toliani, P. Corbella, C. Mamalaki, J. Parrington, M. Fox, M. Jones, and D. Kioussis. 1996. Locus control region function and heterochromatin-induced position effect variegation. Science 271: 1123-
Fiering S., E. Epner, K. Robinson, Y. Zhuang, A. Telling, M. Hu, D. Martin, T. Enver, T. Ley, and M. Groudine. 1995. Targeted deletion of HS2 of the murine β-globin LCR reveals that it is not essential for proper regulation of the β-globin locus. G&D 9: 2203-2213.
Foley, K.P. and J.D. Engel. 1992. Individual stage selector element mutations lead to reciprocal changes in beta- vs. epsilon-globin gene transcription: genetic confirmation of promoter competition during globin gene switching. G&D 6: 730-744.
Forrester, W.C., C. Thompson, J.T. Elder and Groudine, M. 1986. A developmentally stable chromatin structure in the human β-globin gene cluster PNAS 83: 1359-1363.

Forrester, W.C., Novak, U., Gelinas, R. and M. Groudine. 1989. Molecular analysis of the human β-globin locus activation region. PNAS 86: 5439-5443.
Forrester, W.C., Epner, E., Driscoll, M.C., Enver, T., Brice, M., Papayannopoulou, T. and M.Groudine. 1990. A deletion of the human beta-globin locus activation region causes a major alteration in chromatin structure and replication across the entire beta-globin locus. G&D 4: 1637-1649.
Fraser, N.W., P.B. Sehgal and J Darnell. 1978. DRB-induced premature termination of late adenovirus transcription. Nature 272: 590-593.
Fraser, P., Pruzina, S., Antoniou, M. and F.Grosveld. 1993. Each hypersensitive site of the human beta-globin locus control region confers a different developmental pattern of expression on the globin genes. G&D 7: 106-
Fraser, P., J. Gribnau and T. Trimborn. 1998. Mechanisms of developmental regulation in globin loci. Curr. Op. Hemat 5: 139-144.
Furukawa, T., G. Zitnik, K. Leppig, T. Papayannopoulou and G. Stamatoyannopoulos. 1994. Co-expression of gamma and beta globin mRNA in cells containing a single human beta globin locus: results from studies using single-cell reverse transcription polymerase chain reaction. Blood 83: 1412-1419.
Gaensler, K.M., Kitamura, M. and Y.W. Kan. 1993. Germ-line transmission and developmental regulation of a 150-kb yeast artificial chromosome containing the human beta-globin locus in transgenic mice. PNAS 90: 11381-11385.
Garel, A. and R. Axel. 1976. Selective digestion of transcriptionally active ovalbumin genes from oviduct nuclei. PNAS 73: 3966-3970.
Garrick, D., S Fiering, D. Martin, and E. Whitelaw. 1998. Repeat induced silencing in mammals. Nat. Gen. 18: 56-58.
Gasser, S.M. and U.K.Laemmli. 1986. Cohabitation of scaffold binding regions with upstream/enhancer elements of three developmentally regulated genes of D. melanogaster. Cell 46: 521-530.
Gasser, S.M., B. Amati, M. Cardonas and J. Hofman. 1989. Studies on scaffold attachment sites and their relation to genome function. Int. rev. cyt. 119: 57-96.
Georgakopoulos, T., and G. Thireos. 1992. Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. EMBO J. 11: 4145-4152.
Georgakopoulos, T., N. Gounlaki and G. Thireos. 1995. Genetic evidence for the interaction of the yeast transcriptional co-activator proteins GCN5 and ADA2. Mol. & Gen. Genetics 246: 723-728.
Geyer, P. 1997. The role of insulator elements in defining domains of gene expression. Cur. Biol. 7: 242-248.
Giglioni, B., Casini, C. and R. Mantovani, 1984. A molecular study of a family with Greek hereditary persistence of Haemoglobin and β-thalassaemia. EMBO J. 11: 2641-2645.
Gillemans, N., R. Tewari, F. Lindeboom, R. Rottier, T. Wit, M. Wijgerde, F. Grosveld and S. Philipsen. 1998. Altered DNA binding specificity mutants of EKLF and SP1 Show that EKLF is an activator of the β-globin LCR in vivo. In press
Gourdon, G., J. Sharpe, D. Wells, W. Wood and D. Higgs. 1994. Analysis of a 70 kb segment of DNA containing the human • and α-globin genes linked to their regulatory element (HS-40) in transgenic mice. NAR 22: 4139-4147.
Gourdon, G., J. Sharpe, D. Higgs and W. Wood. 1995. The mouse α-globin locus regulatory element. Blood 86: 766-771.
Greaves, D.R., F.D. Wilson, G. Lang and D. Kioussis. 1989. Human CD2 3'-flanking sequences confer high-level, T cell-specific, position-independent gene expression in transgenic mice. Cell 56: 979-986.
Gribnau, J., de Boer, E., Trimborn, T., Wijgerde, M., Milot, E., Grosveld, F. and Fraser, P. (1998). Chromatin interaction mechanism of transcriptional control. Submitted.
Gross, D.S. and W.T. Garrard. 1988. Nuclease hypersensitive sites in chromatin. Annu. Rev. Biochem., 57: 159-197.
Grosveld, G.C., E. deBoer, C.K. Shewmaker and R.A Flavell. 1982. DNA sequences necessary for transcription of the rabbit beta-globin gene in vivo. Nature 295: 120-126.
Grosveld, F., G. Blom van Assendelft, D.R. Greaves and G. Kollias. 1987. Position-independent, high-level expression of the human β-globin gene in transgenic mice. Cell 51: 975-985.
Grosveld, F., N. Dillon and D. Higgs. 1993. The Regulation of Human Globin Gene Expression. In: Balliere's Clinical Haematology: The Haemoglobinopathies. Vol. 6/No. 1, pp. 31-66.
Groudine, M. and H. Weintraub. 1982. Propagation of globin DNAase I -hypersensitive sites in absence of factors required for induction: A possible mechanism for determination. Cell 30: 131-139.

