Spatial organisation of the β-globin locus the Active Chromatin Hub

Robert-Jan Palstra

Spatial organisation of the β-globin locus the Active Chromatin Hub

Ruimtelijke organisatie van het β-globine locus de actieve chromatine hub

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus

Prof.dr. S.W.J. Lamberts

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op
woensdag 5 januari 2005 om 13.45 uur

door

Robert-Jan Theo Sijtze Palstra

geboren te 's-Gravenhage

Promotiecommissie

Promotor:	**Prof.dr. F.G. Grosveld**
Overige leden:	**Prof.dr. C.P. Verrijzer** **Dr. W. Vermeulen** **Dr.ir. D.N. Meijer**
Copromotor:	**Dr. W.L. de Laat**

Dit proefschrift kwam tot stand binnen de vakgroep Celbiologie aan de faculteit der Geneeskunde en gezondheidswetenschappen van het Erasmus Medisch Centrum van de Erasmus Universiteit te Rotterdam. De vakgroep maakt deel uit van het Medisch Genetisch Centrum Zuid-West Nederland. Het onderzoek werd financieel gesteund door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (N.W.O.). Een bijdrage in de drukkosten van dit proefschrift is verkregen van de J.E. Jurriaanse Stichting

SUMMA SCIENTICA NIHIL SCIRE

JOHANN VALENTIN ANDREAE,
DIE CHYMISCHE HOCHZEIT DES CHRISTIAN ROSENCREUTZ
STRASSBURG, ZETNER, 1616

Voor Lencho, Brigitte en mijn ouders

Contents

List of abbreviations

3C	Chromosome Conformation Capture
4-OHT	4-hydroxy-tamoxifen
ACH	Active Chromatin Hub
AGM	Aorta/gonad/mesonephros
ATP	Adenosine triphosphate
bp	Base pairs
CH	Chromatin Hub
ChIP	Chromatin immuno precipitation
CHX	Cycloheximide
CTD	C-terminal domain
DMR	differential methylated domain
DNA	Deoxyribonucleic acid
DNase I	Deoxyribonuclease I
DRB	5,6-dichloro-1β-D-ribofuranosyl benzimidazole
DRED	Direct repeat erythroid-definitive
EKLF	Erythroid kruppel like factor
FISH	Fluorescence *in situ* hybridization
GFP	Green fluorescent protein
GTF	General transcription factor
HAT	Histone acetyl transferase
HDAC	Histone deacetylases
HMG	High mobility group
HMT	Histone methyl transferase
HPFH	Hereditary persistence of fetal hemoglobin
HS	Hypersensitive site
ICR	Imprinting control region
kb	Kilo base pairs (10^3 bp)
LacO	Lac operator
LBD	Ligand binding domain
LCR	Locus Control Region
MEL	Mouse erythroleukemia
mRNA	messenger RNA
OR	Olfactory receptor
PAC	P1 artificial chromosome
PCR	Polymerase chain reaction
PIC	Pre-initiation complex
RNA	Ribonucleic acid
RNAP	RNA polymerase
RNA TRAP	RNA tagging and recovery of associated proteins
rRNA	ribosomal RNA
TBP	TATA-binding protein
Tg	transgenic
TFIIx	Transcription factor II x (e.g. B, D, E, F, H)

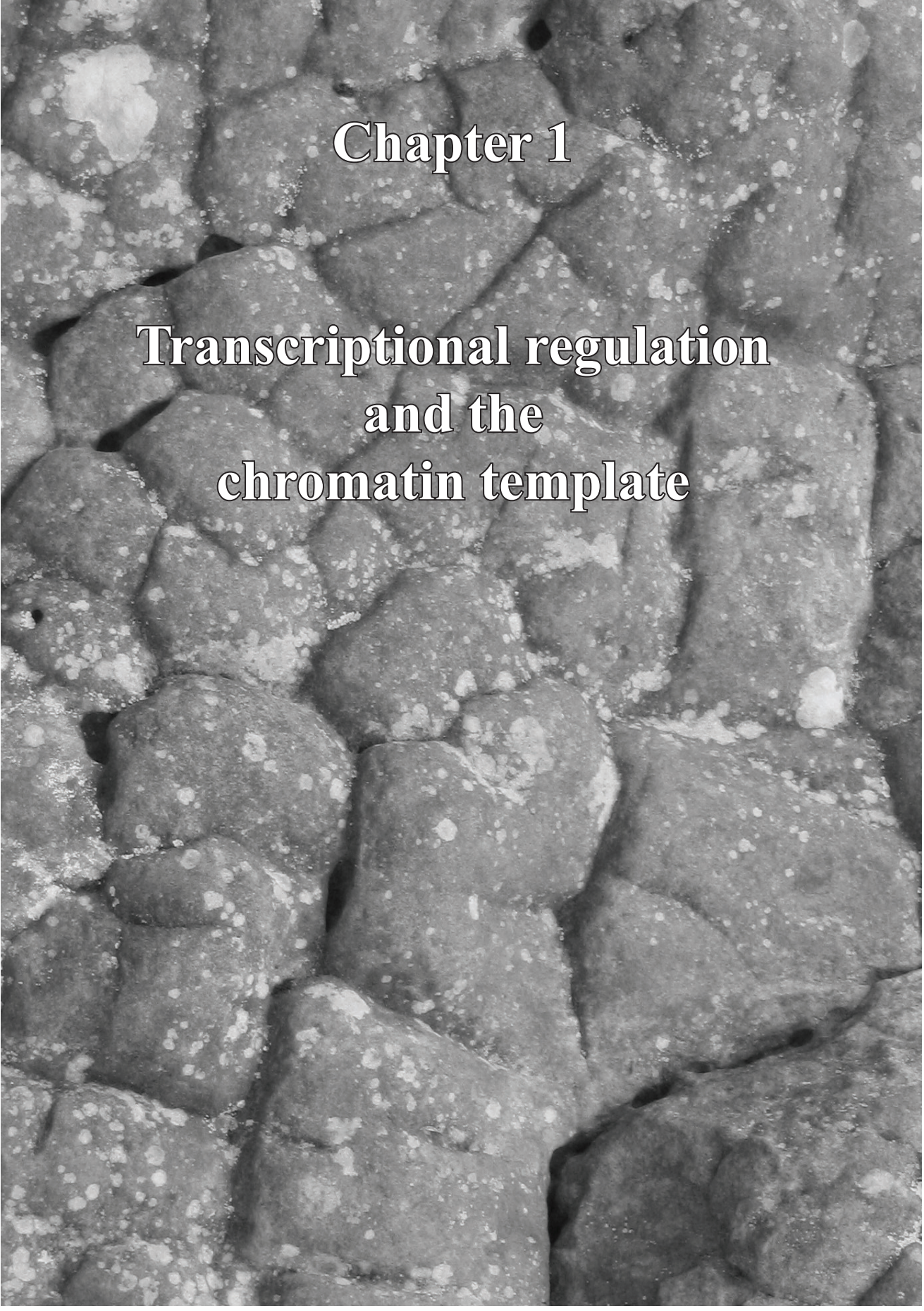
Scope of the Thesis

Multicellular organisms contain a complete set of genes in nearly all of their cells, each cell harbouring the potential to make any protein encoded in their genome. The exact temporal and spatial activation of all these genes is crucial for the identity and survival of the cell. The activation of genes is extensively studied at the level of DNA and chromatin and at the level of chromosome domains and nuclear organisation. However, little is known about the structural organisation of individual gene loci. The aim of this thesis is to unravel the spatial structure of the chromatin fibre of an individual multi gene locus and to investigate how this structure is initiated and modulated during development and transcriptional activation.

Chapter 1 gives an introduction to gene transcription and how this process is influenced by chromatin. Transcription of genes is often regulated by *cis*-regulatory elements that can be located hundreds of kilo bases away from the actual site of transcription and several models are presented that explain how these regulatory sequences might communicate with the promoter sequence. The nucleus is dynamic but nevertheless structurally highly organised and in **chapter 2** an overview of the structural organisation of the nucleus in relation to transcription is given. In this thesis the β-globin locus, which expression is tightly regulated during erythropoeisis, is used as a model system to study transcriptional activation of multi-gene loci. **Chapter 3** introduces the process of erythropoeisis and transcriptional control of the β-globin locus. Current knowledge on nuclear organisation is limited to ultra structural features; how individual gene loci are organised within the nuclear space is largely unknown. Presently methods are emerging that try to address this question. **Chapter 4** discusses the different methods that are available to study the structural organisation of genomic loci.

Chapter 5 describes experiments that address the *in vivo* spatial organisation of the mouse β-globin locus in active and inactive tissues. The data clearly shows that the locus control region (LCR), as well as other regulatory sequences, are in close proximity to the β-globin promoter in the active locus but not in the inactive locus. Intervening inactive sequences do not participate in these interactions and loop out. This spatial clustering of regulatory elements for transcriptional activation is termed an active chromatin hub (ACH). In **chapter 6** these experiments are extended to an earlier time point in development, when a different globin gene is active, showing that proximity of the LCR to a gene correlates with activity of that gene. It is also shown that the ACH is conserved between mice and humans. Moreover, erythroid precursor cells committed to β-globin expression are shown to have a poised pre-structure present, termed the chromatin hub (CH). **Chapter 7** describes the influence of EKLF, an erythroid transcription factor essential for β-globin expression, on the formation of the ACH. The data show that in EKLF knockout mice progression from a CH to an ACH is blocked. Additional experiments show that this effect is not due to a block in differentiation and that EKLF acts directly on the locus.

Each chapter has its own comprehensive specific discussion and in **chapter 8** the general implications of this work and future directions are put forward. All the experimental results are set in a broader context and the implications for our understanding of the organisation of multi gene domains and transcription factories are discussed.



Chapter 1

Transcriptional regulation and the chromatin template

Transcriptional regulation and the chromatin template

The chromatin template

All the information needed for the formation and functioning of an organism is stored as a four "letter" code in DNA. Stretches of different combinations of nucleotides that signify a functional unit, often encoding proteins, are called genes. In eukaryotes the entire DNA has a length of up to 2 meters and is stored in a nucleus with a diameter as small as 3-10 μm. In order to fit the complete DNA into this limited volume it is folded and compacted in a structure called chromatin (Horn and Peterson 2002). The basic building blocks of chromatin are the nucleosomes. The nucleosome consists of an octamer of two copies each of histones H2A, H2B, H3 and H4. About 147 base pairs (bp) of DNA are wrapped 1.7 times around this core histone octamer in a left-handed super helix (Richmond and Davey 2003). Each histone contains two structural and functionally different domains, a globular core domain mediating histone-histone and histone-DNA interactions and a structurally less defined histone tail that can be modified in several ways (see below). A stretch of approximately 10-60 bp of linker DNA is found between two neighbouring nucleosomes, which can be bound by a linker nucleosome such as histone H1 (figure 1.1).

On chromosomes, thousands of nucleosomes are organised on a continuous DNA fibre to form the primary compaction unit of chromatin. *In vitro,* in a low-ionic-strength buffer, this array adopts a "beads-on-a-string" like structure. However when the cation concentration is raised to concentrations similar to those found in the nucleus, these arrays fold into a 30-nm fibre, which can self-associate to form higher order (>30-nm) fibres (Hansen 2002). Condensation is an intrinsic property of the nucleosomal array, it is dependent on the histone tails and is stabilised by linker histones and other proteins (Adkins et al. 2004).

Although it is clear that chromatin primarily exists in a highly ordered state *in vivo*, the precise nature of its structure is still unclear. Microscopy studies revealed only occasional stretches of 30-nm fibre *in vivo* while a substantial amount of 100-nm wide fibres called chromonema fibres were observed (Belmont et al. 1989).

Early in the history of the study of the nucleus two general types of chromatin were recognised; heterochromatin and euchromatin. Chromatin that remains condensed and darkly stained throughout the cell cycle as observed in cytological studies was termed heterochromatin. Heterochromatin was associated with telomeres and pericentric regions of chromosomes. Subsequently several other hallmarks of heterochromatin were identified such as high repetitive sequence content, low gene density and the inaccessibility to nucleases (table 1.1)(Richards and Elgin 2002). Heterochromatin is considered to be the inactive state of chromatin, although some genes appear to depend on heterochromatin for their proper transcription (e.g. Devlin et al. 1990; Eberl et al. 1993; Weiler and Wakimoto 1998; Copenhaver et al. 1999). Two kinds of heterochromatin can be distinguished; constitutive heterochromatin found at pericentric regions that is permanently silent and facultative heterochromatin that can be induced to become active. Active chromatin is called euchromatin and was recognised in early cytological studies as being the part of chromatin that de-condensed during interphase. Additional features of euchromatin are among others a high gene density and accessibility to nucleases (table 1.1)(Richards and Elgin 2002). Recent analyses of lymphoblastoid chromatin

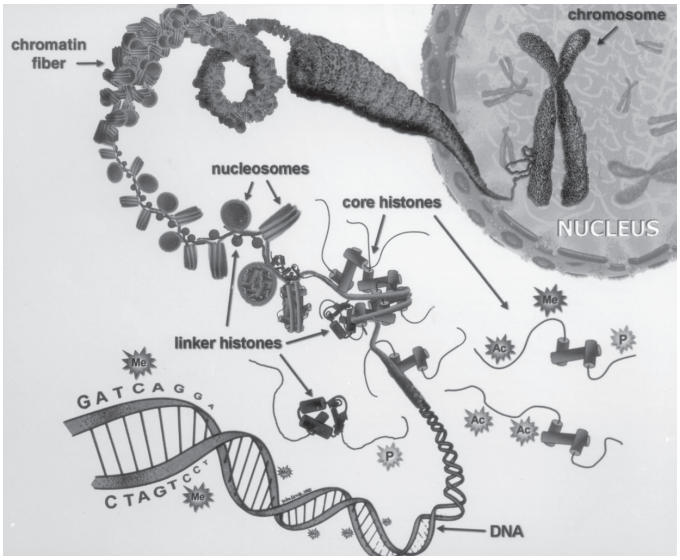


figure 1.1: Organisation of DNA in chromosomes.
DNA is wrapped around nucleosomes, which associate to form a 30nm fibre. Further compaction and loop formation results in higher-order structure and finally mitotic chromosome condensation compacts the chromatin further into

structure across the human genome showed that compact chromatin fibres are composed of heterochromatin and contains the regions of low gene density of which some are active genes (Gilbert et al. 2004).

Conversely, open chromatin fibres contain the regions of highest gene density, are cytological decondensed and have a distinctive nuclear organization, however this does not correlate with gene expression, since inactive genes can be found in the domains of open chromatin (Gilbert et al. 2004). These observations suggest that, at least at this level of resolution, the organisation of the genome is more complex than the simplistic bi-modal separation into euchromatin and heterochromatin.

Transcription on the chromatin template

Transcription on "naked" DNA *in vitro* can be achieved by the combined action of a limited set of general transcription factors. Protein encoding genes are transcribed into an intermediate called messenger-RNA (mRNA) by an enzymatic RNA polymerase II holocomplex. The RNA polymerase II holocomplex is a large protein complex composed of 12 subunits whose sequences are conserved among many diverse species. The largest subunit of the RNA polymerase II holocomplex has a carboxy terminal domain (CTD), which plays a key role in the regulation of transcription initiation and coordination of co-transcriptional RNA processing events. Transcription starts upon assembly of a pre-initiation complex (PIC) on a special recognition DNA sequence, called a promoter. The PIC contains

Table 1.1. Distinctions between Euchromatic and Heterochromatic Domains

Feature	Euchromatin	Heterochromatin
Staining/packaging in interphase	Dispersed/decondensed	Appears condensed, heteropycnotic
DNA sequence	Predominantly unique	Predominantly repetitive (satellites, derivatives of viruses, transposons,etc.)
Presence of genes	High/variable density	Low density
Meiotic (reciprocal) Recombination	Normal frequency	Low frequency
Replication timing	Throughout S phase	Late S phase
Chromatin structure	DNase I HSs, irregular nucleosomes, accessible to nucleases	Loss of DNase I HSs, regular nucleosomal array, less accessible to nucleases
Activity state Euchromatic genes Heterochromatic genes	 Genes inducible Genes silenced (variegated)	 Genes silenced (variegated) Genes inducible
Characteristic modifications	Histone hyperacetylation Histone H3-meK4 present Cytosine (CpG) hypomethylation	Histone hypomethylation Histone H3-meK9 present Cytosine (CpG) hypermethylation

Adapted from (Richards and Elgin, 2002).

a well-characterised collection of general transcription factors. In general TFIID is the first PIC component to assemble on the promoter. One of its components, the TATA-binding protein (TBP), recognises and binds the TATA-box frequently found in promoters (Muller and Tora 2004). Binding of TFIIA and TFIIB is thought to follow binding of TFIID. An early checkpoint for correct promoter complex assembly is the recognition of this complex at the TATA box by an RNA polymerase II- TFIIF complex (Svejstrup 2004). Entry of TFIIE into the complex probably leads to establishment of a correct RNA polymerase II configuration that is prepared for the subsequent consequences of open complex formation and promoter melting catalysed by TFIIH (Reese 2003). In a still poorly understood manner, CTD-phosphorylation, mediated by a kinase component of TFIIH, triggers promoter clearance and thereby defines the initiation-to-elongation transition. Mediator, a complex composed of approximately 25 proteins, binds RNA polymerase II at least partly via its CTD and enhances TFIIH –mediated CTD phosphorylation up to several hundred fold (Svejstrup 2004). The presence of nucleosomes on the DNA template poses a considerable barrier to transcription. Depending on the promoter structure (positioned or randomly situated nucleosomes, higher order chromatin arrangement, etc.) and general chromatin environment (chromatin neighbourhood) of genes a promoter needs to be made accessible by chromatin remodelling factors and transcriptional activators before GTFs and RNAPII can gain access. Over the last

couple of years, chromatin immunoprecipitation (ChIP) experiments have shown that the order of recruitment of chromatin remodelling factors, transcriptional activators, GTFs, and RNAPII can vary significantly depending on the gene in question (e.g. Agalioti et al. 2000; Shang et al. 2000; Hatzis and Talianidis 2002). Not only recognition of the promoter and assembly of a PIC can be inhibited by chromatin but also "read through" of the polymerase is hampered. Factors such as Spt4/5, Spt6, FACT, Elongator and COMPASS play an as yet not entirely defined role in dealing with chromatin during transcription elongation (Shilatifard 2004; Svejstrup 2004).

Higher compaction of nucleosome arrays poses even more formidable barriers to basal transcription. Although it is not known what the exact *in vivo* template for transcription is, in yeast nucleosomes are depleted from active regulatory regions (Bernstein et al. 2004; Lee et al. 2004). Actively transcribed, decondensed regions are thought to retain higher-order structure beyond the 30 nm fibre as judged by cytological studies. For example, upon induction the mouse *HoxB* loci decondenses to an extend that is consistent with unwinding to a 30 nm fibre (Chambeyron and Bickmore 2004a). However a region enriched in open chromatin fibres, as identified in lymphoblastoid cells, has a cytological compaction level that is approximately 10-fold higher tha the 30 nm fibre (Gilbert et al. 2004). Eukaryotes have found several ways to modulate the inhibitory effects posed upon transcription by chromatin and even take advantage of the presence of chromatin to generate a dynamic range of transcriptional regulation.

Modification and remodelling of the chromatin template

The tails of histones can be modified in several ways. One of the most studied modifications is acetylation of the lysine residues present in the histone tails. As early as 1964 Allfrey and colleagues described that the use of acetylated histones in an *in vitro* transcription assay resulted in an increase in transcription as compared to the use of unacetylated histones (Allfrey et al. 1964).

Several transcriptional coactivators, which are needed for efficient transcription on a chromatin template, turned out to possess histone acetyl transferase (HAT) activity (e.g. Brownell and Allis 1996; Brownell et al. 1996; Mizzen et al. 1996). The opposing histone deacetylase (HDAC) activity is often associated with repressive protein complexes (Narlikar et al. 2002). The exact biochemical mechanism behind the activating role of histone acetylation remains elusive but likely mechanisms involve changes in higher order chromatin folding and increasing the affinity of chromatin for other positive modifiers.

Acetylation is not the only post-translational modification known to occur on the histone tails. Other modifications identified include methylation, phosphorylation, ubiquitination and poly-adenyl ribosylation (Peterson and Laniel 2004). Histone modifications can have many different effects. While acetylation of histone tails is generally linked to active chromatin, methylation of histones can either be connected to silenced or active chromatin states depending on which residue is modified (Bannister et al. 2002). Adding to the complexity is for example the fact that each lysine residue can accept one, two or even three methyl groups and an arginine can either be mono- or di-methylated each resulting in a different chromatin state (Bannister et al. 2002). The enzymatic activity needed to place

these modifications is associated with several transcriptional activators and repressors and the modification is often targeted to specific residues in the histone tails (Peterson and Laniel 2004).

The chromatin modifying enzymes are targeted to specific chromatin locations by several mechanisms. For example, histone methyl transferases (HMT) are associated with the elongating RNA polymerase II holocomplex directing histone H3 k4 or k36 methylation to actively transcribed regions (e.g. Miller et al. 2001; Strahl et al. 2002). The enzymes can also be targeted to specific chromatin domains because they recognise specific combinations of previously established histone modifications or interact with other proteins that recognise specific combinations of histone modifications. A good example of the intricate interplay between chromatin marks and chromatin modifying enzymes is heterochromatin spreading. Binding of HP1 to chromatin is dependent on the presence of a methylation mark on lysine 9 of histone H3 and is mediated by the chromodomain of HP1. Su(var)3-9, a histone H3 lysine 9 methylase, binds to a different domain present in HP1 called the chromo shadow domain. In this way the methylation mark is perpetuated resulting in spreading of HP1 binding and heterochromatin formation (Rea et al. 2000; Bannister et al. 2001). The methylation of histone H3 on lysine 9 is influenced by pre-existing modifications of histone H3 and affects other histone modifications, implying a set of functional interactions. The notion that different combinations of chromatin marks can target specific enzymes, which generate unique states of activity, to chromatin domains led to the proposal of a "histone code" (Strahl and Allis 2000; Jenuwein and Allis 2001; Turner 2002). It is currently unclear if this code serves an "instructional code" or is the end result of a transcriptional program providing a "memory" that can be propagated through DNA replication and mitosis, guaranteeing the inheritance of chromatin states to progeny.

Several variants of histones H2A and H3 are known and the replacement of canonical histones with a variant can have profound effects on gene expression. Classical examples of variants associated with a silenced chromatin state are the histone H3 variant CENP-A found at centromeres and macroH2A found at the inactive X chromosome. Histone H2A.Z and H3.3 are found associated with active chromatin and both are deposited in a transcription coupled deposition pathway (Henikoff et al. 2004). Histone modifications associated with active chromatin states are enriched on histone H3.3 while the modifications associated with inactive chromatin are enriched on histone H3 (McKittrick et al. 2004).

A nucleosome positioned over a specific DNA sequence can prevent the binding of transcription factors that have the capability to bind that sequence on naked DNA. Several protein complexes have been identified that facilitate transcription by remodelling or disrupting nucleosomes (Narlikar et al. 2002). These complexes can be divided into three subclasses based on the identity of their ATPase subunits, the SWI/SNF family, the ISWI family and the Mi-2 family. A variety of different assays indicate that SWI/SNF and ISWI, the two best-studied families, act via different mechanisms. They also have a different substrate specificity as is by the fact that an ISWI complex is unable to remodel a tail-less histone while a SWI/SNF complex is able to do so (Clapier et al. 2001; Langst and Becker 2001). The ATP-dependent remodelling complexes all have the capability to expose nucleosomal DNA. Two main mechanisms have been proposed; sliding along DNA and induction of a conformational change of the DNA nucleosome complex and both mechanisms appear to play a role. ATP-dependent remodellers and covalent modifiers work close together to

regulate gene expression. In an additional level of transcriptional regulation, interaction between ATP-dependent remodellers, histone modifiers, and transcription factors can modulate each other's activities in the context of activation and repression of genes (Narlikar et al. 2002). Because both ATP-dependent remodelling complexes and histone modifying complexes can bind non-specifically to DNA, these complexes have to be targeted to specific loci. It has been shown that transcription factors that recognise and bind to specific DNA sequences are able to interact with these complexes. Sequence specific transcription factors can be ubiquitously expressed or in a temporal or cell type specific manner thereby allowing different transcription programs at different times during development and in different cell types in an organism. Transcriptional activators such as nuclear receptors, EKLF, C/EBPβ, c-Myc and MyoD target SWI/SNF and HATs while transcriptional repressors, for example Ikaros, can target ATP-dependent remodellers and HDACs to specific chromatin regions via their recognition sequences. The difference between activators and repressors is not strict and many transcription factors, like e.g. GATA (Rodriguez et al. submitted) have the ability to target either HATs or HDACs to a chromatin region, depending on the context or developmental stage. Not surprisingly, binding sites for sequence specific transcription factors are primarily found in regulatory elements associated with genes.

Cis-regulatory elements

Tight regulation of gene expression is crucial for the survival of a eukaryotic organism. The eukaryotic genome has evolved several DNA elements that ensure the proper temporal and spatial expression of genes present on the same chromatin fibre. These *cis*- regulatory elements are generally stretches of DNA 100-500 bp long, consisting of a collection of recognition sequences for transcription factors and can be detected as DNase I hypersensitive sites (box). Depending on the experimental assay used several distinct functional elements can be recognised. Their functions are often less defined *in vivo* and a given element can have more properties at once, which may depend on its genomic context.

Promoters

Eukaryotic promoter elements are on average 100–200 bp in size, and are found directly upstream of the coding region of genes. Their purpose is to recruit, position and stabilise the transcriptional apparatus and ensure the proper initiation of transcription. The core of the promoter, where the PIC is assembled, is located between approximately –35 and +35 relative to the transcription start site of metazoan genes. In *Drosophila* two distinct classes of core promoters can be identified. About half of the promoters have an AT-rich sequence called the TATA box combined with an initiator (Inr) element overlapping the start site (Burke et al. 1998; Kutach and Kadonaga 2000). The TATA box is located 22 bp to 30 bp upstream of the initiation site (Benoist and Chambon 1981) and is recognised and bound by the general transcription factor TBP, which positions the basal transcription machinery and determines the precise site of transcription initiation. The other half of *Drosophila* core promoters contain an Inr element combined with a downstream promoter element (DPE),

Box: DNase I sensitivity

DNA in the eukaryotic nucleus is packaged into chromatin and therefore has a reduced sensitivity to nucleases as compared to naked DNA. Nucleases can be used to probe the state of chromatin compaction. A common enzyme used is the endonuclease DNase I that hydrolyses double- or single- stranded DNA in a relatively random manner. Transcriptionally inactive DNA is packaged into condensed chromatin such that it is relatively unavailable to the DNase I endonuclease. A change in chromatin structure associated with transcriptional competence can be detected as increased sensitivity of the exposed DNA to digestion with DNase I. This increased susceptibility is referred to as general DNase I-sensitivity. DNase I-sensitive regions appear to represent decondensed or "open" chromatin but can currently not be linked to any specific chromatin modification and the specific biochemical events responsible for establishing DNase I sensitivity are unknown. Some small chromatin regions (~200 bp) are exceptionally sensitive to cleavage by very low concentrations of DNase I and are therefore termed as DNase I hypersensitive sites (HS). These HS often contain recognition sites for transcription factors, and mark cis-regulatory elements such as promoters and enhancers. Nuclease-hypersensitive sites occur in species from yeast to humans and are believed to result from chromatin perturbation or even complete removal of nucleosomes. However, HS are also formed by binding of proteins to a DNA template in the absence of chromatin assembly, indicating that sensitivity to nucleases is not always simply a consequence of nucleosome re-organisation or depletion but also depends on a distorted DNA structure induced by the binding of transcription factors (Leach et al. 2001). Moreover binding of GATA-1 induces a strong hypersensitive site and nuclease digestion patterns of the GATA/nucleosome complex closely resemble those of GATA complexed to free DNA (Boyes et al. 1998).

which is, located approximately 30 bp downstream of the start site (Burke et al. 1998; Kutach and Kadonaga 2000). Mammalian core promoters contain a smaller percentage of TATA boxes and the TATA boxes are less often paired with Inr elements than in *Drosophila* core promoters. DPE elements exist in mammalian core promoters but have been difficult to identify. Many mammalian core promoters lack all three elements and contain GC-rich sequences, the so-called CpG-islands, which are frequently associated with housekeeping genes and usually have multiple sites of transcriptional initiation (Smale 2001).

Other elements, to which ubiquitous or tissue and developmental specific transcription factors bind, can be found more upstream of the core promoter. These transcription factors contribute to the assembly and stabilisation of the PIC on the core promoters at least in part by recruiting ATP dependent remodellers and histone modifying enzymes. Architectural transcription factors like high mobility group (HMG) proteins bind to promoters and serve a more structural function by changing the topology and structure of the promoter (Wolffe 1999). In this way they help to build enhanceosomes via interactions with partner proteins (Alvarez et al. 2003) thereby facilitating multiple rounds of transcription initiation through recurrent RNA-polymerase II binding without the need to reassemble a new PIC for every round of transcription (Dieci and Sentenac 2003).

Enhancers and silencers

Enhancers are *cis*-regulatory elements that are capable of enhancing the basal transcription of a linked promoter. They act orientation independent and have the ability to act over large distances in higher eukaryotes; surprisingly, yeast enhancer-like elements (UASs) only work over a short distance up to about 1 kb (Guarente 1988). Enhancers can be found upstream, downstream and even within the transcription unit of a gene and some

of them overlap with promoters. Structurally enhancers are similar to promoters, in the sense that they contain clusters of transcription factor binding sites, but do not contain genic transcriptional start sites. In yeast, enhancers of active genes are associated with the Mediator complex and binding to the enhancer is mediated via different subunits of mediator as the ones interacting with the basic promoter complex. The physical interaction between enhancer and promoter that is established in this way would make Mediator a primary conduit of regulatory information from enhancers to promoters (Kuras et al. 2003). The activity of enhancer elements is not always promiscuous and can be limited to a single specific gene depending on the context of the promoter (Ohtsuki et al. 1998; Butler and Kadonaga 2001).

Silencer elements are similar to enhancers but differ in their action in that they repress transcription of a *cis*-linked promoter (Ogbourne and Antalis 1998). Co-ordinated stage specific and tissue specific expression of a gene often depends on the intricate co-operation between enhancer and silencer elements.

Models of long-range enhancer action

The current view of transcription is based primarily on the concept of "recruitment" (Ptashne and Gann 1997). According to this view, various factors participating in transcription are recruited to promoters via protein-protein interactions with an activator protein. In essence, activator proteins increase the local concentration of the transcriptional machinery near the promoter. This view is largely based on studies in yeast where most activators bind less than a kilobase upstream of the promoters they activate. However recruitment does not work if the activator protein is bound to an enhancer that is separated in space from the promoter because recruitment of the proteins to the enhancer does not increase their concentration at the promoter. Therefore, other strategies or mechanisms have to function for metazoan enhancers that are often separated by tens of kilobases up to megabases away from the promoter and several models have been proposed for enhancer action over a distance. These models all assume a function for enhancers in transcriptional initiation. However, analysis of mice containing a deletion of the major regulatory element, the LCR, which possesses strong enhancer activity, from the native β-globin locus, showed that PIC assembly and activator recruitment at the βmajor promoter is only mildly affected. Interestingly LCR deletion has a more dramatic effect on Serine 5 phosphorylation of the RNA polymerase II CTD and transcriptional elongation. This suggest that the β-globin LCR functions primarily downstream of activator recruitment and PIC assembly and enhances the transition from transcription initiation to elongation (Sawado et al. 2003) or more likely reinitiation. The exact mode of transcriptional enhancement does not influence the question of how long-range communication between enhancer and promoter or gene is established.

The Looping model

The looping model proposes a direct interaction of an enhancer bound activator protein with a protein at the promoter by bringing them in close proximity in the nuclear space while the intervening chromatin loops out and does not participate in the activation

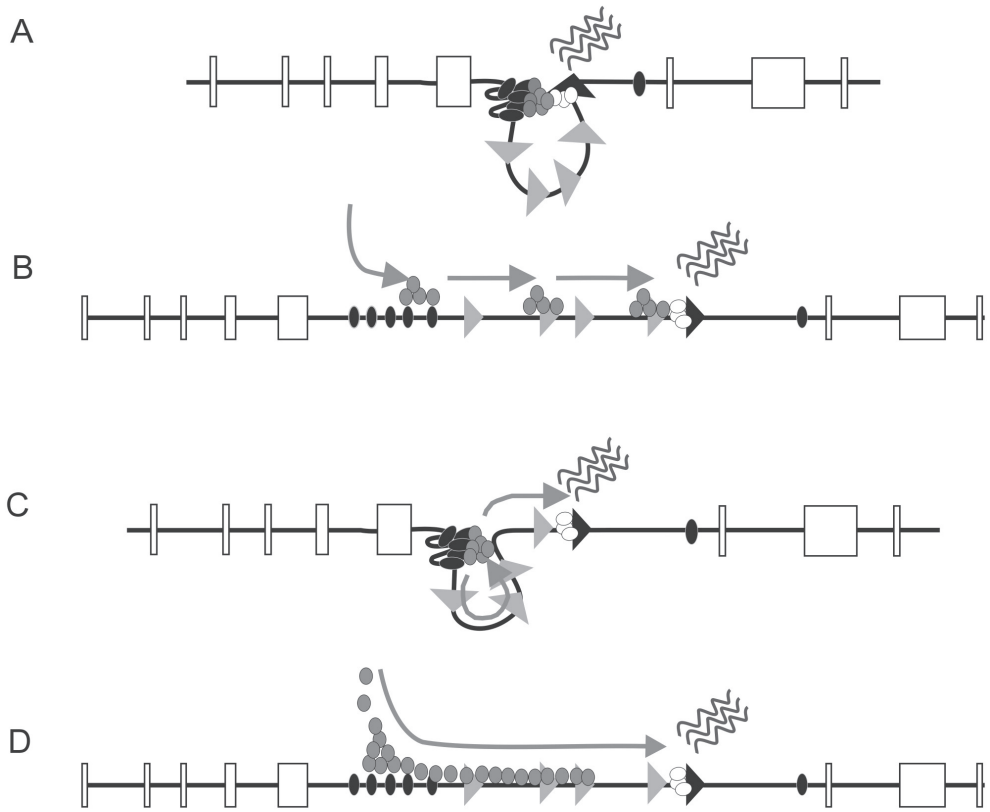


figure 1.2: Proposed mechanisms of long-range enhancer action.
In this figure long-range enhancer action is exemplified by LCR-globin gene promoter communication in the β-globin locus. The globin genes are depicted as triangles with the transcribed gene in black and the non-transcribed genes in grey. DNase I HSs are indicated as black ovals. LCR bound trans-acting factors are shown as grey ovals and promoter bound factors (e.g. general transcription machinery) as white ovals. White boxes are olfactory receptor (OR) genes.
(A) According to the looping model, DNA-bound protein-protein interactions result in direct contacts between an enhancer (LCR 'holocomplex') and the promoter of a gene. The contacts are established by random-collision and are essential for transcription. As a consequence of LCR-gene interactions the intervening chromatin template is looped out. Note that the looping model, unlike the other models, does not depend on the presence of the intervening chromatin template for transcription, rather transcription depends on random-collision between LCR-bound complexes and promoter-bound complexes. (B) The tracking model proposes that factors nucleate at the enhancer/LCR, and track along the intervening chromatin fibre to the promoter of the gene. The grey arrows indicate the movement of the complexes. Once a tracking complex reaches the promoter of a potentiated gene, transcription initiates. (C) The facilitated tracking model unites looping and tracking. Complexes nucleate and form an LCR-protein complex, which tracks along the intervening chromatin fibre towards a potentially active gene (grey arrows indicate movement). Upon direct contact between the LCR-protein complex and the promoter transcription initiates. (D) According to the linking model LCR-promoter communication is established by transmission of a nucleoprotein structure, which gradually polymerises along the intervening chromatin fibre in the direction of the promoter of a gene (grey arrow). Transcription is initiated when the nucleoprotein structure reaches the promoter of a potentiated gene.

process (figure 1.2A) (Ptashne 1986). This mechanism allows for an increase in the local concentration of the transcriptional machinery near the promoter via recruitment by an activator bound to a distally located enhancer.