Gu W. and R. Roeder. 1997. Activation of p53 sequence specific DNA binding by acetylation of p53 c-terminal domain. Cell 90: 595-606.
Hampsey M. 1997. A SAGA of histone acetylation and gene expression. TIG 11: 427-437.
Hanscombe, O.,M. Vidal, J. Kaeda, L. Luzzatto, D. Greaves and F.Grosveld. 1989. High level, erythroid specific, expression of the human β globin gene in transgenic mice and the production of human hemoglobin I murine erythrocytes. G&D 3: 1572-1581.
Hanscombe, O., Whyatt, D., Fraser, P., Yannoutsos, N., Greaves, D., Dillon, N. and F.Grosveld. 1991. Importance of globin gene order for correct developmental expression. G&D 5: 1387-1394.
Hebbes, T., A. Thorne, and C. Crane-Robinson. 1988. A direct link between core histone acetylation and transcriptional active chromatin. EMBO J. 7: 1395-1402
Hebbes, T.R., Clayton, A.L., Thorne, A.W. and C. Crane-Robinson. 1994. Core histone hyperacetylation co-maps with generalised DNAase I sensitivity in the chicken β-globin chromosomal domain. EMBO J. 13: 1823-1830.
Heitz, E. 1928. Das chromatin der Moose. Jahr. Wiss. Botanik 69: 762-818.
Henikhoff, S. 1994. Dosage-dependent modification of PEV in Drosophilla. Bioassays 18: 401-409.
Herendeen, D.R., G.A. Kassavetis and E.P. Geiduschek. 1992. A transcriptional enhancer whose function imposes a requirement that proteins track along DNA. Science 256: 1298-1303.
Higgs, D., W. Wood, A. Jarman, J. Sharpe, J. Lida, I. Pretorius and H. Ayyub. 1990. A major positive regulatory region located far upstream of the human α globin gene locus. G&D 4: 4679-4689.
Hirschhorn, J., S. Brown, C. Clark, and F. Winston. 1992. Evidence that SWI2/SNF2 and SNF5 activate transcription in yeast by altering the chromatin structure. G&D 6: 2288-2298.
Holde van K., J.Zlatanova, G. Arents, and E Moudrianakis. 1995. Elements of chromatin structure. In Chromatin structure and gene expression. ed S. Elgin. IRL press, Oxford, UK. p.p. 8-18.
Holmquist, G. 1987. Role of replication time in the control of tissue-specific gene expression. Am J. Hum. Genet. 40: 151-173.
Hong, L., G. Schroth, H. Matthews, P. Yau, and E. Bradbury. 1993. Studies of the DNA binding properties of histone H4 amino terminus. JBC 268: 305-314.
Hofer, E., Hofer-Warbinek, R. and Darnell J.E. (1982). Globin RNA transcription: A possible termination site and demonstration of transcriptional control correlated with altered chromatin structure. Cell, 29, 887-893.
Hug, B.A., Moon, A.M. and T.J. Ley. 1992. Structure and function of the murine β-globin locus control region 5' HS-3. NAR 20: 5771-5778.
Hug, B.A., Wesselschmidt, R.L., Fiering, S., Bender, M.A., Epner, E., Groudine, M. and T.J. Ley. 1996. Analysis of mice containing a targeted deletion of β-globin locus control region 5' hypersensitive site 3. MCB 16: 2906-2912.
Ito, T., M. Bulger, M. Pazin, R. Kobayshi and J. Kadonaga. 1997. ACF an ISWI- containing and ATP utilising chromatin assembly and remodelling factor. Cell 90: 145-155.
Imbalzano, A.N., H. Kwon, M.R. Green and R.E. Kingston. 1994. Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature 370: 481-485.
Jane, S.M., Ney, P.A., Vanin, E.F., Gumucio, D.L. and A.W. Nienhuis. 1992. Identification of stage selector element in the human γ-globin gene promoter that fosters preferential interaction with the 5' HS2 enhancer when in competition with the β-promoter. EMBO J. 11: 2961-2969.
Janknecht, R., and T. Hunter. 1996. A growing co-activator network. Nature 383: 22-23
Jarman, A.P. and D.R. Higgs. 1988. Nuclear scaffold attachment sites in the human globin gene complexes. EMBO J. 7: 3337-3344.
Jenuwein, T., W. Forrester, L. Fernandez-Herrero, G. Laible, M. Dull, and R. Grosschedl. 1997. Extension of chromatin accessibility by nuclear matrix attachment regions. Nature 385: 269-272.
Jeppesen, P. and B.M. Turner. 1993. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. Cell 74: 281-289.
Jimenez, G., K.B Gale and T. Enver 1992. The mouse β-globin locus control region: hypersensitive sites 3 and 4. NAR 20: 5797-5803.
Jimenez, G., S. Griffiths, A. Ford, M. Greaves and T. Enver. 1992. Activation of the β-globin LCR precedes commitment to the erythroid lineage. PNAS 89: 10618-10622.
John, B. and G. Miklos. 1979. Functional aspects of satellite DNA and heterochromatin. Int. Rev. Cytol. 58: 1-
Jones, B.K., B.R. Monks, S.A. Liebhaber and N.E. Cooke. 1995. The human growth hormone gene is regulated by a multi-component locus control region. MCB 15: 7010-7021.

Kadosh, D., and K. Struhl. 1997. Repression by Ume6 involves recruitment of a complex containing Sin3 co-repressor and Rpd3 histone deacetylase to target to promoters. Cell 89: 365-371.
Kadosh, D., and K. Struhl. 1998. Histone deacetylase activity of Rdp3 is important for transcriptional repression in vivo. G&D 12: 797-805.
Karpen, G. 1994. Position effect variegation and the new biology of heterochromatin. Cur. Opin. Gen. Dev. 4: 281-291.
Kaul, D.K., Fabry, M.E., Nagel, R.L. 1996. The pathophysiology of vascular obstruction in the sickle syndromes. Blood reviews 10: 29-44.
Kellum R., and P. Schedl. 1991. A position effect assay for boundaries of higher order chromosomal domains. Cell 64: 941-950.
Kellum R., J. Raff and B. Alberts. 1995. Heterochromatin protein 1 distribution during development and during cell cycle in Drosophilla embryos. J. Cell Sci. 108: 1407-1418.
Kennison, J. and J. Tamkun. 1988. Dosage dependent modifiers of Polycomb and Antennapedia mutations in Drosophilla. PNAS 85: 8136-8140.
Kim C.G., E.M. Epner, W.C. Forrester and M. Groudine. 1992. Inactivation of the human β globin gene by targeted insertion into the β globin locus control region. G&D: 928-937.
Kioussis, D., E. Vanin, T. deLange, R.A. Flavell and F.G. Grosveld. 1983. Beta-globin gene inactivation by DNA translocation in γβ-thalassemia. Nature 306: 662-666.
Kitsberg D., S. Selig, I. Keshet & H. Cedar. 1993. Replication structure of the human β-globin gene domain. Nature 366: 588-595.
Kollias, G., N. Wrighton, J. Hurst and F. Grosveld. 1986. Regulated expression of human A gamma-, beta-, and hybrid gamma beta-globin genes in transgenic mice: manipulation of the developmental expression patterns. Cell 46: 89-94.
Kollias, G., J. Hurst, E. deBoer and F. Grosveld. 1987. The human beta-globin gene contains a downstream developmental specific enhancer. NAR 15: 5739-5747.
Kong, S., Bohl, D., Li, C. 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. MCB 17: 3955-3965.
Kornberg, R.D. and J.O. Thomas. 1974. Chromatin structure: oligomers of the histones. Science, 184, 865-868
Kornberg, R.D. 1974. Chromatin structure: a repeating unit of histones and DNA. Science 184: 868-871.
Kornberg, R.D. 1977. Structure of chromatin. Ann. Rev. Biochem. 46: 931-954.
Kotkow, K. and S. Orkin. 1996. Complexity of the erythroid transcription factor NF-E2 as revealed by gene-targeting of the mouse p18NF-E2 locus. PNAS 93: 3514-3518.
Kulozik, A., S. Bail, A. Bellan-Koch, C. Bertam, E. Krohne and E. Kleihauer. 1991. The proximal element of the β-globin LCR is not functionally required in vivo. J. Clin. Invest. 87: 2142-2146.
Kuo, M.H., J.E. Brownell, R.E. Sobel, T.A. Ranalli, R.G. Cook, D.G. Edmondson, S.Y Roth. and C.D. Allis. 1996. Transcription-linked acetylation by Gcn5p of histone H3 and H4 at specific lysines. Nature 383: 269-272.
Kuo M., J. Zhou, P. Jambeck, C.Allis. 1998. Histone acetyltransferase activity of yGCN5p is required for the activation of target genes in vivo. G&D 12: 627-638
Kwon, H., A. Imbalzano, P.A. Khavari, R.E. Kingston and MR.Green. 1994. Nature, 370, 477-481.
Laemmli, U.K., E. Kas, L. Poljak and Y. Adachi. 1992. Scaffold-associated regions: cis-acting determinants of chromatin structural loops and functional domains. Curr. Opin. Genet. Dev. 2: 275-285.
Laurent, B., M. Treitel, and M. Carlson. 1990. The SNF5 protein of Saccharomyces Cerevisiae is a glutamine and proline rich transcriptional activator that affects expression of a broad spectrum of genes. MCB 10: 5616-
Laurent, B., I. Treich, and M. Carlson. 1993. 1993. G & D 7: 583-591.
Leder, A., A. Kuo, M. Shen and P. Leder. 1992. In situ hybridisation reveals co-expression of embryonic and adult β-globin genes in the earliest murine erythrocyte progenitors. Development 116: 1041-1049.
Leder, A., C. Daugherty, B. Whitney and P. Leder. 1997. Mouse zeta- and alpha-globin genes: embryonic survival, alpha-thalassemia, and genetic background effects. Blood 90: 1275-1282.
Leighton, P., R. Ingram, J. Eggenschwiller, J. Estratiadis, and S Tilghman. 1995. Disruption of imprinting caused by a deletion of the H19 gene region in mice. Nature 375: 34-39.
Leonard, M., M. Brice, J.D. Engel and T Papayannopoulou. 1993. Dynamics of GATA-1 transcription factor expression erythroid differentiation. Blood 82: 1071-1079.
Li, R., J. Knight, S. Jackson, R. Tjian and M Botchan. 1991. Direct interaction between sp1 and the BPV enhancer E2 protein mediates synergistic activation of transcription. Cell: 493-505.