DNA looping is a common way of communicating among distantly positioned DNA sequences in prokaryotes. For example *lac* and *gal* repressors can make stable DNA loops (reviewed in Matthews and Nichols 1998); similarly, interaction of the AraC activator with RNA polymerase is accompanied by loop formation (Lee and Schleif 1989). Activation of σ^{54}-dependent bacterial promoters by an enhancer occurs a looping in a process that shares several key properties with the action of eukaryotic enhancers i.e. activation is orientation independent, and can take place over a large distance both upstream and downstream of the activated promoter (Buck et al. 2000). Several proteins e.g. Bach1 and Sp1 are able to form *in vitro* looped structures in DNA molecules between their binding sites (Mastrangelo et al. 1991; Su et al. 1991; Yoshida et al. 1999). The SWI/SNF chromatin-remodelling complex is also able to form loops in an ATP independent fashion in *in vitro* reconstituted nucleosomal arrays (Bazett-Jones et al. 1999).

The looping model can readily explain a number of observations that have been made in eukaryotic systems. Transvection is a naturally occurring process in *Drosophila* whereby an enhancer on one chromosome activates a promoter on another chromosome (Morris et al. 1999; Wu and Morris 1999). This phenomenon, together with experiments that show that an enhancer on one DNA molecule can activate a promoter *in trans* on another DNA molecule (Dunaway and Droge 1989; Mueller-Storm et al. 1989; Mahmoudi et al. 2002) demonstrate that a *cis* configuration of a promoter and enhancer is not an absolute prerequisite for interaction, as predicted by the looping model. Gene competition for a single regulator (de Villiers et al. 1983; Wasylyk and Chambon 1983; Hanscombe et al. 1991), leading to alternate transcription by a stochastic "flip-flop" mechanism (Wijgerde et al. 1995; Gribnau et al. 1998; Trimborn et al. 1999) is most easily explained by direct interactions of an enhancer with the promoter, especially because the competitive advantage of the enhancer proximal gene is lost when genes are closely spaced at further distance from the regulator (Heuchel et al. 1989; Hanscombe et al. 1991; Dillon et al. 1997). Strong evidence in favour of the looping model of enhancer-promoter interactions was recently obtained for the mouse β-globin locus using two independent assays (Carter et al. 2002; Tolhuis et al. 2002). Both studies showed that the LCR of the mouse β-globin locus, the major regulatory element of this locus with enhancer properties, is in close proximity to the active β-globin promoter located 50 kb away on the linear template while the intervening chromatin loops out. Proximity of the LCR and active β-globin promoter was also observed in a transgenic human β-globin locus (Palstra et al. 2003).

The mechanism by which the enhancer finds the promoter, in that way establishing the chromatin loop, is unclear. Theoretical calculations (Rippe 2001) and experimental measurements of site-specific recombination between sites separated by several kilobases in mammalian cells (Ringrose et al. 1999) suggest that random diffusion alone is not enough to establish contacts between two sites separated by several kilobases on a chromatin fibre. Therefore mechanisms probably exist that facilitate enhancer-promoter contact over a distance. Many enhancers are transcribed in a RNA polymerase II dependent fashion and it has recently been suggested that transcription of these sites recruits them to transcription factories thereby inducing the looping (Cook 2003). Several DNA elements have been

suggested to facilitate the long range interaction between enhancer and promoter (Calzolari et al. 1999; Calhoun et al. 2002) and proteins like Chip and Nipped B might also be involved in facilitating the long range communication (Rollins et al. 1999; Rollins et al. 2004).

To find the promoter, an enhancer-bound activator can "scan" the surrounding DNA regions by "hopping" and "scanning" mechanisms as proposed for translocation of the *lac* repressor on DNA (Berg et al. 1981; Winter et al. 1981). Such a "looping-scanning" mechanism has also been proposed for eukaryotic enhancers in the facilitated tracking model (figure 1.2C) (Blackwood and Kadonaga 1998). Some evidence for the facilitated tracking model has been obtained from the study of the HNF4α enhancer (Hatzis and Talianidis 2002). In this study a ChIP approach was used to show that enhancer bound proteins (C/EBPα and HNF-3β) could be cross-linked to the 6.5 kb of spacer DNA separating the enhancer and promoter as well as to the enhancer itself. This cross-linking to spacer DNA can only be detected after activation of the enhancer but before actual transcription of the gene at which stage the enhancer bound proteins are cross-linked to the promoter. It has to be mentioned that the data shown are extremely clean and background signals, expected with a technique like ChIP, are never detected, raising concerns about the linearity of amplification or the application of thresholds. Moreover, the enhancer takes about 80 hours to track along the 6.5 kb of spacer DNA and therefore it seems unlikely that this process takes place in larger loci or when fast transcriptional activation is required. Moreover, it was shown that several HS spread throughout the β-globin cluster together in the nuclear space and is very hard to envision which HS is tracking where (Tolhuis et al. 2002). The facilitated tracking model shares some of the limitations with the tracking model.

The tracking model

In the tracking model the enhancer acts as a loading platform for a DNA-tracking protein. Since the tracking protein is leaving the enhancer, the protein may be loaded again and may in this way accumulate in the vicinity of the promoter (figure 1.2B). Tracking has been shown to be the mechanism in the enhancer action of the late genes of bacteriophage T4 (for a review see Kolesky et al. 2002). However, there is no conclusive evidence for tracking in eukaryotes and no activators are known that have to leave the enhancer to activate transcription. The tracking model is attractive because it can explain enhancer blockers in a satisfying way (in a similar way as the linking model can; see futher) by suggesting that the tracking protein is stopped in its tracks by an insulator bound protein. The detection of intergenic transcripts that originate from enhancers (Tuan et al. 1992; Kong et al. 1997; Gribnau et al. 2000) has renewed interest in this model and it has been suggested that RNA polymerase II might be the tracking protein. However, Muller and colleagues showed that introducing a transcriptional terminator between the promoter and enhancer does not terminate the enhancer action (Muller et al. 1990). Furthermore, the tracking model is difficult to reconcile with observations that an enhancer can activate a promoter *in trans* (Dunaway and Droge 1989; Mueller-Storm et al. 1989; Mahmoudi et al. 2002) as well as the observation of alternate transcription by a "flip-flop" mechanism (Wijgerde et al. 1995; Gribnau et al. 1998; Trimborn et al. 1999). Additional mechanisms have to be incorporated into the tracking model to account for these observations.

The linking model

As in the tracking model, the linking model proposes that the enhancer acts as a loading platform for DNA-binding proteins. However, now the bound protein(s) facilitates polymerisation of the proteins in the direction of the promoter thereby coating the chromatin fibre (figure 1.2D) (Bulger and Groudine 1999; Dorsett 1999). The linking model was proposed to explain the properties of the *Drosophila* Chip protein. Chip cannot bind to DNA directly, but can interact with numerous transcription factors and facilitates their action over a distance *in vivo* (Morcillo et al. 1997; Torigoi et al. 2000). It was proposed that Chip is recruited by an activator bound at an enhancer where it functions as a protein "bridge" between the activator bound at the enhancer and other proteins having multiple weak binding sites between enhancer and promoter (Dorsett 1999; Gause et al. 2001). If distances between enhancer and promoter are large a considerable amount of protein is needed to coat the chromatin fibre therefore it has been proposed that the spreading induces formation of small loops (~1.5 kb) between multiple small activator binding sites.

Other models

Based on observations of the activation of a transgene and its nuclear localisation it was proposed that activation by an enhancer involves translocation of a locus to a different nuclear compartment that is more favourable to transcription (Francastel et al. 1999). Recent observations of histone acetylation and nuclear localisation of the β-globin locus showed that translocation of the locus away from centromeric heterochromatin does not require the LCR (Schubeler et al. 2000), while translocation away from its chromosomal territory and its final activation is LCR dependent (Ragoczy et al. 2003). However, these studies were performed on human chromosomes containing the wild type, Hispanic thalassemia, or ΔHS2-5 human β-globin loci after their transfer from a nonerythroid into a mouse erythroid background and these effects are not observed in normal transgenic constructs. Therefore these effects may not be true for loci in general.

Another intriguing model for enhancer action involves chromatin domain opening by the reduction of linkage numbers in supercoiled looped chromatin domains. The topological changes can bypass complexes formed on DNA strands and can be transmitted from localised points over a broad region (Xin et al. 2003). Most models are not mutually exclusive and combinations of the above models, e.g. a combination of linking and looping, have been proposed (Bulger and Groudine 2002).

Insulators: boundary elements and enhancer blockers

Insulators are elements that demarcate domains of autonomous regulated gene expression. Two different effects have to be distinguished; a boundary activity that counteracts the spreading of heterochromatin and an enhancer blocking activity that blocks the activation of a promoter by an enhancer. Importantly, enhancer blockers only work when placed between

the promoter and enhancer. Both activities can be mediated by one kind of element as is exemplified by the *Drosophila* gypsy insulator, which consists of twelve degenerate binding sites for the DNA-binding protein Suppressor of Hairy-wing (Su(Hw)) (Parkhurst et al. 1988; Spana et al. 1988; Scott et al. 1999). In the chicken β-globin HS4 element these activities are separable. A CTCF binding site mediates the enhancer blocking activity and this site is not necessary or sufficient for counteracting heterochromatin spreading. Four other footprints, for which the binding partners have not been identified, mediate the boundary activity of cHS4 (Recillas-Targa et al. 2002). A wide variety of enhancer-blocking elements have been identified. Most of them are found in *Drosophila*, but increasing numbers are identified in vertebrates. Several DNA-binding proteins, without any similarities, that interact with these elements have been identified. Three of these, Su(Hw) (Parkhurst et al. 1988), BEAF (Zhao et al. 1995) and Zw5 (Gaszner et al. 1999) were found in *Drosophila*. CTCF, an eleven zinc-finger protein, provided the first example of a vertebrate enhancer-blocking protein (Bell et al. 1999). Recently the transcription factor YY1 was also shown to possess enhancer-blocking activity (Kim et al. 2003).

Models for enhancer-blocking activity are intimately linked to models explaining enhancer function at a distance and the property most difficult to explain is the position dependence of enhancer blockers. In processive models, like tracking, linking and facilitated tracking, the presence of an enhancer blocker bound by a protein complex will block the signal/activity that progresses from the enhancer to the promoter. However, duplication of a Su(Hw) insulator between an enhancer and promoter neutralised the enhancer-blocking activity instead of enhancing it, as predicted by the processive models (Cai and Shen 2001; Muravyova et al. 2001). The "decoy model" suggests that the enhancer blocker traps the enhancer rendering it inactive (Geyer 1997). Consistent with this model are studies in *Drosophila*, which have shown that enhancer blocking function depends strongly on enhancer and promoter strength (Scott et al. 1999; Cai et al. 2001). The position dependence of enhancer blockers is difficult to reconcile with this model, a decoy element placed on the other side of the enhancer should work equally well.

Recent data indicate a more structural or topological role for enhancer blockers. The *Drosophila* genome contains about 500 binding sites for Su(Hw) which are located at the boundaries between bands and interbands of polytene chromosomes. The 500 insulator sites coalesce into approximately 25 large structures, named insulator bodies, present mostly in the nuclear periphery in diploid cells (Gerasimova et al. 2000). The observation that duplication of a Su(Hw) insulator between an enhancer and promoter neutralised the enhancer-blocking activity, suggests that the enhancer blockers interact with each other and in this way prevent the targeting to an insulator body. (Cai and Shen 2001; Muravyova et al. 2001). Indeed, several of the proteins that bind insulator elements (or their binding partners) are able to form homo- or hetero-dimers and even higher oligomers. Blanton and co-workers obtained direct evidence of interaction between insulator elements. They showed in a ChIP assay, that Zw5 a protein that only binds the *scs* element also crosslinks to the *scs'*-element the binding site for BEAF. Zw5 and BEAF are able to interact with each other *in vitro* and *in vivo*. Furthermore they confirmed that the *scs* element and the *scs'*-element, are in close proximity to each other *in vivo* (Blanton et al. 2003). *Gypsy* insulator elements are also found at the base of chromatin loops in nuclear halos and this depends on the presence of Su(Hw) (Byrd and Corces 2003). Furthermore, artificial tethering of chromosome loops to nuclear pores in yeast

generates isolated domains (Ishii et al, 2002) and formation of stable LacI –induced DNA loops, topologically isolating an enhancer and promoter in a plasmid, resulted in insulator activity (Bondarenko et al. 2003a; Bondarenko et al. 2003b). Thus insulators can organise the chromatin fibre into separate loops forming rosette like structures. It has been proposed that an enhancer in one loop would be unable to communicate with a promoter in another loop. However, clustering of insulators could also serve to target genes to a compartment that is less favourable for transcription, like the nuclear periphery.

The activity of enhancer blockers can be regulated, as exemplified by the regulation of the *Igf2/H19* locus in mammals. *Igf2* is only expressed from the paternal chromosome and is activated by an enhancer located distally of *H19*, a gene which is only expressed from the maternal chromosome. An enhancer blocker called the ICR, containing CTCF binding sites, is located between *Igf2* and *H19*. On the paternal chromosome the imprinting control region (ICR) is methylated, preventing the binding of CTCF and thus allowing activation of *Igf2* while the *H19* promoter is silenced by methylation. Conversely, on the maternal chromosome the ICR is not methylated and CTCF binds thereby blocking activation of *Igf2* by the enhancer which now activates *H19* (Bell and Felsenfeld 2000; Hark et al. 2000). It is interesting to note that the *Igf2/H19* enhancer blocker works through positional looping (mediated by the differential interaction between diferentially methylated regions (DMRs)) of the *Igf2/H19* locus. The chromatin loop acts as a simple epigenetic switch by which *Igf2* (whose promoters are not regulated by DNA methylation) is moved into either an inactive domain or an active domain close to enhancers, depending on the methylation status of the DMRs and CTCF binding (Murrell et al. 2004). Interestingly, 5'HS5 of the human β-globin LCR was recently identified to be a developmental stage-specific enhancer blocker (Wai et al. 2003). The enhancer blockers can also be regulated by post-translational modifications of the bound proteins; Poly(ADP-ribosyl)ation of CTCF is essential for its enhancer blocking activity (Yu et al. 2004).

Several mechanisms have been suggested to be involved in counteracting the spreading of heterochromatin. These include targeting of histone acetylases to boundary elements as in cHS4 (Litt et al. 2001), generation of nucleosome free regions (Bi et al. 2004), and possibly formation of looped structures (Laloraya et al. 2000; Ishii et al. 2002; Ishii and Laemmli 2003).

Locus Control Regions

Locus control regions (LCRs) are regulatory elements that define active domains of gene expression in a dominant fashion. LCRs are operationally defined by their ability to enhance expression of linked genes to physiological levels in a tissue-specific and copy number-dependent fashion regardless of the ectopic chromatin site (Grosveld et al. 1987). Although most data regarding LCR function has come from studies on the human and mouse β-globin LCR, several other LCR containing loci have been identified (Li et al. 2002). LCRs are composed of varying numbers of tissue-specific DNaseI-hypersensitive sites and typically include enhancers and insulators/enhancer blockers. LCRs can take up a well-defined region in a locus, like the β-globin LCR (Grosveld et al. 1987) or the elements can be spread throughout a locus as in the chicken lysozyme, λ5, CD2 and hGH loci (Li et al. 2002).

Studies on the human β-globin and human CD2 LCR showed that that these LCRs are essential for establishing an open chromatin structure (but see e.g. Epner et al. 1998; Schubeler et al. 2001). Histone acetylation appears to be a common theme in LCR action; elements from the LCRs of the murine immunoglobulin heavy chain locus, human growth hormone locus and T-cell receptor γ locus all appear to induce a wide spread increase of histone acetylation (Li et al. 2002). However, targeted deletion of the mouse LCR shows that, even after germline passage of the deletion, the locus is still DNase I sensitive and acetylation of nucleosomes only modestly reduced (Bender et al. 2000a; Schubeler et al. 2001). This shows that in the endogenous mouse β-globin locus, the LCR is dispensable to initiate an open chromatin conformation (see chapter 3).

The T cell-specific TCR α/δ LCR consists of 8 HSs located downstream of the T-cell receptor (TCR) gene and is a bi-functional element regulating both the TCR gene and the adjacent ubiquitously expressed *Dad1* anti-apoptosis gene. Two subregions in the TCRα/δ LCR have been identified: one that constitutes a non-tissue restricted chromatin opening element and one comprising the 4 proximal HSs that restore tissue specificity to the downstream chromatin-opening element (Ortiz et al. 1999). The occurrence of activators and insulators in LCRs appears to be a common theme, suggesting that the interaction of these elements may modulate LCR function. Indeed, the tissue unrestricted HS was bound by several ubiquitously expressed transcription factors when it was linked to a transgene. However, when the complete LCR was present, tissue-specific binding of tissue-restricted proteins to this element was observed (Ortiz et al. 2001). Interestingly, the TCR α/δ LCR is also implicated in keeping specific DNA regions methylation free in a tissue specific manner, thus opposing chromatin closure and gene silencing (Santoso et al. 2000).

An interesting variant of the LCR theme can be found in the human keratin-18 (K18) gene locus. The 2.5- and 3.5-kb regions, which flank the human K18 gene conferred position-independent expression on K18 and two heterologous reporter genes in transgenic mice. The 5'-flanking region contains a 323-bp fragment composed primarily of a transcriptionally active Alu repetitive element. Two copies of the 323-bp Alu fragment, flanking a heterologous transgene, retained much of the position effect blocking activity contained in the larger 5'- and 3'-flanking regions of K18. Its activity in mice is dependent on its RNA polymerase III promoter activity and it was suggested that transcriptionally active Alu elements might eliminate transcriptional interference of neighboring genes (Willoughby et al. 2000).

The diversity and complexity of the LCRs currently identified highlights the importance to consider *in vivo* transcriptional regulation in the context of complete loci, so that essential regulatory elements are not excluded or overlooked.

Chapter 2

The dynamic nucleus and its structural organisation

The dynamic nucleus and its structural organisation

The nucleus contains and protects the genetic material of an eukaryotic cell and is the site of several key processes in the cell, including the on and off switching in time and space of many thousands of genes, the faithful replication of the genome during cell division and repair of damaged DNA. Transcription, RNA processing, DNA repair and replication are functionally tightly coupled and it is clear that these complex processes can only be carried out in a nucleus that is structurally highly organised. Indeed many structures like the nucleolus, the nuclear membrane, heterochromatic and euchromatic compartments can easily be visualised in microscopy studies (figure 2.1).

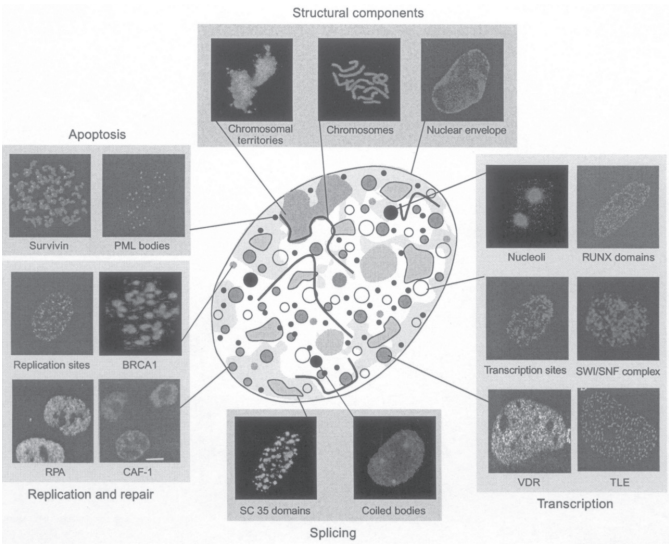


figure 2.1: The nuclear architecture is functionally linked to the organization and sorting of regulatory information.
Immunofluorescence microscopy of the nucleus in situ has revealed the distinct nonoverlapping subnuclear distribution of vital nuclear processes, including: DNA replication sites, DNA damage repair, apoptosis, structural parameters of the nucleus (e.g. the nuclear envelope, chromosomes and chromosomal territories), rRNA synthesis (nucleoli), chromatin remodeling (e.g. mediated by the SWI/SNF complex), transcriptional control and RNA processing and splicing.(Stein et al. 2003)

Chromosome territories

When the cell exits mitosis, the rod-like chromosomes de-condense into their interphase state. This de-condensation is not random and chromosome painting shows that chromosomes occupy relative discrete, non-overlapping, territories (Cremer and Cremer 2001). The discovery that small nuclear proteins (snRNPs) and RNA transcripts were located at the boundary of these chromosome territories, led to the proposal that the transcription and splicing machinery might be confined to a system of channels that run between the chromosome territories and are connected to the nuclear pores. This interchromosome domain

(ICD) model predicted that active genes reside at the exterior of the condensed chromatin domains, which line the interchromatin channels and that are accessible to the transcription and mRNA processing machinery (Zirbel et al. 1993). Support for this model came from the observations that active genes, but not silenced genes or noncoding regions, from several loci reside at the surface of their respective chromosome territories (Kurz et al. 1996; Dietzel et al. 1999). However, the subsequent discovery of transcription sites within the interior of chromosome territories (Abranches et al. 1998; Verschure et al. 1999) necessitated a refinement of the ICD model. Successive high-resolution microscopy studies showed that although chromosomes occupy relatively discrete areas in the nucleus, these domains are highly porous with numerous channels extending throughout more condensed sub domains (Visser et al. 2000). The revised interchromatin compartment (IC) model envisages that active genes not only reside at the surface of the territory, but also within the territory at the surface of channels that extend throughout condensed chromatin sub-domains and are accessible to the transcription and mRNA processing machinery (Cremer and Cremer 2001). Transcriptionally inactive regions would tend to be at the interior of the condensed chromatin sub-domains. Initially immuno-fluorescence microscopy studies indicated that TFIIH and RNA polymerase II (the transcription machinery) and hnRNP-U (splicing factor) are present predominantly in the interchromatin space, inside and between chromosome territories, and are largely excluded by domains of condensed chromatin (Verschure et al. 2002). However, a subsequent live cell study from the same group gave conflicting results and showed that condensed chromatin in living cells is accessible to fluorescently labelled dextrans and proteins (like RNA polymerase II) of various sizes (Verschure et al. 2003). Therefore, the silencing of genes that are incorporated into such domains is not due to the physical inaccessibility of condensed chromatin domains to transcription factors. Recent evidence for the IC domain model has come from a study of the WAGR locus located at chromosome 11p13 and containing the genes *WT1, RCN, PAX6* and *PAXNEB*. The locus is located within the painted chromosome 11 territory and transcriptional activation of the genes did not result in their relocation to the territory surface. However, the active *RCN* gene was found to reside on the surface of a condensed chromatin subdomain, whereas an intergenic probe from the WAGR locus located within a condensed subdomain (Mahy et al. 2002b).

Some active genes can be found on large extended loops outside their painted chromosome territories. The major histocompatibility (MHC) locus located at 6p21 is able to extend out from the chromosome 6 territory on large loops in response to activation of the MHC genes in B-lymphoblastoid cells and interferon-γ treated fibroblasts (Volpi et al. 2000). Similar observations are made for the epidermal differentiation complex (EDC) locus (Williams et al. 2002), the β-globin locus (Ragoczy et al. 2003) and the *HoxB* locus (Chambeyron and Bickmore 2004a). The fact that these loci are gene dense and have a high transcriptional activity, taken together with a study of genomic regions with known gene densities (Mahy et al. 2002a), suggests that the formation of extraterritorial loops is correlated to gene density and transcriptional activity in general. However, inhibition of transcription with Actinomycin D or 5,6-dichloro-1β-D ribofuranosyl benzimidazole (DRB) caused only a minor reduction in the observed extraterritorial signals (Mahy et al. 2002a). In contrast, the MHC locus has an approximately twofold reduced number of extraterritorial signals in cell types where the genes are not expressed compared to cells treated with IFN-γ (Volpi et al. 2000). Therefore, it remains to be established whether the looping out of chromosome territories is a requirement for transcription at these particular loci.

Dynamics of chromatin and chromatin interacting proteins

The use of FISH and immunohistochemistry techniques has the limitation that these techniques only provide snapshots of a nucleus and might give the impression that the organisation of the nucleus is rather static. However, analysis of large numbers of nuclei show that the positions of gene loci with respect to their territory are not absolute but show a statistically significant trend towards a position, suggesting that the loci are dynamic.

Utilisation of green fluorescent protein (GFP), fused to a protein of interest, has led to an upsurge of information regarding the localisation and dynamics of proteins within living cells. By using this approach in combination with photobleaching techniques it is even possible to determine how much of the protein is mobile and the frequency of exchange of mobile proteins between cellular (or nuclear) compartments (Lippincott-Schwartz et al. 2001). Genomic loci can be tagged by *lac* or *tet* operator sequences allowing them to be visualised and followed in time by detecting the bound repressor-GFP fusion proteins (Belmont 2001; Spector 2003).

Experiments in yeast, *Drosophila* and mammalian cells, using lac-repressor protein fused to GFP to visualise randomly integrated *lac* operator repeat arrays in live cells show that chromatin at interphase is continually moving (Marshall et al. 1997; Vazquez et al. 2001; Chubb et al. 2002). However, the chromatin does not roam freely around the nucleus but is constraint to an average range of movement of 0.5 μm, which differs per chromatin region (Chubb et al. 2002). In higher eukaryotes, loci are in this way on average restricted to about $1/1000^{th}$ of the nuclear volume limiting the range of environments that it is able to access. It has been suggested that the sub-nuclear position of a locus contributes to the regulation of gene expression. Movements over several microns have been observed in *Drosophila* (Vazquez et al. 2001) and it remains possible that large movements can occur in specific situations allowing them to access specialised regions of the nucleus such as transcriptionally repressive compartments of a chromosome territory (Ragoczy et al. 2003). Targeting of the acidic activator VP16 to a tagged chromosome region leads to large-scale uncoiling of the chromatin structure and an additional change in the position of a chromosome site from the periphery to the interior of the nucleus (Tumbar et al. 1999; Tumbar and Belmont 2001). Large-scale chromatin decondensation of a chromosome array was also observed upon transcriptional activation of a GR-responsive promoter (Muller et al. 2001). Experiments using an inducible system, composed of a 200-copy transgene array integrated into a euchromatic region, show that the condensed array is heterochromatic. In the condensed state it is associated with HP1, histone H3 methylated at lysine 9, and several histone methyltransferases. Upon transcriptional induction, the locus decondenses within 20 minutes correlating to an increase in RNA levels at the transcription site. HP1alpha is depleted from the locus and the histone variant H3.3 is deposited suggesting that histone exchange is a mechanism through which heterochromatin is transformed into a transcriptionally active state (Janicki et al. 2004).

Biochemical studies have led to the notion that transcription factors and associated factors are in stable complexes, which are bound to chromatin for long periods of time. The observation of GFP tagged proteins in living cells has changed this view and it has become apparent that many, if not most, nuclear proteins are moving quickly, and exchange rapidly with a variety of nuclear targets. The presence of very large protein complexes like RNA

polymerase holoenzymes, roaming the interchromatin compartment is questioned by studies on the assembly dynamics of the RNA polymerase I transcription machinery (Dundr and Misteli 2001). This study suggests that assembly of a productive polymerase holocomplex occurs by stochastic binding of individual subunits at the site of transcription. Similarly, the nuclear dynamics of RNA polymerase II and transcriptional regulators is consistent with such a model (Becker et al. 2002; Kimura et al. 2002). These observations were in line with previous investigations of other transcription factors. Transcription factors like the hormone bound glucocorticoid receptor and the cofactor GRIP-1 exchange rapidly between the nucleoplasm and multi copy repeats of reporter elements (McNally et al. 2000; Becker et al. 2002). Furthermore the transcription initiation and DNA repair factor TFIIH is also transiently associated with RNA polymerase II complexes (Hoogstraten et al. 2002). Even components of "stable" heterochromatin like histone H1 (Misteli et al. 2000) and HP1 (Cheutin et al. 2003) are surprisingly mobile when compared with the core histones (Kimura and Cook 2001). A recent extensive study of the basic *in vivo* biophysical properties of a wide range of chromatin proteins of diverse functions demonstrated that most chromatin associated proteins have a residence time in the order of seconds and that transient binding is a common property of chromatin-associated proteins. Regardless, at steady state the major population of each protein is bound to chromatin rather than present in a soluble form in the nucleoplasm (Phair et al. 2004). This data support a model in which a single molecule of a chromatin-associated protein resides on chromatin for a few seconds and then dissociates and diffuses for a relatively short period of time before it associates with a new site. This suggests that chromatin-binding proteins find their binding sites largely by three-dimensional scanning of the genome space and that dynamic interaction networks play a critical role in the control of gene expression (Misteli 2001b).

Because of the essentially stochastic nature of the binding of individual proteins, the assembly of functional complexes is an inefficient, possibly rate limiting process. The overall flux of the reaction may therefore be extremely sensitive to any factors that affect protein binding and dissociation, such as histone modifications, nucleosome positioning and the local concentration of both protein and target. In this situation nuclear compartmentalisation might have large effects on the efficiency of nuclear processes (Chubb and Bickmore 2003).

Nuclear compartments

Several nuclear proteins, like RUNX (Zeng et al. 1998), accumulate at discrete sites in the nucleus and this accumulation appears to be linked to the function of the protein. However, the compartmentalisation observed in the nucleus is very different from compartmentalisation in the cytoplasm where factors are concentrated behind membranes. The structural organisation of nuclear compartments represents a dynamic steady state rather than a static situation and the compartments themselves are established by particular nuclear functions (Misteli 2001a). A compartment can be defined as a quantity of a substance that has uniform and distinguishable kinetics of transportation or transport (box).

Many nuclear compartments can be identified in the nucleoplasmic space and include nuclear speckles containing spliceosomal components (Misteli et al. 1997), Cajal bodies involved in snRNP biogenesis (Gall 2000) and PML nuclear bodies enriched in transcriptional regulators (Zong et al. 2000). The exact function of these nuclear bodies is unclear but they

Box: compartments

From a mathematical point of view, a compartment is defined as a quantity of a substance that has uniform and distinguishable kinetics of transport (Godfrey, 1983). The interior of a nucleus is clearly non-homogeneous and therefore compartmentalised. The nuclear compartments are not separated from the nucleoplasm by a physical barrier and macromolecules inside the nucleus of a living cell are in continuous flux between compartments. For most proteins the exchange is rapid, and the residence time of most proteins is in the order of a minute or less. Compartments are thus extremely dynamic, yet overall stable, structures, and their morphology represents the equilibrium of release and binding of proteins. The functional status of exchanging proteins critically determines composition and morphological appearance of a compartment suggesting that compartments are formed and maintained by principles of self-organisation. Self-organisation involves the physical interactions of molecules in a dynamic steady-state structure and the properties that determine the organisation are the intrinsic properties of the structure's components. Self-organised structures do not require complex mechanisms to establish, maintain, and regulate their architecture and thus, self-organisation is a simple but effective way to optimally organise nuclear structures. (Misteli 2001a; Misteli 2001b; Carmo-Fonseca 2002)

are all in constant flux and highly dynamic (Belmont 2003; Janicki and Spector 2003). The best-characterised nuclear compartment is the nucleolus, which is dedicated to the RNA polymerase I dependent transcription of rDNA genes and ribosome biogenesis. Within each nucleolus, three distinct compartments can be further identified: the fibrillar centre, which contains the rDNA genes, the dense fibrillar component, where maturation of pre-rRNA transcripts occurs, and the granular component, dedicated to assembly of pre-ribosomal particles (Carmo-Fonseca et al. 2000). The integrity of nucleoli is dependent on ongoing RNA polymerase I transcription (Carmo-Fonseca et al. 2000).

Transcription factories

Activation of a gene requires remodelling of chromatin and subsequent assembly of the transcription machinery. Chromatin modification and assembly of the core transcriptional apparatus is based on stochastic processes and is therefore relatively inefficient. The relative low efficiency of transcription apparatus assembly is counterbalanced by the co-operative nature of binding of many transcription factors and by the formation of transcription factories, which act to increase the local concentration of components (Misteli 2001b).

The first notion of transcription factories came from estimates of the amount of nascent RNA transcripts in HeLa cells and the actual number of sites with ongoing transcription. The number of nascent RNA transcripts per sub-tetraploid HeLa nucleus was determined to be approximately 90,000 of which 65,000 were contributed by RNA polymerase II (Jackson et al. 1998). Several estimates of polymerase II sites in the nucleoplasm give a number of 5000 to 8000 active sites per HeLa nucleus (Iborra and Cook 1998; Jackson et al. 1998; Pombo et al. 1999). Furthermore, most of the active transcription units, either *bona fide* gene or non-coding, are associated with one RNA polymerase II molecule (Jackson et al. 1998). These results imply that every transcription site contains approximately eight polymerases along with about eight transcription units. Therefore clustering of RNA polymerase II transcribed genes in transcription factories appears to be very similar to the clustering of RNA polymerase I dependent rDNA genes in the nucleolus. Recently it was show that in erythrocytes, during transcription *in vivo*, distal active genes co-localise at high frequencies

to the same transcription factory containing the β-globin gene. The authors propose that active genes migrate to pre-assembled transcription sites rather than recruiting and assembling transcription complexes and the movement into or out of these factories results in activation or abatement of transcription (Osborne et al. 2004).

Historically it has been thought that the polymerase moves along the template as it transcribes. However, the sheer size of the transcription complex argues against a polymerase tracking along the tangled DNA template and at the same time preventing the transcript from becoming entwined (Cook 1999). More recent models suggest that the polymerase remains immobilised while the DNA template is pulled through the active site. The RNA polymerase is perfectly capable of reeling in the DNA template and is the most efficient molecular motor known (Gelles and Landick 1998). An immobilised RNA polymerase also fits well with the data on transcription factories, where active transcripts are restricted to an area with a diameter in the range of 40-70 nm (Jackson 2003). How the polymerases are attached in the factories is unknown but protein : protein, protein : DNA and protein : RNA interaction might be enough to hold the factory together while transcription is underway. Disruption of the presumed nuclear network formed from nuclear lamins results in a reduction of RNA polymerase transcription (Spann et al. 2002). This suggests that transcription factories might be tethered to a nuclear substructure. Experiments show that the majority of a GFP tagged RNA polymerse II is highly mobile when they are not engaged in transcription (Kimura et al. 2002) and are not fixed to any substructure. Therefore if transcription factories exist they must be highly dynamic and are only assembled, and possibly linked to a nuclear substructure, upon transcriptional activation.

Nuclear Matrix and Chromatin loops

Nuclei appear to have a complex structure of nuclear compartments, which generally represent centres of nuclear function. In general these centres are not stable entities and contain components that are inherently dynamic. Remarkably, the spatial organisation of these sites is unaffected if nuclei are depleted of chromatin (Nickerson 2001). This observation suggests that a nuclear solid phase, the nuclear matrix, exists upon which these active sites assemble. It is generally assumed that this structure that can be seen in the microscope after extraction of chromatin is the same as the amorphous structure that can be isolated from nuclei by hypertonic extraction following nuclease digestion (Hancock 2000; Nickerson 2001) or the nucleo-protoneous core that is found when nuclear halo's are prepared.

Several proteins implicated in the regulation of gene expression, like CTCF (Yusufzai and Felsenfeld 2004), SatB1 (Dickinson et al. 1992; Yasui et al. 2002) and Su(Hw) (Byrd and Corces 2003) reportedly interact with the nuclear matrix. DNA sequences called MAR/SARs that bind to the nuclear matrix have been identified and these are able to modulate gene expression (Bode et al. 2000). Insulators like *gypsy* (Byrd and Corces 2003) and cHS4 (Yusufzai and Felsenfeld 2004) are also often found to interact with the nuclear matrix. Transcription dependent association of genes with the nuclear matrix has also been reported (Maya-Mendoza et al. 2003) and transcription processes are dependent on proteins that could function as the structural building block for a matrix like lamin (Spann et al. 2002), nuclear actin or actin related proteins (Blessing et al. 2004). Recently it was shown that emrin, an inner nuclear membrane protein, is able to cap the pointed ends of actin filaments providing

evidence for and actin cortical network at the nuclear inner membrane (Holaska et al. 2004). However, structural polymers of the proteins supposed to form a nuclear matrix structure have not been characterised or visualised in live cells. Therefore concerns have been raised that the nuclear matrix is an artefact formed by a-specific interactions induced during the severe extraction procedures (Hancock 2000; Pederson 2000). The concept of a nuclear matrix remains therefore enigmatic and highly controversial.

Chapter 3

Temporal and spatial regulation of transcription: The β-globin locus as a model system

Temporal and spatial regulation of transcription: the β-globin locus as a model system

Erythropoiesis

After fertilisation, when an egg cell and sperm cell fuse, a single cell embryo is formed, which is omnipotent. This single cell embryo has the potential to generate all different tissues needed to form a complete organism. While proliferating, this capacity is maintained in embryonic stem cells, but upon differentiation, which happens during early embryogenesis, this omnipotency is gradually lost and so called multipotent stem cells are formed. These self-renewing cells only have the capacity to contribute to limited lineages of tissue, and it is believed that these cells are maintained through adult life.