Li, Q. and G. Stamatoyannopoulos. 1994. Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. Blood 84: 1399-1401.
Li, Q and G. Stamatoyannopoulos. 1997 Binary transgenic mouse model for studying the trans control of globin gene switching: evidence that GATA-1 is an in vivo repressor of human epsilon gene expression. PNAS 94: 2444-2448.
Liebhaber, S., Z. Wang, F. Cash, B. Monks and E. Russel. 1996. Developmental silencing of the embryonic • -globin gene: concerted action of the promoter and 3'-flanking region combined with stage-specific silencing by the transcribed segment. MCB: 2637-2646.
Lindenbaum M.H. and F. Grosveld. 1990. An in vitro globin gene switching model based on differentiated embryonic stem cells. G&D 4: 2075-2085.
Liu, Q., J. Bungert and J.D. Engel. 1997. Mutation of gene-proximal regulatory elements disrupts human ε-, γ-, and β-globin expression in yeast artificial chromosome transgenic mice. PNAS 94: 169-174.
Locke J., M. Kotarski, and K. Tartof. 1988. Dosage dependent modifiers of PEV in Drosophilla and a mass action model that explains their effect. Genetics 120: 181-198.
Lohuizen van, M. 1998. Cell. Mol. Life Sci. 54: 71-79.
Lowrey, H., D. Bodine and A. Nienhuis, 1992. Mechanism of DNase I hypersensitivity site formation within the human globin locus control region. PNAS 89: 1143-1147.
Maandag, E., M. Valk, M. Vlaar, C. Feltkamp, A. Berns, H., te Riele. 1994. Developmental rescue of an embryonic lethal mutation in the retinoblastoma gene in chimaeric mice. EMBO J. 13: 4260-4268.
Magram, J., K. Chada and F.Costantini. 1985. Developmental regulation of a cloned adult β-globin gene in transgenic mice. Nature 315: 338-340.
Marcus, G., S. Silverman, S. Berger, J. Horiuchi and L. Guarente. 1994. Functional similarity and physical association between Gcn5 and ADA2, putative transcriptional adapters. EMBO J. 13: 4807-4815.
Marcus, G., J. Horiuchi, N. Silverman, L. Quarante. 1996. ADA5/SPT20 links the ada and spt genes, which are involved in yeast transcription. MCB16:3197-3203.
Marshall, N.F. and D.H. Price. 1992. Control of formation of two distinct classes of RNA polymerase II elongation complexes. MCB 12: 2078-2090.
Marshall, N.F., J. Peng, Z. Xie and D.H. Price. 1996. Control of RNA polymerase II elongation potential by a novel carboxyl-terminal domain kinase. JBC 271: 27176-27183.
Martin, D.I.K., Fiering, S. and M. Groudine. 1996. Regulation of β-globin gene expression: straightening out the locus. Curr. Opin. Genet. Dev. 6: 488-495.
Matrangelo, I., A. Courey, J. Wall, S. Jackson and p. Hough. 1991. DNA looping and the Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. PNAS 5670-5674.
Matzke, M. and J. Matzke. 1998. Epigenetic silencing of plants transgenes as a consequence of diverse cellular defence responses. Cell. Mol. Life Sci. 54: 94-103.
Marshall, N.F. and D.H. Price. 1992. Control of formation of two distinct classes of RNA polymerase II elongation complexes. MCB 12: 2078-2090.
Mignotte, V., J.F. Eleouet, N. Raich and P.H. Romeo. 1989. Cis- and trans-acting elements involved in the regulation of the erythroid promoter of the human porphobilinogen deaminase gene. PNAS 86: 6548-6552.
Miller, I.J. and J.J. Bieker. 1993. A novel, erythroid cell-specific murine transcription factor that binds to the CACCC element and is related to the Kruppel family of nuclear proteins. MCB 13: 2776-2786.
Milot, E., J. Stouboulis, 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.
Mirkovitch, J., M.E. Mirault and U.K. Laemmli. 1984. Organisation of the higher-order chromatin loop: specific DNA attachment sites on nuclear scaffold. Cell 39: 223-232.
Mizzen, C.A., X.J. Yang, T. Kokubo, J.E. Brownell, A.J. Bannister, T. Owen-Hughes, J. Workman, L. Wang, S.L. Berger, T. Kouzarides, Y. Nakatani, and C.D. Allis. 1996. The $TAF_{II}250$ subunit of TFIID has histone acetyltransferase activity. Cell 87: 1261-1270.
Moon, A. and T.J. Ley. 1990. Conservation of the primary structure, organisation, and function of the human and mouse β-globin locus-activating regions. PNAS 87: 7693-7697.
Morley, B., C. Abbot, J. Sharpe, J. Lida, P. Chan-Thomas and W. Wood. 1992. A single β-globin locus control region element 5'HS2 is sufficient for developmental regulation of human globin genes in transgenic mice. MCB 12: 2057-Mueller-Storm, 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.

Muller, H.J. 1930. Types of visible variations induced by X-rays in Drosophilla. J.Genet. 22: 299-334
Neish, A., S. Anderson, B, Schlegel, W. Wei, and J. Parvin. 1998. Factors associated with the mammalian RNA polII holoenzyme. NAR 26 847-853.
Ney, P.A., Sorrentino, B.P., McDonagh, K.T. and Nienhuis, A.W. 1990. Tandem AP-1-binding sites within the human beta-globin dominant control region function as an inducible enhancer in erythriod cells. G&D 4: 993-
Ney, P.A., Sorrentino, B.P., Lowrey, C.H. and Nienhuis, A.W. 1990. Inducibility of the HS II enhancer depends on binding of an erythroid-specific nuclear protein. NAR 18: 6011-6017.
Nicolis, S., C. Bertini, A. Ronchi, E. Moroni, B. Giglioni and S. Ottolenghi. 1991. An erythroid specific enhancer upstream of the enhancer to the gene encoding the cell-type specific transcription factor GATA-1. NAR 19: 5285-5291.
Nuez, B., Michalovich, D., Bygrave, A., Ploemacher, R. and F. Grosveld. 1995. Defective Haematopoiesis in foetal liver resulting from inactivation of the EKLF gene. Nature 375: 316-318.
Ogryzko, V., R. Schiltz, V. Russanova, B. Howard, and Y. Nakatani. 1996. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell 87: 953-959
O'Neill, L.P. and Turner, B.M. 1995. Histone H4 acetylation distinguishes coding regions of the human genome from heterochromatin in a differentiation-dependent manner. EMBO J., 14, 3946-3957.
Orlando, V. and R. Paro. 1995. Chromatin multi-protein complexes involved in the maintenance of transcription patterns. Curr. Opinion Gen. Dev. 5: 174-179.
Orkin, S.H. and H.H. Kazazian, 1984. The mutation and polymorphism of the human beta-globin gene and its surrounding DNA. Ann. Rev. Genet. 18: 131.171.
Orkin, S.H., S.E. Antonarakis and H.H. Kazazian. 1984. Base substitution at position -88 in a beta-thalassemic globin gene. Further evidence for the role of distal promoter element ACACCC. JBC 259: 8679-8681.
Paro, R. and D. Hogness. 1991. The polycomb protein shares a homologous domain with a heterochromatin associated protein of Drosophilla PNAS 88: 263-267.
Paulson, J. and U. Laemli. 1977. The structure of histone depleted metaphase chromosomes. Cell 12: 817-828.
Pawlik K., and T. Townes. 1995. Autonomous erythroid specific DNAse I HS formed by human β-globin LCR 5'HS2 in transgenic mice. Dev. Biol. 169: 728-732.
Peng, J., Y. Zhu, J.T. Milton and D.H. Price. 1998. Identification of multiple cyclin subunits of human P-TEFb. G&D 12: 755-762.
Perkins, A.C., Sharpe, A.H., and S.H. Orkin. 1995. Lethal β-thalassemia in mice lacking the erythroid CACCC-transcription factor EKLF. Nature 375: 318-322.
Perkins, A.C., Gaensler, K.M.L., and S.H. Orkin. 1996. Silencing of human foetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF. PNAS 93: 12267-12271.
Peschle, C., F. Mavilio, A. Carè, G. Migliaccio, A.R. Migliaccio, G., Salvo, P. Samoggia, S. Petti, R. Guerriero, M. Marinucci, D. Lazzaro, G. Russo and G. Mastroberardino, 1985. Haemoglobin switching in human embryos: asynchrony of $\zeta \rightarrow \alpha$ and $\varepsilon \rightarrow \gamma$-globin switches in primitive and definitive erythropoietic lineage. Nature 313: 235-238.
Peterson, C. and I. Herskowitz. 1992. Characterisation of the yeast SWI1, SWI2 and SWI3 genes, which encode a global activator of transcription. Cell 68 : 573-583.
Peterson K.R., Clegg, C.H., Huxley, C., Josephson, B.M., Haugen, H.S., Furukawa, T., and G. Stamatoyannopoulos. 1993. Transgenic mice containing a 248-kb yeast artificial chromosome carrying the human beta-globin locus display proper developmental control of human globin genes. PNAS 90: 7593-7597.
Peterson, K.R. and G. Stamatoyannopoulos. 1993. Role of gene order in developmental control of human γ and β-globin gene expression. MCB 13: 4836-843.
Peterson, K.R., C. Clegg, P. Navas, E. Norton, T. Kimrough and G. Stamatoyannopoulos. 1996. Effect of deletion of HS3 or HS2 of the human β-globin locus control regioin on the developmnetal regulation of globin gene expression in β-globin locus YAC transgenic mice. PNAS 93: 6605-6609.
Pevny, L., M. Simon, E. Robertson, W. Klein, S_F Tsai, V. Dágati, S. Orkin and F. Constantini. 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. Nature 349: 257-260.
Pevny, L., C.S. Lin, V. D'Agati, M.C. Simon, S.H. Orkin and F. Constantini, 1995. Development of hematopoietic cells lacking transcription factor GATA-1. Development. 121: 163-172.
Pham C.T., D.M. MacIvor, B.A. Hug, J.W. Heusel, and T.J. Ley. 1996. Long range disruption of gene expression by a selectable marker cassette. PNAS 93: 13090-13095.

Philipsen, S., D. Talbot, P. Fraser and F. Grosveld. 1990. The β-globin dominant control region: hypersensitive site 2. EMBO J. 9: 2159-2167.
Philipsen S., S. Pruzina and F. Grosveld. 1993. The minimal requirements for activity in transgenic mice of hypersensitive site 3 of the β-globin locus control region. EMBO J. 12: 1077-1085.
Pillus, L., and M. Grunstein. 1995. Chromatin structure and epigenetic regulation in yeast. In: chromatin structure and gene expression. S. Elgin, ed. Oxford university press. UK. pp 123-146.
Pirotta, V. 1995. Chromatin complexes regulating gene expression in Drosophilla. Cur. Biol. 5: 466-472.
Pirotta, V. 1997. PcG complexes and chromatin silencing. Cur. OP. Gen. Dev 7: 249-258.
Pissard, S., A. M'rad, Y. Beuzard and P. Romeo. 1996. A new type of HPFH: HPFH tunisia beta+ (+C-200)G gamma. Br J. Haematology 95: 67-72.
Ploeg van der, H., L. Konings, A. Oort, M. Bernini and R. Flavell, 1980. Gamma-beta thalassemia: deletion of the gamma and delta genes influences beta globin gene expression in man. Cell 283: 637-642.
Ponez, M., P. Henthorn, C. Stoeckert and S. Surrey. 1988. Globin gene expression in hereditary persistence of foetal haemoglobin and delta-beta thalassaemia. In Oxford Surveys on Eukaryotic genes, (ed. N. McLean). Oxford University Press. D 5, 163-203.
Polach, K.J. and J. Widom. 1995. Mechanism of protein access to specific DNA sequences in chromatin: a dynamic equilibrium model for gene regulation. JMB 254: 130-149.
Proudfoot, N. 1996. Transcriptional interference and termination between duplicated α globin gene constructs suggest a novel mechanism for gene regulation. Nature 322: 562-565.
Pruzina, S., O. Hanscombe, D. Whyatt, F Grosveld and S.Philipsen. 1991. Hypersensitive site 4 of the human β-globin locus control region. NAR 19: 1413-1419.
Ptashne, M. 1986. Gene regulation by proteins acting nearby and at a distance. Nature 322: 697-701.
Raich, N., T. Enver, B. Nakamoto, B. Josephson, T. Papayannopoulou and G. Stamatoyannopoulos 1990. Autonomous developmental control of human embryonic globin switching in transgenic mice. Science 250: 1147-Raich, N., T. Enver, B. Nakamoto, B. Josephson, T. Papayannopoulou and G. Stamatoyannopoulos. 1990. Autonomous developmental control of human embryonic globin gene switching in transgenic mice. Science 250: 1147-1149.
Raich, N., T. Papayannopoulou, G. Stamatoyannopoulos and T. Enver. 1992. Demonstration of a human epsilon-globin gene silencer with studies in transgenic mice. Blood 79: 861-864.
Raich, N., C.H. Clegg, J. Grofti, P.H. Romeo and G. Stamatoyannopoulos. 1995. GATA1 and YY1 are developmental repressors of the human epsilon-globin gene. EMBO J. 14: 801-809.
Rippe, K., P. von Hippel, and J. Langowski. 1995. Action at a distance: DNA-looping and initiation of transcription. Trends Bioch. Sci. 20: 500-506.
Roseman, R., V. Pirotta and P. Geyer. 1993. The su(Hw) protein insulates expression of the Drosophilla. White gene from chromosomal position effects. EMBO J. 12: 435-442.
Ryan, T., R. Behringer, D. Martin, T. Townes, R. Palmiter, and R. Brinster. 1989. A single erythroid specific DNAse hypersensitive site activates high level s of human β-globin expression in transgenic mice. G&D 3: 314-323
Schedl, P. and F. Grosveld. 1995. Domains and boundaries. In Chromatin structure and gene expression. S. Elgin, ed. Oxford university press. UK. pp 174-196.
Sharp, J., P. Can-Thomas, J. Lida, H. Ayyub, W. Wood, and D. Higgs. 1992. Analysis of the human • -globin upstream regulatory element (HS-40) in transgenic mice. EMBO J. 11: 4564-4570.
Sharpe, J.A., D.J. Wells, E. Whitelaw, P. Vyas, D.R., Higgs. and W.G. Wood. 1993. Analysis of the human alpha-globin gene cluster in transgenic mice. PNAS 90: 11262-11266.
Shehee, W.R., P. Oliver and O. Smithies. 1993. Lethal thalassemia after insertional disruption of the mouse major adult beta-globin gene. PNAS 90: 3177-3181.
Shih, D.M., R.J Wall and S.G. Shapiro. 1990. Developmentally regulated and erythroid specific expression of the human em bryonic ß-globin gene in transgenic mice. NAR 18: 5465-5472.
Shih, D.M., R.J. Wall and S.G. Shapiro. 1993. A 5' control region of the human epsilon-globin gene is sufficient for embryonic specificity in transgenic mice. J. Biol. Chem., 268: 3066-3071.
Shivdasani, R.A. and S.H. Orkin. 1995. Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2. PNAS 92: 8690-8694.
in gene targeted GATA-1 mouse embryonic stem cells. Nature Genetics 1:92-98.
Simpson, R.T. 1979. Mechanism of a reversible thermally induced conformational change in chromatin core particles. JBC 254: 10123-10127.