The hematopoietic stem cell (HSC) is responsible for all blood formation and resides in the adult bone marrow. Under the influence of hematopoietic growth factors these cells either proliferate to expand the HSC compartment or differentiate, via several more

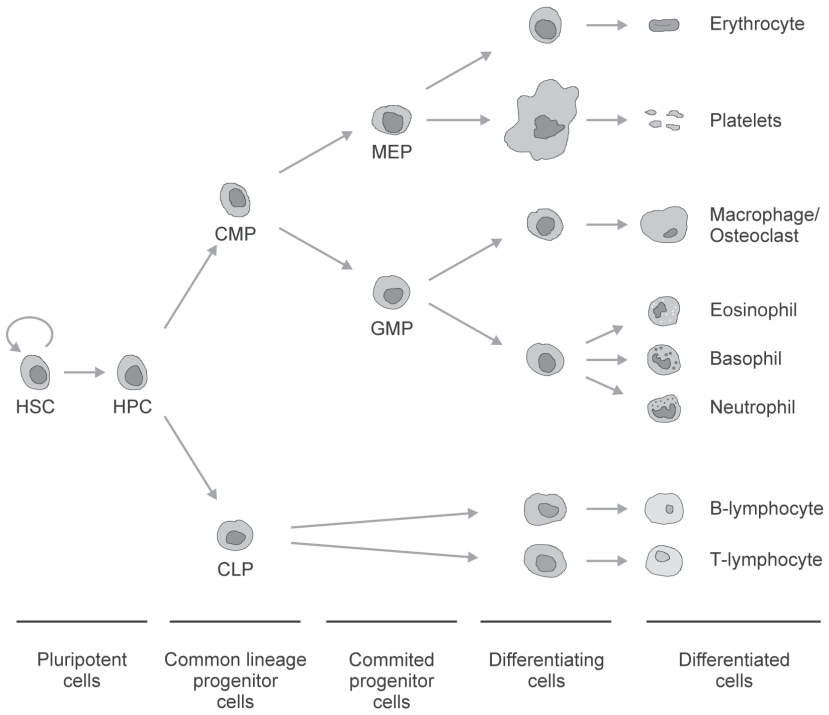


figure 3.1: Schematic representation of the hematopoietic hierarchy.
The hematopoietic stem cell gives rice to all blood cells. During differentiation, cells loose their pluripotent capacity and become progressively more restricted to one of the blood cell lineages. HSC= hematopoietic stem cell; HPC= hematopoietic progenitor cell; CMP= common myeloid progenitor; CLP= common lymphoid progenitor; MEP= myeloid and erythroid progenitor; GMP= granulocyte monocyte progenitor.

committed progenitor cells, into the different blood lineages (figure 3.1). This differentiation program is controlled by the expression of a combinatorial set of broad spectrum and cell lineage specific transcription factors (Perry and Soreq 2002). Modulation of transcription factor expression patterns drives the differentiation into different lineages (Hirasawa et al. 2002; McNagny and Graf 2002).

The most abundant blood cell in the mammalian system is the red blood cell or erythrocyte. Erythroid differentiation takes place through a series of intermediate precursors that progressively gain erythroid features and gradually lose proliferative capacity. The CFU-E (colony forming unit erythroid) is the last intermediate that still has considerable proliferative capacity and is characterised by its capacity to form small colonies in semisolid medium (Wong et al. 1986). A well co-ordinated cohort of transcription factors regulates the formation, survival, proliferation and differentiation of multipotent progenitors into the erythroid lineage (Perry and Soreq 2002). Expression of a hierarchical program of genes, most notably α-globin and β-globin of which the protein products combine to form the oxygen carrying molecule haemoglobin, accompany differentiation. During the terminal stages of differentiation the erythrocyte sheds its nucleus, endoplasmatic reticulum and mitochondria, leaving a cell membrane, filled with approximately 90% haemoglobin, specialised to transport oxygen and carbondioxyde. Since this cell is unable to grow and proliferate and in humans dies after spending approximately 120 days in the circulation, red cells have to be continuously replenished from a pool of HSC. In embryos erythropoesis is initiated in the yolk sac shifting to the liver during fetal development and finally resides in the adult bone marrow. The erythroid compartment is a well-characterised and accessible system, making it a popular model for studying the mechanisms involved in differentiation and development.

The β-globin locus

The α-globin and β-globin genes were among the first to be studied by biochemical and molecular biological methods due to the easy availability of globin protein and globin mRNA from erythroid cells. For example haemoglobin (horse methaemoglobin) was one of the first proteins to be crystallised (Rupley 1964). Furthermore, the first eukaryote introns were discovered in the β-globin gene (Jeffreys and Flavell 1977; Tilghman et al. 1978) and thalassemia and sickle cell anaemia were the first genetic disorders discovered to be caused by mutations in, or deletions of, a specific gene (the globin genes) (Weatherall 2004). More recently the globin genes were used as an important model system to study tissue-specific and developmentally regulated transcription. Study of the globin genes has played a major role in identifying *cis*-regulatory DNA elements, which contribute to the proper regulation of transcription.

The mouse and human β-globin locus

The mammalian β-globin loci and especially the mouse and human β-globin locus are highly conserved (figure 3.2).The mouse β-globin locus contains four functional genes

arranged in the order of their temporal expression starting 5' with εy followed by βh1, βmajor and βminor at the 3' side of the locus (figure 3.2). Upstream of the locus lies a cluster of erythroid specific *cis*- regulatory elements specified by DNase I hypersensitive sites, called a Locus Control Region (LCR). More erythroid-specific distally located *cis*-regulatory DNA elements are found in the mouse β-globin locus. 3'HS1 is found 68 kb downstream of the εy cap site (Tuan et al. 1985), and two sets of recently discovered hypersensitive sites are located 85/84 kb (5' HS-84/-85) and 62/60 kb (5' HS-60/-62) upstream of εy cap site (Farrell et al. 2000; Bulger et al. 2003). Therefore the locus spans to our current knowledge over 150 kb. The mouse β-globin locus is embedded in an array of olfactory receptor (OR) genes and is located on chromosome 7 (Bulger et al. 1999).

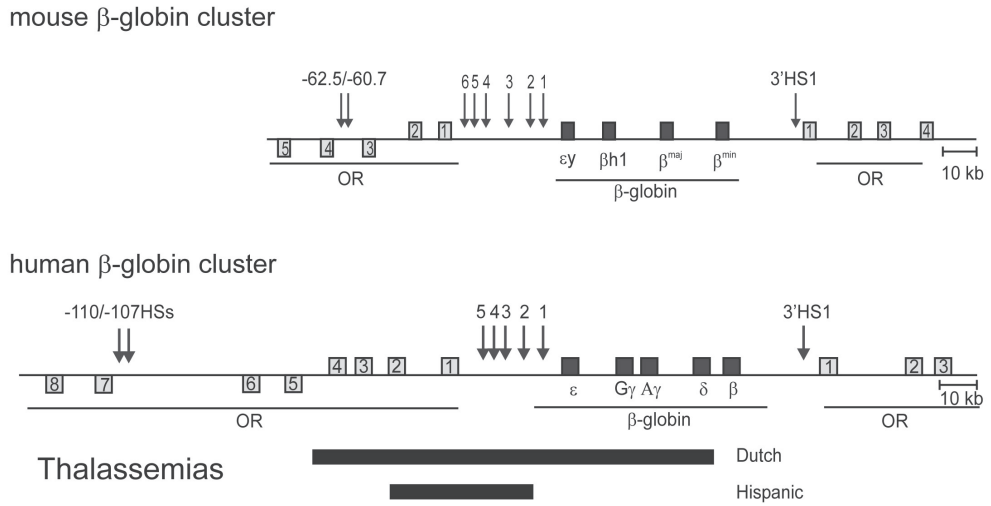


figure 3.2: The mouse and human β-globin loci.
The β-globin loci of man and mouse are shown schematically. Boxes above the horizontal lines represent genes transcribing from left to right, boxes below represent genes transcribed in the opposite direction. The globin genes are indicated (black) and corresponding erythroid cell-specific DNase I HSs are represented by arrows. The cis-regulatory DNA elements of the loci are positioned in distinct regions. Each region contains one or more HSs and names of each site are specified. Olfactory receptor genes (OR, grey) flank and partially overlap both loci. Below the human locus the size of two deletions (Dutch and Hispanic) that are known to cause β-thalassemia are indicated.

The human β-globin locus is located on chromosome 11 and has a similar, but not identical, genomic organisation as the mouse β-globin locus (figure 3.2). The human locus contains five functional genes and they are arranged in the order 5' ε-^{G}γ -^{A}γ-δ-β 3'. The *cis*-regulatory DNA elements and genes are highly conserved between the two loci and both loci have a GC-content of approximately 40% and an almost identical percentage of repetitive DNA (Moon and Ley 1990; Hardison et al. 1997; Bulger et al. 1999). In humans, a homologue of the mouse 5' HS-62 was found at 110kb and an additional site at 107 kb upstream of ε (5' HS-107/-110) but the *cis*-regulatory element corresponding to the mouse 5' HS-84/-85 is absent (Bulger et al. 2003).

The β-globin genes are relatively small genes of about 1500 bp and encode proteins of roughly 146 amino acids. The β-like globin genes have three coding regions

(exons) separated by two intervening sequences (introns). Regulatory elements, which include promoters, enhancers and silencers, flank the genes *in cis* and these are sufficient and necessary for tissue-specific expression and developmental timing of the individual genes.

The globin promoters are located within 200 bp of the respective transcriptional start sites and can be detected as DNase I HSs in the chromatin of erythroid cells. The DNase I hypersensitive state of the globin promoters depends on the stage of development; in human fetal livers, the HSs of the fetal and adult genes can be identified, while in bone marrow only the sites near the adult genes are present (Groudine et al. 1983). All globin promoters contain a TATA box and different combinations of GATA, CCAAT and CACC boxes (figure 3.3). These sequences serve as recognition sites for tissue-specific transcription factors like GATA-1, NF-E2 and EKLF, thus providing the erythroid-specificity to the promoter. The significance of these elements is illustrated by several naturally occurring mutations like for example found in the CCAAT box region of the γ-genes resulting in re-activation of γ-gene expression (Collins et al. 1985) and a mutation in the proximal CACC box of the β-globin promoter resulting in β thalassemia (Kulozik et al. 1991). Like the promoters, the enhancers, which are located in close proximity or even within the globin like genes (Behringer et al. 1987; Bodine and Ley 1987; Kollias et al. 1987), are also detectable as DNase I HSs (Groudine et al. 1983). Their importance is illustrated by the deletion of the 3' β-globin enhancer, which displayed decreased transcription of the β-globin gene without the concurrent upregulation of the γ-globin genes (Liu et al. 1997). Furthermore, silencer elements can be found at various positions in the β-globin locus (Cao et al. 1989b; Cao et al. 1989a; Ronchi et al. 1996).

The β-globin LCR

While early transgenic experiments in mice showed that the proximal regulatory elements of the adult β-globin gene were sufficient for tissue specificity and developmental timing of expression, it appeared that the expression levels of the transgene were severely influenced by the site of integration (Kollias et al. 1986; Behringer et al. 1987; Kollias et al. 1987).

A Dutch γδβ thalassemia patient provided the first indication of the existence of a distally located *cis*-regulatory DNA element. This patient carried a large 100kb deletion on one chromosome resulting in the absence of the ε- and γ-genes while the β-globin gene together with its proximal *cis*-regulatory DNA elements was unaffected (figure 3.1). The affected chromosome failed to express any β-like globin genes. Since the β-like globin genes from the unaffected chromosome express normally, the conclusion was drawn that a deletion far from the β-globin gene caused the suppression of its activity rather than a defect in *trans*-acting environment (Kioussis et al. 1983). Further analysis of the deleted region showed that several erythroid specific DNase I HSs were located upstream of the globin genes (Tuan et al. 1985; Forrester et al. 1986). Linkage of these hypersensitive sites to a β-globin gene resulted in tissue-specific, position-independent and copy-number dependent expression in transgenic mice (Grosveld et al. 1987).

The human β-globin LCR was the first LCR to be identified and was functionally defined as a DNA element that provides high levels of tissue specific expression to a *cis*-linked gene in a copy number dependent, integration site independent manner (Grosveld et al. 1987). The mouse β-globin LCR was subsequently identified on the basis of sequence

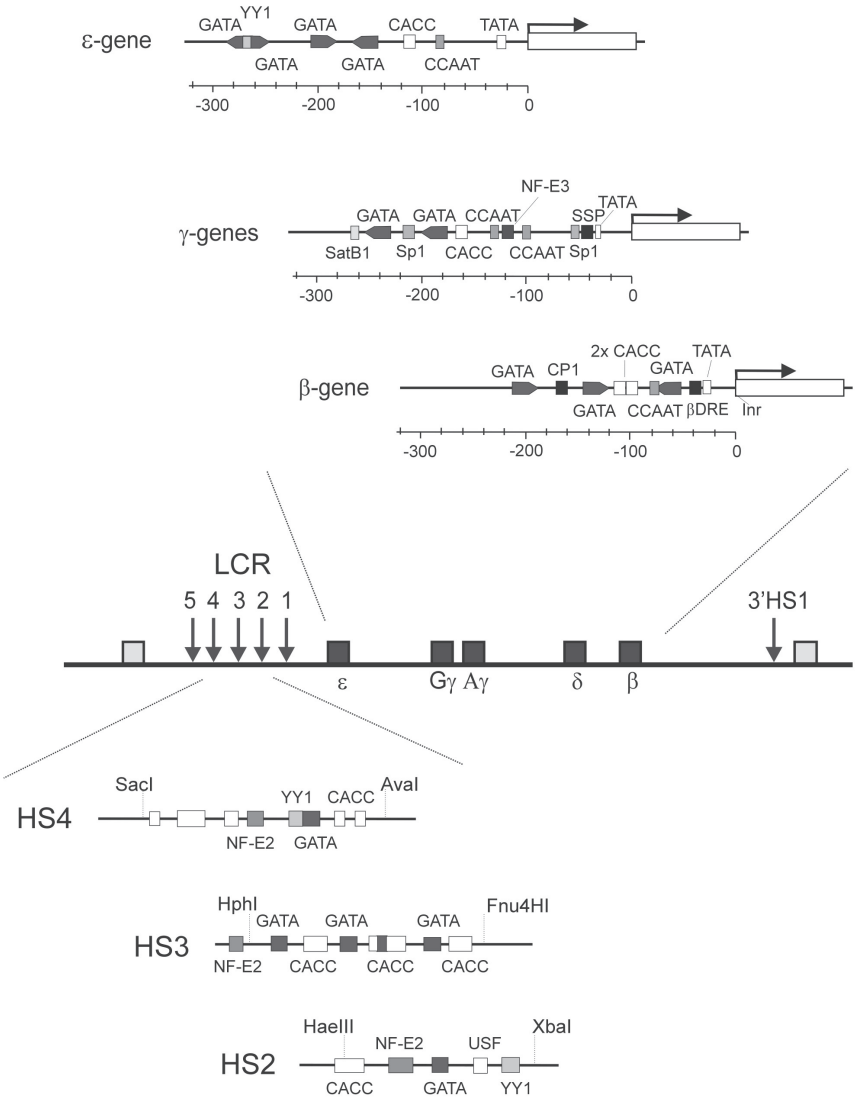


figure 3.3: The promoter regions of the human ε-, γ-, and β-globin genes and the core regions of 5' HS2, 3,and 4 of the human LCR.
The locations of putative conserved trans-acting factor binding sites in the promoters and LCR of the human β-globin locus are shown. The type of trans-acting factor that associates with a particular type of binding site can differ per promoter.

homology between the human and mouse loci (Moon and Ley 1990). The human β-globin LCR consists of five erythroid specific, developmentally stable hypersensitive sites each encompassing about 200 to 300 bp and containing binding sites for several transcription factors (Hardison et al. 1997). These transcription factors include GATA-1, NF-E2, EKLF, and Sp1 and binding to their recognition sites is required for hypersensitive (Goodwin et al.

2001). Formation of the HSs seems to precede β-globin transcription (Blom van Assendelft et al. 1989) but the LCR needs to be linked to an active promoter to stay hypersensitive (Reitman et al. 1993; Guy et al. 1996; Tewari et al. 1996).

Deletion of any single hypersensitive site from the β-globin LCR in a transgene abolishes position independent expression, suggesting that all the hypersensitive sites act together as a single entity or holocomplex (Milot et al. 1996). Analysis of β-globin pre-mRNA transcripts in single cells supported the notion of a holocomplex by demonstrating that the β-globin genes are alternately transcribed, i.e. only one gene is transcribed at any given moment (Wijgerde et al. 1995; Gribnau et al. 1998). A similar alternating transcription was observed for the mouse β-globin genes (Trimborn et al. 1999). However experiments in transgenic mice and cell lines, which tested the individual contributions of single hypersensitive sites, nevertheless showed that their functional properties are not equivalent (Tuan et al. 1989a; Collis et al. 1990; Fraser et al. 1990; Ney et al. 1990a; Ney et al. 1990b; Moon and Ley 1991).

Linking 5' HS1 of the human β-globin LCR directly to a transgene does not result in expression (Fraser et al. 1990; Fraser et al. 1993). However, HS1 does contribute to LCR function in the context of the complete locus (Milot et al. 1996). The only element that has classical enhancer properties is 5' HS2 (Tuan et al. 1989b) and this activity is mediated by a tandem NF-E2 binding site (Ney et al. 1990a; Ney et al. 1990b; Talbot and Grosveld 1991; Caterina et al. 1994). 5'HS2 contributes equally to overall expression of the locus throughout development (Fraser et al. 1993). Multi-copy integration of 5' HS2 linked transgenes is sufficient to drive position-independent and copy-number dependent expression (Ellis et al. 1993). 5"HS3 is able to control transcription of single copy integrations and is capable of establishing/maintaining DNase I HSs. This suggests a chromatin opening or remodelling activity (Ellis et al. 1996), which depends on the binding of EKLF (Wijgerde et al. 1996; Gillemans et al. 1998). 5'HS3 is the most active site during the embryonic period, and the only site capable of conferring high-level expression of the γ- and β-genes during fetal hematopoiesis (Fraser et al. 1993). A core deletion of 5'HS3 resulted in reduced expression of embryonic/fetal genes in primitive erythrocytes and position dependent β-globin expression in definitive erythrocytes (Navas et al. 1998). Substitution of 5' HS3 with 5' HS4 results in significant expression changes at every developmental stage (Bungert et al. 1995). 5'HS4 is very active in the adult stage of development (Fraser et al. 1993) but is unable as a single copy to drive expression of a linked β-globin gene (Ellis et al. 1996). 5' HS5 was originally identified as a constitutive HS with boundary/insulator activities (Tuan et al. 1985; Li and Stamatoyannopoulos 1994) and contains binding sequences for an enhancer-blocking transcription factor, called CTCF (Farrell et al. 2002). However, recently it was shown that 5' HS5 is erythroid cell-specific with enhancer-blocking activities only at the embryonic stages of development but not at later stages (Wai et al. 2003). Differences between deletion size and experimental set up make this data obtained from transgenic constructs complex and difficult to interpret; moreover the constructs are likely to suffer from multiple integrations and position effects. The data generated by targeted deletions of hypersensitive sites of the endogenous mouse β-globin locus shows discrepancies with the data obtained for the human β-globin locus by transgenic studies.