Singh, G., D. Supp, C. Schreiner, J. McNeish, H. Merker, N. Copeland and W. Scott. 1991. Legless insertional mutation: morphological, molecular and genetic characterisation. G&D 5: 2245-2255.
Skow, L., B. Burkhart, F. Johnson, R. Popp, S. Goldberg, W. Anderson, L. Barnett and S. Lewis. 1983. A mouse model for β-thalassemia. Cell 34: 1043-1052.
Spofford, J. 1976. PEV in Drosophilla . In Ashburner, M and E. Novotski. The genetics and biology of Drosophilla, Vol 1c. Acad Press, New York. Pp. 955-1018.
Stanworth, S.J., N.A. Roberts, J.A. Sharpe, J.A. Sloane-Stanley and W. Wood. 1995. Established epigenetic modifications determine the expression of developmentally regulated globin genes in somatic cell hybrids. MCB 15: 3969-3978.
Stamatoyannopoulos, J., C. Clegg and Q. Li. 1997. MCB 17:240-247.
Stamatoyannopoulos G. and A.W. Nienhuis. 1994. Hemoglobin switching. In G. Stamatoyannopoulos, A. Nienhuis (ed.), The molecular basis of blood diseases. W.B. Saunders, Philidelphia, USA.
Steger, D.J. and J.L. Workman 1996. Remodelling chromatin structures for transcription: what happens to the histones? BioEssays 18: 875-884.
Strouboulis, J., N. Dillon and F. Grosveld. 1992a. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. G&D 6: 1857-1864.
Strouboulis, J., N. Dillon and F. Grosveld. 1992b. Efficient joining of large DNA fragments for transgenesis. NAR 20: 6109-6110.
Struhl, K. 1998. Histone acetylation and transcriptional regulatory mechanisms. G&D 12: 599-606.
Studitsky, V.M., D.J. Clark and Felsenfeld, G. 1995. Overcoming a nucleosomal barrier to transcription. Cell 83: 19-22
Strutt, H., and R. Paro. 1997. The polycomb protein complex of Drosophilla has different compositions at different target genes. MCB 17: 6773-6783.
Su, W., S. Jackson, R. Tjian and H. Echols. 1991. DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. G&D 5: 820-826.
Superti-Furga, G., A. Barberis, G. Schaffner and M. Busslinger. 1988. The -117 mutation in Greek HPFH affects the binding of three nuclear factors to the CCAAT region of the γ-globin gene. EMBO J. 7: 3099-3107.
Talbot, D., P. Collis, M. Antoniou, M. Vidal, F. Grosveld and D. Greaves. 1989. A dominant control region from the human β-globin locus conferring integration site-independent expression. Nature 338: 352-355.
Talbot, D., S. Philipsen, P. Fraser and F. Grosveld. 1990. Detailed analysis of the site 3 region of the human β-globin dominant control region. EMBO J. 9: 2169-2178.
Talbot, D. and F. Grosveld. 1991. The 5'HS2 of the globin locus control region enhances transcription through the interaction of a multimeric complex binding at two functionally distinct NF-E2 binding sites. EMBO J. 10: 1391- 1396.
Talbot, D., P. Descombes and U. Schibler. 1994. The 5'flanking region of the rat LAP gene can direct high level, position independent, copy number dependent expression in multiple tissues in transgenic mice. G&D 6: 1857-1868.
Tang, D.C., D. Ebb, R.C. Hardison and G..P. Rodgers. 1997. Restoration of the CCAAT box or insertion of the CACCC motif activate δ-globin gene expression. Blood 90: 421-427.
Taunton J., C. Hassig and S. Schreiber. 1996. A mammalian histone deacetylase related to the yeast transcriptional regulator Rdp3. Science 272: 408-411.
Tewari, R., N. Gillemans, , M. Wijgerde, M. von Lindern,., F. Grosveld and S. Philipsen. 1998. Erythroid Krüppel-like factor (EKLF) is active in primitive and definitive erthroid cells and is required for the function of 5'HS3 of the β-globin locus control region. EMBO J. 17: 2334-2341.
Thompson, E., E. Christians, M. Stinnakre, and J-P. Renard. 1994. SARs stimulate HSP70.1 expression in mouse pre-implantation embryos but not in differentiated tissues. MCB 14: 4694-4703.
Townes, T.M., J.B. Lingrel, H.Y. Chen, R.L Brinster and R.D. Palmiter. 1985. Erythroid-specific expression of human beta-globin genes in transgenic mice. EMBO J. 4: 1715-1723.
Treisman, R., S.H. Orkin and T. Maniatis. 1983. Specific transcription and RNA splicing defects in five cloned beta-thalassemia genes. Nature 302: 591-596.
Trudel, M., J. Magram, L. Bruckner and F. Constantini. 1987. Upstream Gγ-globin and downstream β-globin sequences required for stage specific expression in transgenic mice. MCB 7: 4024-4029.
Tsai, F.Y., G. Keller, F.C Kuo,.M. Weiss, J. Chen, M. Rosenblatt, F.W. Alt and S. Orkin. 1994. An early hematopoietic defect in mice lacking the transcription factor GATA-2. Nature 371: 221-226.

Tsukiyama, T. and C. Wu, 1995. Purification and properties of an ATP-dependent nucleosome remodelling factor. Cell 83: 1011-1020.
Tsukiyama, T. and Wu, C. 1997. Chromatin remodelling and transcription. Curr. Opin. Genet. Dev. 7: 182-191.
Tuan, D., W. Solomon, Q. Li and I.M London. 1985. The "beta-like-globin" gene domain in human erythroid cells. PNAS 82: 6384-6388.
Tuan, D.Y., W. Solomon, I.M. London and D. Lee. 1989. An erythroid specific, development specific enhancer far upstream of the β-like genes. PNAS 86: 2554-2558.
Tuan, D., S., Kong and K. Hu. 1992. Transcription of the hypersensitive site HS:2 enhancer in erythroid cells. PNAS 89: 11219-11223.
Turner, B., A. Birley, and J. Lavender. 1992. Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophilla polytene nuclei. Cell 69: 375-384
Tweeten, K.A. and G.R. Molloy. 1981. Induction of premature termination of transcription of the mouse β globin gene by DRB. NAR 9: 3307-3319.
Ura and Wolffe. 1995. A positive role for nucleosome mobility in the transcriptional activity of chromatin templates: restriction by linker histones. EMBO J. 14: 3752-3765.
Varga-Weisz, P.D. and P.B. Becker. 1995. Transcription factor-mediated chromatin remodelling: mechanisms and models. FEBS 369: 118-121.
Varga-Weisz, P.D., , M. WilmBonte, K. Dumas, M. Mann and P.B. Becker. 1997. Chromatin-remodelling factor CHRAC contains the ATPase ISWI and topoisomerase II. Nature 338: 598-601.
Vidali, G., L.C. Boffa, E.M. Bradbury and V.G. Allfrey. 1978. Butyrate suppression of histone deacetylation leads to accumulation of multi-acetylated forms of histones H3 and H4 and increased DNAase I sensitivity of the associated DNA sequences. PNAS 75: 2239-2243.
Vitale, M., R Di Marzo, R. Calzolari, S. Acuto, D. O'Neill, A. Bank, and A. Maggio. 1994. Bioch. Bioph. Res Com. 204: 413-418.
Wawrzyniak, C. and R. Popp. 1987. Expression of the two adult β-globin genes in mouse yolk sac and fetal liver erythrocytes. Dev. Biol. 119: 299-301.
Wakimoto, B., and M. Hearn. 1990. The effect of chromosomes rearrangements on the expression of heterochromatic genes in chromosomes. 2l of Drosophilla. Genetics 125: 141-151
Weaver, R.F. and C. Weissman. 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15s β globin mRNA precursor and mature 10s β globin mRNA have identical map co-ordinates. NAR 6: 1175-1193.
Weintraub, H. and M. Groudine. 1976. Chromosomal subunits in active genes have an altered conformation. Science 193: 848-856.
Weintraub, H. 1988. Formation of stable transcription complexes as assayed by analysis of individual templates. PNAS 85: 5819-5823.
Weiss, M.J. and S. Orkin. 1995. GATA transcription factors: Key regulators of haematopoiesis. Exp. Haematology 23: 99-107.
Whitelaw, E., S. Tsai, P. Hogben, and S. Orkin. 1990. Regulated expression of globin chains and erythroid transcription factor GATA-1 during erythropoiesis in the developing mouse. MCB: 6596-6606.
Whitney, J.B. 1977. Differential control of the synthesis of two hemoglobin β chains in normal mice. Cell 12: 863-Whyatt, D., A. Karis, I. Harkes, A. Verkerk, N. Gillemans, A. Elefanty, G. Vairo, R. Ploemacher, F. Grosveld and S. Philipsen. 1997. The level of tissue specific factor GATA-1 affects the cell-cycle machinery. Genes & Function 1: 11-
Wijgerde, M., F. Grosveld and P. Fraser. 1995. Transcription complex stability and chromatin dynamics in vivo. Nature 377: 209-213.
Wijgerde, M., J. Gribnau, T. Trimborn, B. Nuez, S. Philipsen, F. Grosveld and P. Fraser. 1996. The role of EKLF in human β-globin gene competition. G&D 10: 2894-2902.
Wilson, C., H.J. Bellen and Gehring, W.J. 1990. Position effects on eukaryotic gene expression. Ann. Rev. Cell. Biol. 6: 679-714.
Wilson, C.J., D.M. Chao, A.N. Imbalzano, G.R. Schnitzler, R.E. Kingston and R.A. Young. 1996. RNA polymerase II holoenzyme contains SWI/SNF regulators involved in chromatin remodelling. Cell 84: 235-244.
Wolffe A. 1994. The transcription of chromatin templates. Cur. Opin. Gen. Dev. 4: 245-
Wolffe A. 1995. In Chromatin. Ed A. Wolffe, 2nd edition, Academic press, Harcourt Brace & company, New York, USA. pp 96-99.

Wong, P.M.C., S. Chung, J.S. White, S.M. Reicheld, M. Patterson, B.J. Clarke, and D.H.K. Chui. 1983. Adult hemoglobins are synthesized in murine fetal hepatic erythropoietic cells. Blood 62: 1280-1288.
Wood, D. 1976. Haemoglobin synthesis during human development. Brit. Med. Bull 32: 282.
Wright, S., E. de Boer, F.G. Grosveld and R.A Flavell. 1983. Regulated expression of the human ß-globin gene family in murine erythroleukemia cells. Nature 305: 333-336.
Wright , S., A. Rosenthal, R. Favell and F Grosveld. 1984. DNA sequences required for regulated expression of ß-globin genes in murine erythroleukemia cells. Cell 38: 265-273.
Wu, C., P. Bingham, K. Livak, R. Holmgren and S. Elgin. 1979. The chromatin structure of specific genes: I evidence for higher order domains of defined DNA sequence. Cell 16: 797-806.
Yang, W.M., C. Inouye, Y. Zeng, D. Bearss and E. Seto. 1996. Transcriptional repression by YY1 is mediated by interaction with a mammalian homolog of the yeast global regulator RPD3. PNAS 93: 12845-12850.
Yang, X.J., Ogryzko, V.V., Nishikawa, J.I., Howard, B.H. and Y. Nakatani. 1996. A p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature 382: 319-324.
Yoshinaga, S., C. Peterson, I. Herskowitz and K. Yamamoto. 1992. Roles of SWI1,SWI2 and SWI3 proteins for transcriptional enhancement by steroid receptors. Science 258: 1598-1603.
Zafarana, G., S. Raguz, S. Pruzina, F. Grosveld, and D. Meyer. 1996. The regulation of human β-globin gene expression: the analysis of HS5 in the LCR. In Haemoglobin switching. G. Stamatoyannopoulos (ed.). Intercept Ltd. Hampshire, UK.
Zhao, K., E. Kas, E. Gonzalez and U.K Laemli. 1993. SAR-dependent mobilisation of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. EMBO J. 12:3237-3247.
Zink, D. and R. Paro. 1995. Drosophilla polycomb group protein regulated chromatin inhibits the accessibility of trans-acting factors to its target DNA. EMBO J. 14: 5660-5671.

Summary

The human β-globin locus serves as a model system to study the dynamics of chromatin in a multi-gene locus. The locus consists of five developmentally regulated genes, and these genes are arranged in the same order as they are expressed during development (5'- ε, $^{G}\gamma$, $^{A}\gamma$, δ, β - 3'). Upstream of the ε-gene is the locus control region (LCR. The LCR was identified as having DNase I hypersensitive sites, and was shown to be important for the correct developmental expression of the β-globin genes. The LCR confers position-independent, high level and copy number dependent expression to linked transgenes in mice.

Several models have been proposed to explain the mechanism of transcriptional activation by the LCR, all these models have in common that the LCR is thought to create an open chromatin conformation. The scanning model assumes that the "activating machinery" traverses along the DNA activating the first gene that it encounters. The random activation model proposes that the transcriptional environment determines random activation of the different genes. The looping model presumes that the LCR stably interacts with individual genes, thereby activating it. Since the distance of the LCR to the genes varies from 5-50 kb it is thought to loop out the intervening DNA. This model predicts that there are two parameters that are important for the level of a given gene: the frequency of LCR/gene interaction and the duration of this interaction (stability)

The looping model has been the working-model in our lab. Recent experiments have all provided additional proof to this idea of transcriptional regulation by the LCR.

The frequency of productive LCR/gene interaction was suggested to be determined mainly by the relative distance of the LCR to the different genes. In **chapter 4** this idea was tested in transgenic mice by placing a second (marked) β gene at different positions in the globin locus. The results showed that placing a gene in a proximal position (to the LCR) results in that gene being transcriptionally activated more often and at the expense of the downstream gene (Dillon et al.,)

The stability of the LCR/gene interaction has been proposed to depend on the transcription factor environment. Targeted disruption of EKLF indicated that reduced levels of this transcription factor results in a decrease in transcription of the β-gene (due to lower LCR/gene stability) with a concomitant increase in active γ transcription (γ gene transcription is known to be EKLF independent) (Wijgerde et al., 1996).

Using novel kinetic analyses we have shown that the human β-globin genes, as well as the mouse β-globin genes are activated via an alternating single gene transcription mechanism. In other words, the LCR activates only one gene at a time (**chapter 5** and **chapter 6**).

Recently we have shown that the mouse β-globin locus is transcriptionally activated in a manner very similar to the human locus. Suggesting that this type of regulation is evolutionary conserved and might be observed in other multi-gene loci (**chapter 6**).