In mice, the different properties of the individual HSs are less well defined and seem to be redundant. Targeted deletions of 5' HS1, 5' HS2 and 5' HS4 reduce expression

levels of the β-globin genes to a similar extent as was observed for the human sites (Fiering et al. 1995; Bender et al. 2001). A targeted deletion of 5' HS3 results in a minimal reduction in expression levels of embryonic genes and a 30% reduction of the adult genes (Hug et al. 1996). A deletion of 5' HS5 has minimal effects on expression of the globin genes consistent with the fact that this is only a weak hypersensitive site (Bender et al. 1998). Recently, two more 5' HSs, 5' HS6 and 5' HS7 of the LCR, have been identified in mice of which the equivalent human elements have not been identified yet (Bulger et al. 1999). Targeted deletion of 5' HS6 from the endogenous mouse locus has minimal effects on expression (Bender et al. 1998). None of the hypersensitive site deletions affected the formation of the remaining sites (Bender et al. 1998; Bender et al. 2000b). This suggests that in the endogenous mouse locus no dominant or initiating site, whose formation must precede the other HSs, exists and that the hypersensitive sites have an additive effect on transcription (Bender et al. 2001) as had been suggested before for the hypersensitive sites of the human β-globin LCR (Fraser et al. 1993). The discrepancies found between human and mouse may reflect differences in functional assays used (transgenic mice versus targeted deletions), the differences in number, structure and affinities of the individual factor binding sites, or the existence of redundant elements in the endogenous mouse locus, capable of chromatin opening, which are missing from the transgenic constructs. Prime candidates for these regulatory elements are the distally located HSs (see figure 3.1) with unknown function. Recent studies indicate an active role of these elements in β-globin gene transcription (Tolhuis et al. 2002; Palstra et al. 2003) and deletion of these sites, in combination with LCR deletions, might prove to be particularly informative.

Chromatin structure of the β-globin locus

A domain of erythroid specific DNase I sensitivity extends from approximately 10 kb upstream of 5' HS-60/-62 to a few kb downstream of 3' HS1 in the endogenous mouse locus. The borders of this DNase I sensitivity domain are not marked by any regulatory sequences (Bulger et al. 2003). Acetylation of histone H3, histone H4 and methylation on lysine 4 of histone H3 (H3 meK4) is found in a domain overlapping with, but not identical to the domain of DNase I sensitivity (Forsberg et al. 2000; Bulger et al. 2003). Erythroid cell-specific histone H3 and H4 hyperacetylation and histone H3 meK4 is found at *cis*-regulatory DNA elements within the active β-globin locus (Kiekhaefer et al. 2002; Bulger et al. 2003). In definitive erythroid cells, four regions within the locus exhibit hyperacetylation: 5' HS-62/-20, the LCR and the transcriptionally active βmajor and βminor globin genes (Forsberg et al. 2000; Bulger et al. 2003). Other regions such as the inactive genes and 3' HS1 are modestly enriched in histone acetylation and methylation (Bulger et al. 2003). In erythrocytes isolated from the embryonic yolk sac, which expresses εy and βH1, the LCR and both active and inactive promoters were found to be hyperacetylated (Forsberg et al. 2000). Methylation of lysine 79 of histone H3 (H3 meK79) is found at the active βmajor promoter, but not on inactive promoters or HS2 of the LCR. In EryP cells, which express the embryonic genes, H3 meK79 is found at a domain spanning at least from 3' of the εy gene to the (active) βh1 promoter (Im et al. 2003). Hematopoietic-specific activators like GATA-1 and NF-E2 govern establishment of these chromatin modification patterns (Kiekhaefer et al. 2002; Im et al. 2003).

Chromatin analyses of mouse erythroid precursor cells, which express the globin genes at basal levels (Hu et al. 1997; Delassus et al. 1999), showed that the βmajor promoter has low to moderate acetylation levels and that 5'HS3 and the βmajor promoter are already accessible to DNase I. In fully differentiated cells, which express the genes at high levels, acetylation and accessibility is further increased (Bottardi et al. 2003). Targeted deletion of the mouse LCR shows that, even after germline passage of the deletion, the locus is still DNase I sensitive and that the remaining HSs are normally formed. This shows that in the mouse β-globin locus, the LCR is dispensable to initiate an open chromatin conformation and this suggests that elements elsewhere in the locus are sufficient to establish and maintain an open chromatin domain (Bender et al. 2000a). Acetylation of nucleosomes in the region between the LCR and the βmajor and βminor genes is only modestly reduced in the LCR deletion as compared to wild type. Similarly, hyperacetylation of histones at the βmajor and βminor promoter was similar as in the wild-type locus although the rate of transcription of the β-major gene in the LCR deletion was reduced to about four percent of the wild-type level (Schubeler et al. 2001).

The human locus, like the mouse β-globin locus, is also more sensitive to DNaseI I than 'bulk' DNA in erythroid cells (Weintraub and Groudine 1976). In a Hispanic thalassemia patient, a deletion was found that extended from 9.5 kb to 39 kb upstream of the ε-globin gene, removing 5' HS2-5 of the LCR but leaving the distally located HSs (i.e. 5' HS-107/-110 and 3' HS1) and all the β-globin genes intact. This deletion resulted in the inactivation of the intact genes, the locus had become DNase I-resistant and the remaining HSs were not formed (Driscoll et al. 1989; Forrester et al. 1990). Acetylation levels of the human locus were determined in MEL cell hybrids carrying human chromosome 11 generated by chromosomal transfer from a patient heterozygous for the Hispanic deletion. A human chromosome carrying a wild-type locus containing recombinase sites around 5' HS2-5 of the LCR showed similar DNase I-sensitivity and acetylation levels throughout the locus before as well as after recombinase mediated deletion of 5' HS2-5. Conversely, the locus on a chromosome from the thalassemia patient with the Hispanic deletion remained DNase I-insensitive and was hypoacetylated (Reik et al. 1998; Schubeler et al. 2000). These findings and experiments with the human β-globin locus in transgenic mice (see e.g.Fraser and Grosveld 1998; Bulger and Groudine 1999) suggested that the LCR is required for chromatin opening but not maintenance of the human locus. However, these conclusions are in apparent contradiction with the data on the endogenous mouse locus (Higgs 1998; Grosveld 1999).

Recent data on erythroid precursor cells of humans and mice might give an explanation for the observed functional differences between the mouse and human LCR. The authors show that the two loci are subject to different epigenetic control mechanisms in hematopoietic progenitor cells and this difference between the two loci is maintained when the human locus is introduced in transgenic mice suggesting that the primary DNA sequence rather than the organism determines the difference in this epigenetic code (Bottardi et al. 2003).

Developmental regulation of β-globin expression

In all species that contain β-globin genes, a switch in globin gene expression coincides with changes in morphology of the erythroid cell, the site of erythropoiesis, and hemoglobin composition (Stamatoyannopoulos and Grosveld 2001). The order of the mouse and human globin genes along the linear chromatin fibre reflects their expression pattern during ontogeny (figure 3.1) (Strouboulis et al. 1992; Stamatoyannopoulos and Grosveld 2001). The expression of the embryonic genes occurs in primitive cells derived from the embryonic yolk sac. In contrast the fetal/adult genes are expressed in definitive cells, which originate from stem cells derived from the AGM region of the developing embryo (Muller et al. 1994; Ling and Dzierzak 2002). In humans, two switches are observed; one switch from embryonic (ε) to fetal (γ) globin gene expression and a second switch from fetal (γ) to adult (δ/β) globin expression. Only a single switch from embryonic (εy/βh1) gene expression to adult (βmajor/βminor) gene expression is observed in mice (figure 3.4).

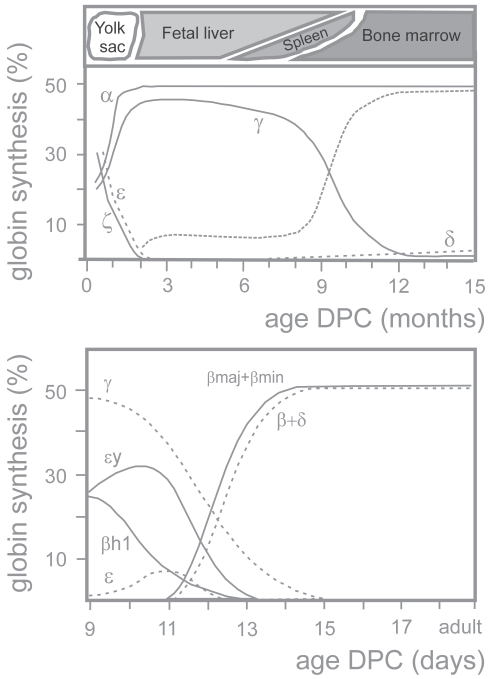


figure 3.4: Developmental regulation of globin synthesis.
Top panel shows the expression levels and the site of expression of the individual human globin genes during development. The ζ- and α-globin genes are located in the α-globin locus. The α-like and β-like globin proteins that are expressed at the same stage of development can form different types of hemoglobin. Bottom panel shows the expression levels of the individual mouse β-globin genes (solid lines) and human β-globin transgenes when expressed in mice (dotted lines).

A dual mechanism of hemoglobin gene switching has been proposed: autonomous gene control and gene competition for a direct interaction with the LCR (Hanscombe et al. 1991; Peterson and Stamatoyannopoulos 1993). The concept of autonomous gene control is largely based on experiments with human globin genes with only proximal *cis*-regulatory DNA elements (i.e. promoters, enhancers and silencers) in transgenic mice. These experiments showed that individual human transgenes express tissue-specifically and with the correct developmental timing, albeit in a position dependent manner (e.g. Magram et al. 1985; Chada et al. 1986; Kollias et al. 1986; Behringer et al. 1987). Transgenes of the human ε- and γ-

globin genes, including their proximal *cis*-regulatory DNA elements, linked to an LCR are properly silenced at later stages of development (Dillon and Grosveld 1991; Raich et al. 1992) and these silencing elements reside in their promoter regions. The proximal regions of the embryonic and fetal genes are sufficient and necessary for autonomous silencing. The first indications for gene competition emerged from HPFH patients. These patients have mutations in the γ-globin genes, which cause an increased γ-gene expression in adult life. Importantly, this leads to a decrease of β-globin gene expression from the same allele while expression of the non-mutated allele is unaffected (Giglioni et al. 1984). Expression of a β-globin transgene was found at all developmental stages when it was directly linked to the LCR (Grosveld et al. 1987; Enver et al. 1990). However, introduction of a γ- or α-globin gene in-between the LCR and the β-gene could restore correct developmental timing of β-globin gene expression. This correct timing was not achieved when the β-gene was the globin gene closest to the LCR (Hanscombe et al. 1991). Introduction of a 'marked' β-gene (βm) proximal to the LCR with respect to the β-gene resulted in a competitive advantage of proximal βm-gene expression. When the distance between the LCR and the proximal βm-gene was increased, and thus the distance between the genes decreased, the competitive advantage of the βm-gene over the β-gene was diminished (Dillon et al. 1997). In addition, inverting the gene order of the locus, thereby altering relative distances, activated the β-gene at early stages and abolished ε -gene expression (Tanimoto et al. 1999). All these experiments suggest competition of the genes for LCR function based on gene order or relative distance. Analysis of β-globin pre-mRNA transcripts in single cells supported the competition mechanism by demonstrating that the actual switching between γ- and β-gene transcription is a continuous dynamic process (a flip-flop mechanism) in which both genes can be alternately transcribed until the γ-genes are autonomously silenced (Wijgerde et al. 1995).

The *trans*-acting environment plays a crucial role in autonomous gene silencing and hemoglobin gene switching. As such, *trans*-acting factors can favour activation or silencing of specific β-globin genes (Peterson and Stamatoyannopoulos 1993). For instance, a transgenic human β-globin locus introduced in mice, expressed the γ-genes at much higher levels than was expected (Strouboulis et al. 1992). This suggests that the *trans*-acting environment in mouse embryonic erythroid cells favours activation of the γ-genes over activation of the ε-gene possibly caused by multiple changes in DNA sequences between the promoters (Strouboulis et al. 1992).

The ε-globin gene is autonomously silenced and a protein complex, direct repeat erythroid-definitive (DRED) has been identified that binds to a ε-promoter element *in vitro* and which is required for ε-gene silencing (Tanimoto et al. 2000). It has been suggested that DRED acts as a stage-specific negative regulator of embryonic and fetal transcription in definitive erythroid cells (Tanabe et al. 2002). The erythroid Krüppel-like transcription factor EKLF, an erythroid-specific member of the Sp/XKLF-family (Miller and Bieker 1993) plays a key role in β- and γ-globin gene competition during hemoglobin gene switching. Fetal livers of mice with a reduced EKLF concentration display a reduced number of transcriptionally active β-genes while a reciprocal increase in the number of transcriptionally active γ-genes is observed. However, the timing of γ-globin gene silencing remains normal in these cells. Total absence of EKLF leads to further increase in the number of transcriptionally active γ-genes, while β-gene transcription is abolished (Wijgerde et al. 1996). Recent data shows that in mice, EKLF is required for initiating or stabilising contacts between the LCR and the actively transcribed βmajor globin gene in definitive erythroid cells (Drissen et al. 2004). In another

study it was suggested that the human β-globin locus is divided into three differentially activated and developmentally regulated chromatin subdomains by specific intergenic transcripts (Gribnau et al. 2000). The authors suggested that intergenic transcription is required for chromatin remodeling of chromosomal subdomains and determines which genes are able to interact with the LCR in erythroid cells. However another study using a different approach failed to detect domain specific transcript although locus wide intergenic transcripts were detected (Plant et al. 2001).

In conclusion; the adult genes have the potential to be active at all stages of development but are silenced by the competitive advantage that the proximal embryonic and fetal genes have at this stage. A change in the *trans*-acting environment silences the early genes and shifts the competitive advantage to the adult genes that become activated.

Erythroid specific transcription factors and β-globin regulation

The *trans*-acting environment depends on which transcription factors are available and plays an important role in globin gene regulation. Several hematopoietic transcription factors that regulate β-globin expression have been identified like GATA-1, NF-E2, Friend of GATA-1 (FOG), FKLF, DRED, COUP-TFII, SSP, Id2, CBF1, and EKLF (Harju et al. 2002). Additionally, ubiquitous transcription factors such as Sp1, YY1, and USF are involved in control of ß-globin gene expression. The most important factors for β-globin expression are GATA-1, NF-E2 and EKLF, which are the best studied and will be discussed in more detail.

GATA-1

GATA-1 was the first identified member of a family consisting of six proteins, all recognising the consensus GATA motif (Orkin 1992). The GATA motif is found throughout the β-globin locus but especially in the regulatory elements of the globin genes. GATA-1 was first identified as an erythroid specific factor, however, later it was shown that GATA-1 is also present in other hematopoietic lineages such as mast cells, megakaryocytes and eosinophils (Crotta et al. 1990; Martin and Orkin 1990). GATA-1 knockout embryos die at embryonic day 10 or 11 from severe anaemia caused by the production of erythroid precursors arrested at the pre-erythroblast stage, which subsequently undergo apoptosis (Pevny et al. 1991; Weiss et al. 1994; Fujiwara et al. 1996). Overexpression of GATA-1 stimulated proliferation of pro-erythroblast cells and results in inhibition of differentiation (Whyatt et al. 1997). GATA-1 has the capacity to activate and repress gene expression. In the β-globin locus, GATA-1 plays a role as a transcriptional activator of globin expression and in agreement with this recruitment of RNA polymerase II to sites in the LCR and the βmajor promoter depends on GATA-1 (Johnson et al. 2002). GATA-1 is also involved in setting specific histone modification marks in the β-globin locus (Kiekhaefer et al. 2002; Im et al. 2003). GATA-1 has been shown to interact with several transcription factors such as EKLF, Sp1 (Merika and Orkin 1995), p300/CBP (Blobel et al. 1998) and FOG (Tsang et al. 1997; Tsang et al. 1998). The interactions of GATA-1 with these transcription factors suggest that it forms transcription factor complexes at specific sites in for example the globin locus, potentially controlling expression at different developmental time-points. Recently it was shown that GATA-1 forms distinct complexes with the essential hematopoietic transcription factor Gfi-1b, with the repressive MeCP1

complex and with the ACF/WCRF chromatin-remodelling complex. The repressive GATA-1, FOG-1, MeCP1 complex binds to repressed early hematopoietic genes and genes of the eosinophillic lineage, GATA-1/Gfi1b binds to repressed proliferative genes while the GATA-1/TAL-1 complex binds to active genes (Rodriguez P et al., submitted).

NF-E2

The erytroid transactivator NF-E2 was initially purified from MEL cells and it was shown that NF-E2 binds as a heterodimer and consists of a hematopoietic specific subunit called NF-E2p45 that contains a transcriptional activation domain and a more widely expressed subunit called NF-E2p18 or MafK (Andrews et al. 1993a; Andrews et al. 1993b; Ney et al. 1993). It was initially found to bind AP-1 sites in the promoter of the human porphobillinogen deaminase (PBGD) gene (Mignotte et al. 1989) and subsequent studies showed that the AP-1 sites present in 5'HS2 enhanced expression of reporter constructs in transfected cells (Ney et al. 1990b) and in transgenic mice (Talbot et al. 1990; Caterina et al. 1994). Studies using MEL cells indicate that NF-E2 plays a major role in globin gene expression. MEL cells that do not express NF-E2p45, cannot sustain high levels of globin expression but reintroduction of the NF-E2p45 subunit restores expression of globin genes (Lu et al. 1994). In un-induced MEL cells, both subunits of NF-E2 are present. However, one study suggested that these subunits occupy different nuclear compartments; NF-E2p45 is located in the euchromatic compartment whereas the NF-E2p18 subunit is found primarily in the centromeric heterochromatic compartment (Francastel et al. 2001). In the same study the inactive β-globin locus was found to preferentially associate with centromeric heterochromatin or the nuclear periphery. Upon differentiation, the β-globin gene loci relocate away from these heterochromatin compartments and this relocation correlates with both transcriptional activation of the globin locus and the relocation of NF-E2p18 away from heterochromatin (Francastel et al. 2001). Using chromatin immuno precipitation, a recent study showed that the NF-E2 complex is recruited to both the LCR and the active globin promoters upon differentiation of MEL cells (Sawado et al. 2001). In MEL cells before differentiation, NF-E2p18 is associated with Bach1 and recruits corepressor complexes, including NuRD, which is associated with repression of β-globin transcription (Brand et al. 2004).