Samenvatting voor vrienden

Het onderzoek waar ik de laatste 4 jaar bij betrokken ben geweest, richt zich op de β-globine genen en met name op de vraag hoe de activiteit van deze genen gedurende de ontwikkeling van de mens gereguleerd wordt. Ik zal in deze samenvatting allereerst uitleggen wat onder de term β-globine verstaan dient te worden en vervolgens wat het belang, de doelstelling en de resultaten van het onderzoek zijn geweest en hoe één en ander in dit proefschrift is terug te vinden.

Organismen –zoals bijvoorbeeld bacteriën en mensen- zijn opgebouwd uit cellen. Bacteriën zijn eencellige organismen, de mens is opgebouwd uit meerdere cellen (drie miljard). Maar de mens ontstaat uit één cel, de zygote, die ontstaat na de versmelting van de eicel met het zaadje. Er vinden dus een groot aantal celdelingen plaats tijdens de ontwikkeling tot een volwassen mens.

Nu is allang duidelijk dat deze celdelingen niet zomaar een verdubbeling van cellen inhouden, cellen blijken namenlijk verschillende eigenschappen te hebben. Zo zijn er vroeg in de ontwikkeling van een organisme verschillende categorieën cellen aantoonbaar: cellen die uiteindelijk de huid gaan vormen, cellen die uiteindelijk het bloed gaan vormen, etc. DNA blijkt een belangrijke rol te spelen in het coördineren van de correcte ontwikkeling van de verschillende cellen.

DNA is aanwezig in vrijwel alle cellen. Het zijn lange strengen moleculen, die chromosomen heten. Een stukje DNA kan een boodschap bevatten, wat in feite een bouwplan is om een eiwit te maken. Zo'n stukje DNA met een boodschap, oftewel een stukje DNA dat een eiwit maakt, noemen we een gen. Het DNA van de mens bevat een groot aantal (~ 80.000) van dit soort stukjes DNA (genen dus). Een stuk DNA met een aantal genen erop, die betrokken zijn bij vergelijkbare processen noemen we een locus van genen.

De genen die ik bestudeerd heb maken het eiwit β-globine. Samen met α-globine vormt β-globine het hemoglobine. Hemoglobine is een eiwitcomplex in onze bloedcellen dat er voor zorgt dat zuurstof van de longen naar de verschillende organen in ons lichaam wordt vervoerd.

Indien er iets fout gaat bij het maken van het globine eiwit onstaat er een bloedcel die deze taak niet naar behoren kan uitvoeren. Er zijn twee groepen patiënten met zulke bloedcellen: patiënten die wel een globine eiwit maken, echter het eiwit heeft de verkeerde samenstelling en functioneert daarom niet en patienten die helemaal geen globine eiwitten maken. In beide gevallen de is aangetoond dat dit wordt veroorzaakt door foutjes in het DNA. Het doel van het onderzoek is om te begrijpen hoe het het β-globine locus gereguleerd wordt.

Het β-globine locus bestaat uit vijf verschillende genen die gedurende de ontwikkeling van de mens worden geactiveerd om het β-globine eiwit te maken. Het gen dat de eerste acht weken van de ontwikkeling actief is heet ε. Vervolgens zijn tot aan de geboorte de genen $^{G}\gamma$ en $^{A}\gamma$ actief. Na de geboorte worden δ en β geactiveerd. Er zijn gedurende de ontwikkeling dus een aantal stadia, dat er meerdere genen actief zijn.

De regulatie van de globine genen is, zoals voor een groot deel van alle genen, met name afhankelijk van een stukje DNA dat direct voor het gen zit en promoter wordt genoemd. De promoter van een gen bevat vaak een aantal stukjes DNA waaraan eiwitten zich kunnen binden. Het zich binden van eiwitten aan de promoter zorgt er voor dat het gen actief wordt of juist inactief, afhankelijk van de eiwitcombinatie die zich op dat moment bindt. Een aantal jaren geleden is aangetoond dat de globine genen naast de promoter ook afhankelijk zijn van een stuk DNA dat relatief ver van de genen ligt: de Locus Control Region (LCR). De globine genen worden alleen op het juiste tijdstip geactiveerd of geïnactiveerd als er een LCR aanwezig is.

De doelstelling van het onderzoek waar ik de laatste vier jaar bij betrokken ben geweest was voornamelijk het verkrijgen van meer inzicht in de activatie van de 5 verschillende globine genen en de rol van de LCR in dit proces. Er was een model geformuleerd dat verklaarde hoe de LCR de verschillende genen van het β-globine locus zou kunnen activeren. In dit model functioneert de LCR als een eenheid en gaat het een interactie aan met de promoter van slechts één globine gen. Door deze interactie wordt dit globine gen geactiveerd. Het DNA dat tussen de LCR en het gen ligt vormt dan als het ware een lus en heeft geen functie. De totale hoeveelheid genprodukt die gemaakt wordt in dit model is afhankelijk van de frequentie (hoe vaak gaan en de LCR en de promoter een interactie aan) en de stabiliteit van deze interactie (hoe lang blijft de LCR vastzitten aan de promoter). Tevens impliceert het model dat als er genprodukten worden aangetoond van twee verschillende genen, dit het gevolg is van een wisselende activatie van deze twee genen (de LCR activeert bijv. eerst een aantal minuten het γ gen om vervolgens in dezelfde bloedcel het β-gen te activeren). De proeven beschreven in dit proefschrift beproeven dit model aan de hand van experimenten met het menselijke β-globine locus in muizen, maar ook door studie van het β-globine locus van de muis zelf.

Hoofdstuk 1 en 2 van dit proefschrift vormen de inleiding. Daarin wordt allereerst een inleiding gegeven in chromatine (DNA ingepakt in eiwitten), hoe is het opgebouwd en wat is de functie. Hoofdstuk 2 geeft een samenvatting van de literatuur van het β-globine onderzoek

van de laatste 10 jaar. In Hoofdstuk 3 worden proeven beschreven waarbij de LCR gemuteerd wordt en wordt het effect van deze mutaties bestudeerd. De LCR bestaat uit vijf eenheden, die hypersensitive sites (HS) heten, HS1, HS2, HS3, HS4 en HS5. We hebben bestudeerd wat het effect is op de activiteit van de genen als één van de HS's wordt weggehaald. De resultaten laten zien dat het niet mogelijk is om stukjes DNA weg te laten uit de LCR, omdat de verschillende globine genen dan niet meer naar behoren worden gereguleerd. De resultaten suggereren dat de LCR inderdaad gezien moet worden als een eenheid die alleen zijn functie uitoefent als die eenheid onaangetast is. In Hoofdstuk 4 hebben we onderzocht of de frequentie van LCR-gen interactie bepaald wordt door de afstand tussen de LCR en de genen. In onze proef werd een extra gen op twee verschillende plaatsen in het globine locus geplaatst en werd bepaald of de afstand tot de LCR van invloed was op de activiteit van het gen. Het werd duidelijk dat de afstand inderdaad de frequentie van de de interactie tussen LCR en promoter bepaalt. Het gen dat dichtbij de LCR geplaatst was, bleek veel vaker actief dan het gen dat verder van de LCR was geplaatst. In Hoofdstuk 5 wordt aangetoond dat de LCR slechts één gen tegelijk kan activeren. We zijn tot deze conclusie gekomen door een stofje te gebruiken dat een gen kan uitschakelen. Door vervolgens dit stofje weer weg te wassen nadat het gen uitgeschakeld is, kan het gen weer geactiveerd worden. Indien de LCR slechts één gen kan activeren verwachten we direkt na het wegwassen van het stofje voornamelijk activiteit van één enkel gen. Als de LCR echter meerdere genen tegelijk kan activeren verwachten we direkt na het wassen activiteit bij meerdere genen. We hebben kunnen aantonen dat onze hypothese, dat de LCR slechts één gen activeert juist is. In Hoofdstuk 6 laat ik zien dat het mechanisme van gen-activatie geconserveerd is tijdens de evolutie. Het globine locus in de muis, dat sterk lijkt op het globine locus in de mens, werd uitvoerig bestudeerd. Ook in het globine locus van de muis activeert de LCR slechts één gen door een interactie met de promoter aan te gaan.

List of abbreviations

Bp	Base pair
DNA	Deoxy-nucleic acid
gDNA	Genomic DNA
HEL	Human erythroid cell line
HS	Hypersensitive sites
IVS	Intervening sequence (intron)
Kb	Kilo base pairs (i.e. 10^3 base pairs)
KD	Kilo Dalton
K562	Human erythroid cell line
LCR	Locus control region
MAR	Matrix attachment region
MEL	Mouse erythroleukemia line
Mb	Mega base (10^6 base pairs)
Nm	Nano meter
PE	Position effect
PEV	Position effect variegation
RNA	Ribonucleic acid
SAR	Scaffold attachment region
TF	Transcription factor
TK	Thymidine kinase
YAC	Yeast artificial chromosome

Curriculum Vitae

Name	Tolleiv O. Trimborn
Born	15 January 1967, Oldebroek, The Netherlands
1994 to date	PhD program at Erasmus University in Rotterdam at the lab of Professor Frank Grosveld
1993 - 1994	Research experience at the University of Amsterdam (VU), using transposons to tag resistance genes in Tomato plants
1986 - 1993	Biology at the University of Leiden (MSc)
1986 - 1992	Chemistry at the University of Leiden (MSc)
1979 – 1986	VWO at the Lambert Francken College, in Elburg

Publications

* **Milot E., J. Strouboulis, T. Trimborn, F. Grosveld and P. Fraser.** 1996. Heterochromatin effects on the frequency and duration of LCR mediated gene transcription. Cell. 87: 105-112.

* **Wijgerde M., J. Gribnau, T. Trimborn, B. Nuez, S. Philipsen, F. Grosveld and P. Fraser.** 1996. The role of EKLF in human beta-globin gene competition. Genes & Dev. 10: 2894-2901.

* **Grosveld F., M. Antoniou, M. Berry, P. Fraser, J. Haley, T. Trimborn and M. Wijgerde.** 1996. Drug discovery and transcriptional control of the human beta-globin gene locus.Genomes, molecular biology and drug discovery. CH 8.

* **Dillon N., T. Trimborn, J. Strouboulis, P. Fraser and F. Grosveld.** 1997. The effect of distance on long range chromatin interactions. Molecular Cell. 1: 131-139.

* **Grosveld F., E. de Boer, N. Dillon, J. Gribnau, T. McMorrow, E. Milot, T. Trimborn, M. Wijgerde and P. Fraser.** 1998. The dynamics of globin gene expression and position effects. In: Epigenetics. Wiley, Chichester (Novartis Foundation Symp 214) pp 67-96.

* **Grosveld F., E. de Boer, N. Dillon, J. Gribnau, E. Milot, T. Trimborn, M. Wijgerde and P. Fraser.** 1998. The dynamics of globin gene expression and genetherapy vectors. Seminars in hematology. 35: 105-109.

* **Fraser P., J. Gribnau and T. Trimborn.** 1998. Mechanisms of developmental regulation in globin loci. Current opinions in Hematology 5: 139-144.

* **Trimborn T., J. Gribnau, F. Grosveld and P. Fraser.** Mechanisms of developmental control of transcription in the murine α and β globin loci. (submitted).

* **Gribnau, J., E. de Boer, T. Trimborn, M. Wijgerde, E. Milot, F. Grosveld and P. Fraser.** Chromatin interaction mechanisms of transcriptional control in vivo. (in press).

Nawoord

Lieve pap en mam, zonder jullie was dit boekje er nooit gekomen. En natuurlijk niet alleen omdat ik er dan niet was geweest. Jullie hebben mij door de jaren heen gestimuleerd en ondanks het feit dat jullie niet bekend waren met het academische wereldje en het sociale leven in Leiden, hebben jullie me altijd door dik en dun gesteund. Hierin moet ik natuurlijk ook Vir betrekken, lieve zus àls jij in Nederland bent is het altijd bere-gezellig, blijf maar.

Eveline, ik ben blij dat ik deze laatste episode niet alleen heb hoeven volbrengen en dat ik deze alinea aan jou kan wijden. Vanaf 1 november 1997, de dag dat we elkaar geregeld hebben, leef ik in een roes. Het sprookje werd nog mooier toen bleek dat je wel zin had om mee te gaan naar Californië. Dit alles heeft er ook aan bijgedragen dat het schrijven van dit boekje zo voorspoedig verliep.

Lab 702, het was een waar genoegen. Ik heb vier fantastische jaren beleefd in dit, in alle opzichten zonnige lab, al was het wel even schrikken (letterlijk) toen ik naast Skippy moest beginnen. Goeie muziek, Irish coffee, Bachi's (from Italy) en bitterkoekjes hebben die onwennige start snel doen vergeten. Ook bij het totstandkomen van dit boekje zijn jullie allemaal nauw betrokken geweest. Met name Tara, Arthur, Joost, Mark, Ali en Eric veel dank hiervoor. Ernie, ik zal het gezellige koffie-leuten wel missen.

Frank en Peter, veel dank voor het vertrouwen dat jullie in mij stelden, toen ik vier jaar geleden bij jullie aanklopte als AIO met een kas vol tomatenplanten. Frank, ik vind het mooi dat jouw deur altijd en voor iedereen openstaat. Ik heb veel geleerd, met name van de gezamenlijke uren achter de computer. Dezelfde (en meer) gedrevenheid die je wetenschappelijk hebt, heb ik gelukkig ook mogen beleven op de squashbaan, op de fiets en op het voetbalveld. Peter, it was a great pleasure working with you. You have convinced me not to cut corners when performing experiments but I do not think we will ever agree on how to play soccer.

Over voetbal gesproken, iedere donderdagavond werd er een titanenstrijd geleverd, toch waren de mooiste potten oud tegen jong oftewel de routine van Ravenelli, DoelpaalDies, Chris, Peter, Niels en spelverdeler Frank tegen het frivole onbezonnen braziliaanse spel van de AIO's.

Squashen met Boerrrrman en Bimo op maandagavond was altijd diep gaan, om vervolgens tot diep in nacht wakker te liggen van de adrenaline. Veel dank voor de lessen jongens.

Kurt, doeshie we zijn beide een totaal verschillende kant op gegaan, en al maakt dit de kans klein, toch hoop ik dat we snel weer eens naast elkaar aan de bench staan te pipetteren. Het is jammer dat je er waarschijnlijk niet bij kunt zijn de 11e.

Zomaar een greep uit een aantal hele mooie herinneringen van de laatste jaren, pantomime, labdag, wandelen op Terschelling (afhankelijk van de grollen van cyber-Johnnie), om te realiseren dat de 7e een unieke afdeling is. In dit rijtje horen natuurlijk ook de drie fantastische zeil-weekenden in Goingarijp thuis. Jan, het duurde een jaartje voordat we elkaar gevonden hadden, maar toen bleek dat we redelijk op één lijn zitten en niet alleen als het om zeilen gaat. En hoe zou dit alles geweest zijn zonder Fokke, Anja, Marella en Wouter (als ie niet aan het wouteren is), ik wil er niet eens over nadenken.

Marieke en Rita natuurlijk bedankt voor alle goede zorgen en het bijkletsen 's ochtends onder het genot van een bakkie kof. Melle (vieze moppen) en Rein voor het inkopen. Mirko voor de cursussen Corel Draw en Adobe en..... Lien voor het geduld door de jaren heen als het weer eens een puinhoop was in mijn stal in het muizenhuis.

Buiten het lab zijn er natuurlijk de Heeren van Koevoet. Ed, Ian, Kees, Ru en Pat met jullie heb ik al zoveel mooie momenten meegemaakt, ik zal dat erg missen als ik in Californië zit.

Arthur en Kees, mijn paranimfen, alvast heel veel dank