A first step in the switch from the repressed state to the active state of the β-globin locus in MEL cells possibly involves the heme-induced displacement of the NF-E2p18 dimerisation partner Bach1 (Sun et al. 2004). Upon erythroid differentiation an exchange of NF-E2p18-binding partners, from Bach1 to NF-E2p45, is associated with the formation of a NF-E2p18/p45 heterodimer activator complex, resulting in β-globin gene expression (Brand et al. 2004). Moreover, NF-E2p45 is involved in setting specific histone modification marks in the β-globin locus (Kiekhaefer et al. 2002; Im et al. 2003) and recruitment of RNA polymerase II to the β-globin promoter (Johnson et al. 2001; Johnson et al. 2002) and *in vitro,* NF-E2p45 is able to facilitate transfer of RNA polymerase II from the LCR to the β-globin gene (Vieira et al. 2004). Despite the importance of NF-E2 for globin expression in MEL cells, the role of NF-E2 in development is less clear. Mice deficient for NF-E2p45 only show a subtle reduction in globin expression and erythroid maturation presumably because of functional redundancy between NF-E2p45 and other bZIP family proteins. However, these mice do suffer from the loss of production of circulating platelets (Shivdasani and Orkin 1995; Shivdasani et al. 1995).

EKLF

Many promoters, including those of the β-globin genes, contain CACC motifs and are bound by a number of proteins, like Sp1 and other Kruppel related proteins. The erythroid Krüppel-like transcription factor EKLF, an erythroid-specific member of the Sp/XKLF-family, was identified using a cDNA subtraction assay between lymphoid and erythroid transcripts (Miller and Bieker 1993). EKLF contains three zinc fingers that bind specifically to the CCACACCCT sequence, which is found in the β-promoter and 5'HS3 of the LCR (Feng et al. 1994; Gillemans et al. 1998). The transcriptional activity of EKLF seems to be downstream of GATA-1, as indicated by the presence of GATA-1 binding sites in the EKLF-promoter (Crossley et al. 1994). Although EKLF is present throughout development, and different sites containing the CACC box are present in diverse combinations in all globin gene promoters (except for the δ-globin promoter), EKLF only enhances β-globin gene expression. 5'HS3 has been implicated in chromatin opening of the β-globin locus (Ellis et al. 1996) and contains binding sites for EKLF. In fact EKLF has been shown to play a direct role in LCR function via this site by means of chromatin remodelling (Gillemans et al. 1998). EKLF interacts *in vitro* with the chromatin remodelling complex E-RC1 (Armstrong et al. 1998) and DNase I hypersensitivity of 5'HS3 is markedly reduced in EKLF knockout mice. Mice deficient for EKLF express the embryonic globin genes apparently at normal levels (Nuez et al. 1995; Perkins et al. 1995). However, these EKLF knockout mice die in utero from anaemia around dpc14, due to a failure to express the β-globin gene, although the red cells already have other defects from the embryonic period onwards (Drissen et al., submitted). Mainly due to its positive role in β-globin gene expression, EKLF plays a major role in the γ- to β-globin gene competition. In transgenic mice that contain the human β-globin locus and have a reduced EKLF level because they are heterozygous for the knockout EKLF allele, β-globin expression is reduced with a concomitant increase in γ-globin expression (Perkins et al. 1996; Wijgerde et al. 1996). The authors suggested that EKLF is important in stabilizing the interaction between the LCR and the β-globin gene. This conclusion was recently confirmed in a study investigating the role of EKLF in the active spatial organisation of the β-globin locus (Drissen et al. 2004).

The β-globin locus has in the past proven to be an excellent model system to study fundamental aspects regarding transcriptional regulation and genomic organisation of multi gene loci. Many discoveries have been made using this model system and several disputes over the correctness of models for gene regulation have been fought on the battleground of the β-globin locus. Over the last two decades a vast amount of knowledge has accumulated on the transcriptional regulation of the β-globin locus, the organisation of its *cis*-regulatory elements and its primary sequence. It is therfore perhaps not surprising that the β-globin locus was the first locus of which detailed information of its spatial organisation would be obtained (Carter et al. 2002; Tolhuis et al. 2002; Palstra et al. 2003; Drissen et al. 2004).

Chapter 4

Methods to study the 3D organisation of gene loci

Methods to study the 3D organisation of gene loci

Since the early days of the study of the nucleus, its organisation has been a major field of interest. Current knowledge on nuclear organisation is limited to ultra structural features such as chromosome domains, the nucleolus and different kinds of nuclear bodies. However, it is largly unknown how individual gene loci are organised within the nuclear space and how this organisation changes upon processes like transcription. However, new methods are emerging that try to address this question.

Microscopy techniques

Microscopy is one of the obvious techniques to study the structure of the nucleus as it could provides powerfull "direct" visual information. However, limitations in the resolution of light microscopy or the ability to find the locus of interest in electron microscopy, limits its use in studying the structure of individual loci. The influence of different fixation techniques on nuclear structure is also not clear and is an area of great concern. Recent advances in light microscopy are re-establishing the technique at the forefront of nuclear research. The current method of choice is fluorescent in situ hybridisation (FISH) and immunocytochemistry. In these techniques labelled nucleic acids and fluorescently labelled antibodies to nuclear factors are used as probes to detect specific gene loci, sites of ongoing transcription or the accumulation of nuclear proteins. DNA-FISH was recently used to show re-localisation of loci from one territory to another upon transcriptional activation (Schubeler et al. 2000; Kosak and Groudine 2002; Kosak et al. 2002; Ragoczy et al. 2003). Compaction of the Immunoglobulin heavy-chain locus upon activation of recombination was recently visualised (Kosak et al. 2002; Fuxa et al. 2004) although resolution did not permit the elucidation of the exact nature of this compaction, which probably reflects looping of the locus. Extended looped domains have been visualised on nuclear halos for the *Drosophila cut* locus (Byrd and Corces 2003) and the human dystrophin gene (Iarovaia et al. 2004) but the *in vivo* relevance of these nuclear halos is under debate.

Another strategy used is the tagging of loci with bacterial Lac-operon or Tet-operon sequences. Loci can then be followed *in vivo* in live cells and in real time by detecting the bound GFP-repressor fusion proteins. This approach has been used to show de-condensation of genomic loci upon transcription (Tumbar et al. 1999; Janicki et al. 2004). An attempt has been made to differentially tag the LCR and β-globin gene using Lac-operon and Tet-operon sequences (Tolhuis, Thesis 2003). However, this resulted in aberrant expression of the β-globin genes and problems with detection levels, raising concerns about the possible aberrant behaviour of the artificial systems.

Further advances in light microscopy, possibly combined with EM, will most likely make these techniques suitable for visualising the organisation of individual gene loci.

RNA/DNA-TRAP

Recently a novel assay called RNA-TRAP, closely related to the FISH assay, was developed to study interactions between active β-globin promoters and the regulatory elements of the β-globin LCR (Carter et al. 2002). In this assay probes against nascent RNA transcripts were used to target horseradish peroxidase (HRP) to the site of active β-globin transcription. The localised HRP activity deposits biotin on chromatin that is in close proximity to the site of active transcription. The chromatin is fragmented; the labelled chromatin is enriched and quantified using a PCR assay (figure 4.1). Using this method Carter and colleagues showed that elements of the LCR are in close proximity to the active transcribing β-globin gene. This assay depends on actively transcribing genes meaning that the measured values only represent loci in transcribing cells and this allows measurements even in mixed populations. However, because non-transcribing loci cannot be analysed it is impossible to determine if these interactions are specific for the transcribing locus or might also be present in silent loci. In addition it is very difficult to carry out the adequate controls. The dependence on nascent transcript may also hamper the application of RNA-TRAP to loci that are less efficiently transcribed. Development of a DNA-TRAP, which involves the hybridisation to specific DNA sequences might circumvent some of the problems and would allow the detection of spatial interactions between non-transcribed regions.

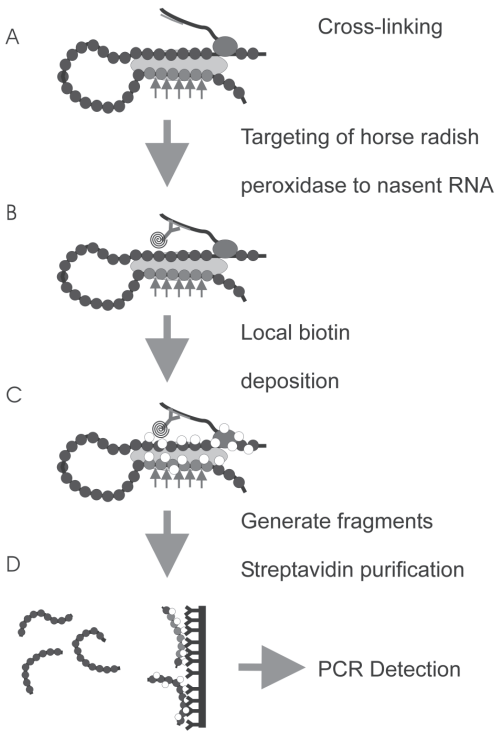


figure 4.1: RNA TRAP (tagging and recovery of associated proteins).
Cells are fixed with formaldehyde (indicated by grey oval). Gene-specific intron probes (oligonucleotides) labeled with digoxigenin are hybridized to primary transcripts (A). A Fab fragment conjugated with horseradish peroxidase (spirals) is added to localize horseradish peroxidase activity to the site of transcription. (B). Biotin-tyramide is added, which, on contact with horseradish peroxidase, becomes a highly reactive radical intermediate that covalently attaches to electron-rich moieties (principally tyrosines) in the immediate vicinity (white circles) (C). Chromatin is fragmented by sonication, and soluble chromatin is purified by affinity chromatography on a streptavidin–agarose column and DNA sequences are quantified by PCR or slot blot analysis in affinity-purified versus input chromatin (D).
See text for details (Carter et al. 2002)

Chromatin immuno precipitation

Results obtained from chromatin immuno precipitation (ChIP) experiments have been used to draw conclusions about the spatial organisation of genomic loci. In the ChIP assay cross-linked chromatin is pulled down with an antibody against a DNA-binding protein. The enriched *in vivo* bound DNA-sequences are then detected. If sequences are detected that are not known to contain binding sites for a certain transcription factor it suggests that these sites are in close proximity to a genomic site that contains a known binding site (figure 4.2). A ChIP assay to detect NF-E2 binding in the mouse β-globin locus detected NF-E2 binding at the β-globin promoter although no known NF-E2 consensus binding sites are present in the promoter (Sawado et al. 2001). One of the explanations given was a proposed looping between the NF-E2 binding sites containing HS2 of the LCR and the β-globin promoter. Later it transpired that NF-E2 was also detected at the β-globin promoter in mice that contained a LCR deletion (Bulger et al. 2002) and subsequently a cryptic NF-E2 binding site was identified in the β-globin promoter (Leach et al. 2003). Hatiz and colleagues used ChIP and an extensive panel of antibodies against factors with binding sites only in the enhancer or promoter of the HNF-4α gene to study its activation during the differentiation of CaCo-2 cells. After induction of differentiation they found recruitment of an enhancer specific complex to the enhancer and assembly of the RNA pol-II holoenzyme at the promoter of HNF-4α, respctively. A time dependent movement in the direction of the promoter via the intervening sequences of the enhancer complex followed this. Upon arrival of the enhancer complex at the promoter a stable enhancer-promoter complex was formed and transcription of the HNF-4α gene commenced (Hatzis and Talianidis 2002). In another study close proximity of the *scs* and *scs'* insulators in *Drosophila* was demonstrated using an antibody against the *scs* binding protein Zw5 to pull down the non-Zw5 binding *scs'* insulator. This interaction was verified

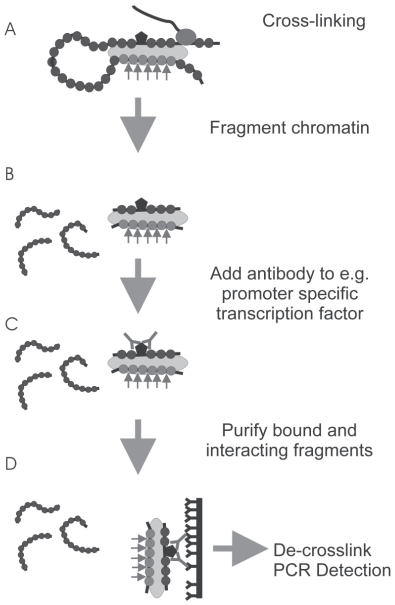


figure 4.2: Detection of chromatin interactions with ChIP.
Cells are fixed with formaldehyde (indicated by grey oval) (A). The chromatin is fragmented by sonication (B). An antibody against a specific transcription factor in e.g. the promoter region is added. A chromatin region can also be modified by introduing bacterial operator repeats in which case an antibody against the associated transgenicly expressed bacterial repressor is used. (C). The chromatin is affinity purified, de-crosslinked and interacting DNA sequences in affinity-purified versus input chromatin are quantified by PCR (D).
See text for details.

by 3C and genetic experiments (Blanton et al. 2003).

The main problem with the use of ChIP to determine interactions between different genomic loci is the difficulty to exclude the presence of cryptic binding sites or binding via an intermediary protein. One way to circumvent these problems is to tag genomic sites of interest with binding sites for bacterial or yeast transcription factors, express these transcription factors from a transgene and use antibodies against these transcription factors to pull down sequences that interact with the tagged genomic region. Tagging of the *H19* DMR by knocking in a GAL4-UAS site, expressing a GAL4-myc fusion protein and performing a pull down with antibodies against the myc epitope revealed parent-specific interactions with two DMRs located near the *igf-2* gene (Murrell et al. 2004). Interactions between promoters and terminators of large yeast genes were detected in a similar assay using antibodies against the GFP part of a *lac*I-GFP fusion protein that is able to bind to the *lac*O sequence introduced 5' of the *FMP27* promoter (O'Sullivan et al. 2004). In both cases the results were confirmed with Chromosome conformation capture. Tagging of a genomic locus of interest and expressing the associated protein is time consuming and laborious, moreover great care has to be taken that the introduction of a foreign sequence into a locus does not disrupt the normal expression pattern.

Chromosome conformation capture

Chromosome conformation capture (3C) is a powerful technique which enables the mapping of *in vivo* interactions between chromosomal regions independent of their transcriptional status and without the need to modify the locus of interest. 3C was initially used to determine the conformation of yeast chromosomes (Dekker et al. 2002). Fitting their data to different polymer models (Rippe 2001) allowed the authors to produce a population-average 3D model of chromosome III of *Saccharomyces cerevisiae* (Dekker et al. 2002). The technique was adapted to analyse the *in vivo* conformation of a 200kb region spanning the mouse β-globin locus in its active and inactive transcriptional state (Tolhuis et al. 2002).

In 3C, formaldehyde is used to trap interactions between chromatin segments by cross-linking proteins to other proteins and to DNA. The cross-linked chromatin is subsequently digested using a restriction enzyme followed by ligation under dilute conditions. In this method intra-molecular ligations of the fragments that are cross-linked to each other are favoured over inter-molecular ligations of free chromatin fragments. The relative abundance of restriction fragments preferentially ligated to a restriction fragment of interest is determined by semi-quantitative PCR and is proportional to the frequency with which the various restriction fragments interact (figure 4.3)(Dekker et al. 2002; Splinter et al. 2004). When determining the relative interactions between different genomic elements, randomly ligated DNA fragments are included in the assay to correct for differences in PCR efficiency between different primers. In order to correct for differences in template amount and quality when comparing two different samples, cross-link frequencies between two restriction fragments in a locus that transcribes similarly in both samples are measured (Tolhuis et al. 2002).

3C has several important advantages over other methods to determine the spatial organisation of a locus; cross-linking frequencies can be determined at many positions within a locus (e.g. active or inactive genes, intergenic sequences or regulatory elements) with the

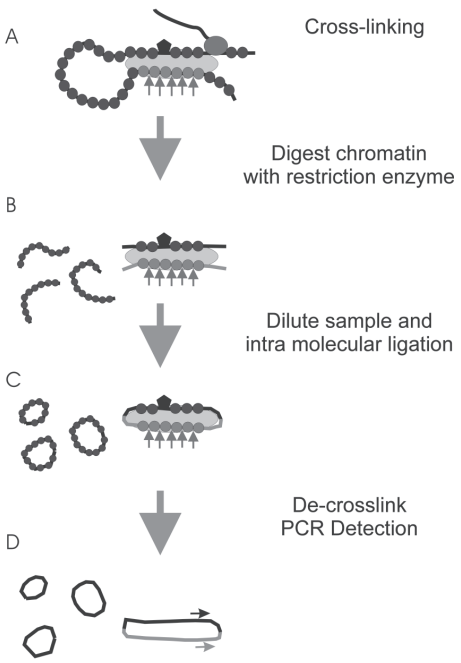


figure 4.3: 3C (Chromosome Conformation Capture). Cells are fixed with formaldehyde (indicated by grey oval) (A). The chromatin is digested with restriction enzymes (B). The chromatin is diluted followed by ligation. Intramolecular ligation is favoured as a consequence of to the dilution (C). The chromatin is decrosslinked and specific interactions are quantified by PCR. An intramolecular ligation sample is compared with a random ligation sample (D). See text for details (chapter 5; Dekker et al. 2002; Tolhuis et al. 2002)

position of restriction sites as the only limitation and it can measure cross-linking frequencies independent of the transcriptional status of a given locus. These important features make the detected interactions particularly meaningful because they can be correlated to a phenotype, e.g. transcriptional status, and it can be shown that the interactions only take place between certain specific genomic elements like *cis*-regulatory elements. Of course 3C also has its drawbacks and these will be discussed in detail in chapter 8.

The mouse β-globin locus was studied with chromosome conformation capture (3C) technology; this procedure was optimised to study the conformation of chromosomal regions in (much more complex) mammalian cells. The next chapters describe experiments that address the *in vivo* spatial organisation of a 200 kb region spanning the mouse β-globin locus in active and inactive tissues during development and erythroid differentiation and the influence that transcription factors can have on this organisation.

Chapter 5

Looping and interaction between hypersensitivity sites in the active β-globin locus

Based on: Molecular Cell 10, 1453-1465, (*2002*).

Looping and interaction between hypersensitive sites in the active β-globin locus.

BasTolhuis[1], Robert-Jan Palstra[1], Erik Splinter, Frank Grosveld and Wouter de Laat[2]

Department of Cell Biology and Genetics Faculty of Medicine Erasmus University, PO Box 1738, 3000DR Rotterdam, The Netherlands, Fax: +31-10-4089468
[1] These authors contributed equally; [2] Corresponding author: w.delaat@erasmusmc.nl

Summary

Eukaryotic transcription can be regulated over tens or even hundreds of kilobases. We show that such long-range gene regulation *in vivo* involves spatial interactions between transcriptional elements, with intervening chromatin looping out. The spatial organisation of a 200 kb region spanning the murine β-globin locus was analysed in expressing erythroid and non-expressing brain tissue. In brain, the globin cluster adopts a seemingly linear conformation. In erythroid cells the hypersensitive sites of the Locus Control Region (LCR), located 40-60kb away from the active genes, come in close spatial proximity with these genes. The intervening chromatin with inactive globin genes loops out. Moreover, two distant hypersensitive regions participate in these interactions. We propose that clustering of regulatory elements is key to creating and maintaining active chromatin domains and regulating transcription.

Introduction

Transcriptional activation in higher eukaryotes frequently involves the long-range action of a number of regulatory DNA elements. Although this has been recognised for more than 20 years, it is still not clear how enhancers (Banerji et al. 1981; Wasylyk et al. 1983), LCRs (Grosveld et al. 1987) or insulators/boundaries (Kellum and Schedl 1991; Dorsett 1999; Gerasimova and Corces 2001; West et al. 2002) exert their effect on the process of chromatin modification and transcription over distance (up to hundreds of kilobases). Many different models have been put forward to explain distant effects. The 'looping model' states that enhancers and promoters communicate through direct interactions between proteins bound to the DNA elements, with the intervening DNA looping out (Ptashne 1986; Muller and Schaffner 1990; Hanscombe et al. 1991). Other models imply a role for the DNA in between to support the transmission of some signal from enhancer to promoter. Direct support for the latter type of models comes from bacteria. Here, activation of the phage T4 late genes was found to involve loading on and sliding from the enhancer of trimeric gp45 along the DNA to the promoter to allow the forming of the transcription initiation complex (Herendeen et al. 1992). The 'looping model' also receives support from studies on transcriptional regulation of many different prokaryotic genes. In fact, the model was originally based on work on bacterial

and phage repressor proteins, like the Gal-, AraC and λ repressor proteins, which were found to function only when homo-multimerized and bound to two separate operator sites. Electron microscopy visually demonstrated the DNA in between to loop out (reviewed in Ptashne 1986). Thus, both type of mechanisms appear to function in bacteria. Eukaryotes have more complex gene clusters with regulatory elements functioning over much greater distances. To date, there are no data that unambiguously demonstrate one (or more or combinations) of the models to be correct for the regulation of a given eukaryotic locus. Support for models has come from indirect and/or *in vitro* observations and often the distinction between the activation and actual transcription of a locus is not made. However with respect to transcription, a number of observations can only easily be explained by the 'looping model'. The first type of experiments involves studies on *trans*-activation, i.e. the ability of an enhancer to activate a promoter present on a physically separate DNA molecule. Most important in this respect is the naturally occurring phenomenon of transvection in *Drosophila* (Bickel and Pirrotta 1990). In addition, Schaffner and co-workers demonstrated *in vitro* that enhancers can stimulate transcription *in trans*, by coupling an enhancer- to a promoter-containing plasmid via a biotin-streptavidin bridge (Mueller-Storm et al. 1989). Similarly, *trans*-activation of transcription was observed when enhancer-containing and promoter-containing plasmids were injected as intertwined catenates into frog oocytes (Dunaway and Droge 1989). More recently, transient transfection assays with reporter plasmids and GAGA as a DNA-bridging factor also demonstrated transcriptional activation *in trans* in mammalian cells (Mahmoudi et al. 2002). All these studies on *trans*-activation demonstrate that a cis configuration of enhancer and promoter is not an absolute prerequisite for interaction, as predicted only by the 'looping model'.

Moreover, gene competition for a single regulator (de Villiers et al. 1983; Wasylyk et al. 1983; Hanscombe et al. 1991), leading to alternate transcription (Wijgerde et al. 1995; Gribnau et al. 1998; Trimborn et al. 1999), is also most easily explained by 'looping', particularly because the competitive advantage of the enhancer-proximal gene is lost when the genes are closely spaced at further distance from the regulator (Heuchel et al. 1989; Hanscombe et al. 1991; Dillon et al. 1997). Finally, in yeast, a downstream enhancer was recently demonstrated to activate gene expression from a distance, by making use of loops induced by telomeres (de Bruin et al. 2001). However all these experiments were either done *in vitro* or are indirect in nature. None of them directly shows *in vivo* that two distal elements linked *in cis* interact by coming in close spatial proximity with intervening DNA looping out.

Here, we provide evidence that looping occurs during transcription *in vivo.* We demonstrate that the murine β-globin LCR is in physical proximity to the active globin genes *in vivo* in expressing tissue with the intervening DNA looping out. Interaction and looping are not observed in non-expressing tissue. In addition, DNase I hypersensitive sites at both end of the locus participate in these interactions, again by looping out intervening DNA. Thus, multiple hypersensitive sites spread over 130 kilobases interact to form a cluster in the nuclear space. On the basis of these data we propose that direct interactions between distal DNase I hypersensitive sites and looping out of chromatin is crucial in establishing an open chromatin domain and activating transcription.

Results

Applying 3C technology to the murine β-globin locus

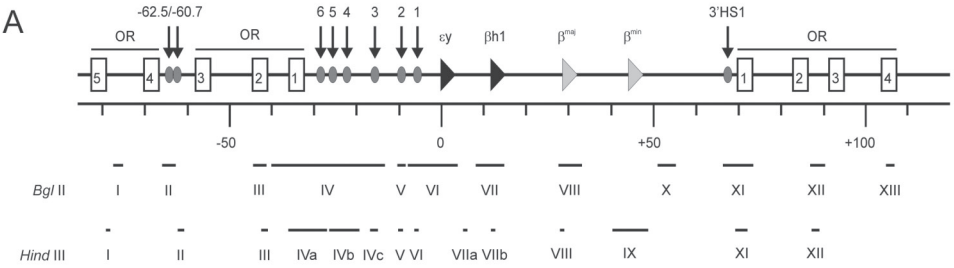
We applied methodology recently developed by Dekker et al. (2002) to gain insight into long-range interactions between the LCR and the genes in the murine β-globin locus. The principle of this technique, Chromosome Conformation Capture (3C), is that cells are treated with formaldehyde to cross-link proteins to other proteins nearby and DNA (see also figure 5.1B). The resulting DNA-protein network is then subjected to cleavage by a restriction enzyme, which is followed by ligation at low DNA concentration. Under such conditions, ligations between cross-linked DNA fragments, which is intramolecular, is strongly favored over ligations between random fragments, which is intermolecular (Dekker et al. 2002). After ligation, the cross-links are reversed and ligation products are detected and quantified by polymerase chain reaction (PCR). The cross-linking frequency of two specific restriction fragments, as measured by the amount of corresponding ligation product, is proportional to the frequency with which these two genomic sites interact (Dekker et al. 2002). Thus 3C analysis provides information about the spatial organisation of chromosomal regions *in vivo*.

A schematic presentation of the murine β-globin locus is given in figure 5.1A. Briefly, the locus contains an LCR, comprising 6 HSs (5'HS1-6), two embryonic genes, εy and βh1 (expressed in the yolk sac), and two adult genes, βmajor and βminor (expressed in fetal liver and adult spleen/bone marrow). The LCR is required for high levels of expression of all β-globin genes. Similar to what is observed in the human β-globin locus, the murine β-globin locus is flanked by olfactory receptor (OR) genes, which are inactive in globin-expressing erythroid tissue (Bulger et al. 1999; Bulger et al. 2000). Also similar to the human locus are the strong erythroid-specific DNase I HS at the 3' side (3'HS1) between βminor and the OR genes and two closely spaced HSs (HS –60.7 and HS-62.5) at the far 5' side located between 5'OR3 and 5'OR4 (Farrell et al. 2000).

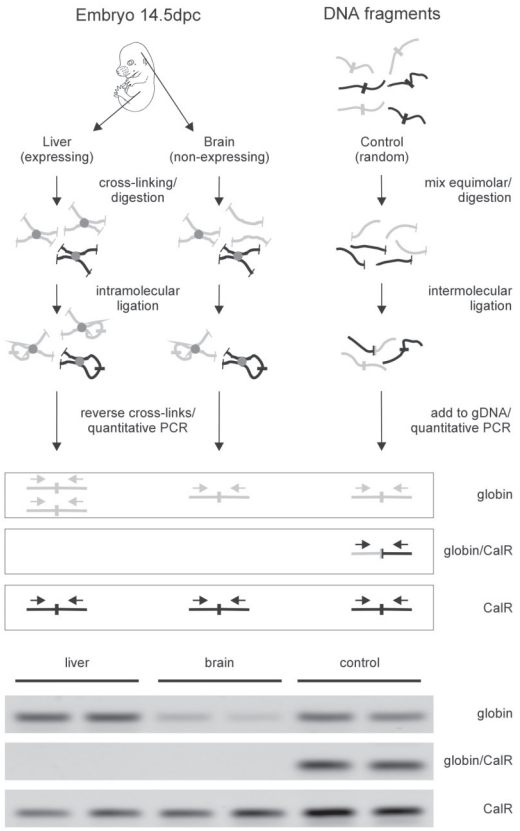
---

figure 5.1: 3C technology in the murine β-globin locus
(A) Schematic presentation of the murine β-globin locus. Black arrows and grey ellipses depict the individual HSs. The globin genes are indicated by triangles, with active genes (βmajor and βminor) in light grey, and inactive genes (εy and βh1) in black. The white boxes indicate the olfactory receptor (OR) genes (5'OR1-5 and 3'OR1-4). The two sets of restriction fragments (*Bgl*II and *Hind*III) that were used for 3C-analysis are shown below the locus. The individual fragments are indicated by roman numbers. Identical numbering between *Bgl*II and *Hind*III indicates that two fragments colocalize. Distances (roman numerals) are in kb counting from the site of initiation of the ey gene.
(B) Schematic outline of the 3C-analysis. Globin fragments (grey), CalR fragments (black), restriction sites (perpendicular bars on fragments), cross-links and PCR primers are indicated. Examples of PCR results (always done in duplo) show products obtained with *Hind*III globin fragments VIII and IV-b (top), globin fragment VIII and one of the *Hind*III CalR fragments (middle) and the two *Hind*III CalR fragments (bottom). Tissue lanes in middle panel were always empty, with every globin fragment tested. The CalR products (bottom) were used for normalizing signals.
(C) Equation, used to calculate relative cross-linking frequency between two given globin fragments [X(gl)]. A(gl): Peak area (determined with ImageQuant 5.2) of PCR signal obtained with a given globin-globin ligation product. A(CalR): Peak area of PCR signal obtained with the CalR-CalR ligation product. A(gl) and A(CalR) are determined for both the tissue (fetal liver or brain) and the control (random ligation, see results). The calculation gives a relative ligation cross-linking frequency for each tissue since it corrects for differences in PCR amplification efficiencies, cross-linking and ligation efficiencies, amounts of template and size of PCR fragments (see text). These values are plotted on the Y-axis in figure. 5.2-5.6 for the various globin fragments.

A



B



C

$$X(\mathrm{gl}) = \frac{[A(\mathrm{gl}) / A(\mathrm{CalR})]_{\mathrm{tissue}}}{[A(\mathrm{gl}) / A(\mathrm{CalR})]_{\mathrm{control}}}$$

Two independent sets of restriction fragments (*Bgl*II- and *Hind*III-fragments, respectively) were used for 3C-analysis of the β-globin locus. Each set covers the 200 kb region depicted in figure 5.1A, with intervals between analysed DNA fragments of approximately 20 kilobases or smaller. Analysis was performed on 14.5 dpc mouse fetal livers, which express the most distal globin genes, βmajor and βminor. Brain from the same 14.5 dpc embryos was simultaneously analysed as a non-expressing control tissue.

A number of experimental controls were included. First, we checked the efficiency of restriction enzyme digestion. Southern blotting and PCR analysis showed that the restriction sites analysed were cleaved without any preference for any particular region(s) after overnight incubation with an excess of enzyme (data not shown). Secondly, we determined the range of amount of template that shows linear PCR product formation. Similar ranges were found with both liver and brain template (data not shown), and roughly equal amounts (~300 ng DNA template per reaction) were used in all subsequent experiments. Thirdly, to correctly interpret signal intensities obtained with a given primer set by quantitative PCR, one needs to correct for the PCR amplification efficiency of that set. Thus, a control template is required in which all possible ligation products are present in equimolar amounts. In yeast, this was done by digesting and randomly ligating non-cross-linked genomic DNA (Dekker et al. 2002). For mammalian cells, with a genome one hundred times the size of the yeast genome, we found that random ligation of two specific loci is too rare an event to be detected by PCR. We therefore enriched for ligation products of interest by mixing equimolar amounts of DNA fragments that span each of the restriction sites analysed (see figure 5.1B). After digestion and ligation, this mix was added to genomic DNA to serve as a control template (see also Experimental Procedures). As a result, the cross-linking frequency between two loci can be expressed as the ratio of signal obtained by quantitative PCR on cross-linked template versus that obtained on control template. Fourthly, we measured the cross-linking and ligation efficiencies in both tissues to be able to compare cross-linking frequencies. This was done by comparing the cross-linking frequency between two restriction fragments present on an unrelated locus situated on another chromosome. Two neighbouring fragments were used, with the restriction sites analysed ~1.5 kilobases apart, in the transcribed part of the calreticulin locus (CalR) on chromosome 8 (the β-globin locus is on mouse chromosome 7). The CalR locus, embedded in an area of ubiquitously expressed genes, is expressed at similar levels in 14.5 dpc brain and liver (WdL, unpublished results). It is therefore reasonable to assume that it adopts a similar spatial conformation in both tissues.

Thus by normalizing each cross-linking frequency to the cross-linking frequency observed between the CalR-fragments within a tissue, we could correct for differences in amount and quality of template. Similarly, by normalizing the observed random ligation efficiency of two given fragments to that observed of the CalR fragments, we corrected for differences in the amount of control template between experiments. The equation used to calculate the relative cross-linking frequency is given in figure 5.1C. As a result of this normalisation, the "cross-linking frequency" value 1 arbitrarily corresponds to the cross-linking frequency between our control CalR fragments. Finally, the cross-linking frequencies between globin fragments and CalR fragments were always measured as an additional control. As expected for the interaction between two unrelated loci, globin-CalR cross-linking frequencies were always found to be zero (no PCR signals observed in tissues, see figure 5.1B).

The β-globin locus adopts a linear conformation in non-expressing brain cells

We performed 3C-analysis on expressing and non-expressing tissue from 14.5 dpc embryos to be able to relate the spatial conformation of the β-globin locus to its transcriptional status. figure 5.2 shows results obtained in a non-expressing tissue, the brain. Depicted are locus-wide cross-linking frequencies for two different *Bgl*II fragments ('fixed' fragments), one in the middle of the 200 kb region (fragment V) and one at the 5' end (fragment II). The central fragment V, a relatively small fragment containing HS2 of the LCR, showed the highest cross-linking frequency with the closest fragments IV and VI. Cross-linking frequency gradually decreased with fragments located further away on the linear DNA template (figure 5.2A). No significant peaks of interactions were observed between fragment V and more distal DNA fragments. Similar results were obtained for the DNA fragment at the 5' end of

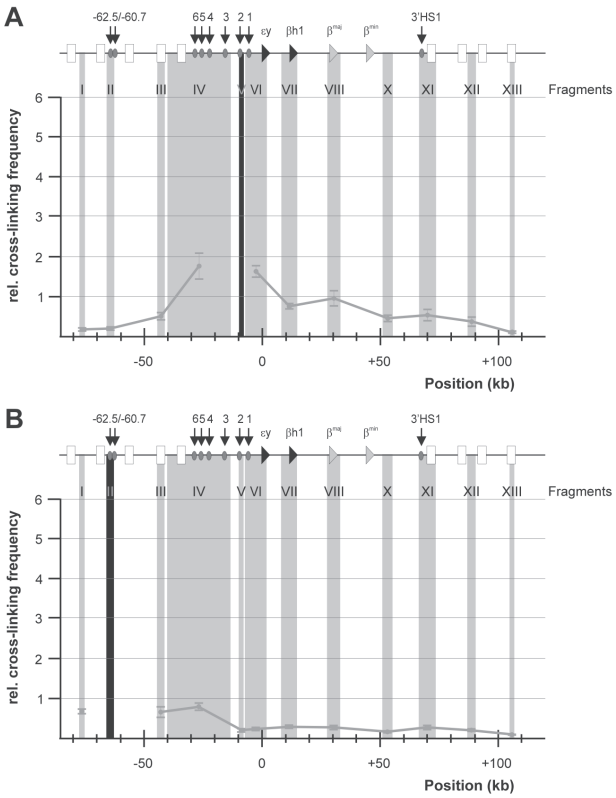


figure 5.2: Linear conformation of the β-globin locus in non-expressing brain cells.
The murine β-globin locus is depicted on top of each graph (for explanation of symbols, see figure 5.1a). X-axis shows position in the locus. Black shading shows the position and size of the 'fixed' fragment. Grey shading indicates position and size of other fragments. Standard-error-of-mean is indicated. Cross-linking frequency with a value of 1 arbitrarily corresponds to the cross-linking frequency between two neighbouring CalR control fragments (with restriction sites analysed being 1.5 kb apart). Scaling on Y-axis (from 0 to 6) allows direct comparison with figures 3-6.
(A) Relative cross-linking frequencies between 'fixed' *Bgl*II fragment V (5'HS2 in LCR) and the rest of the locus.
(B) Relative cross-linking frequencies between 'fixed' *Bgl*II fragment II (5'HS-62.5/60.7) and the rest of the locus.

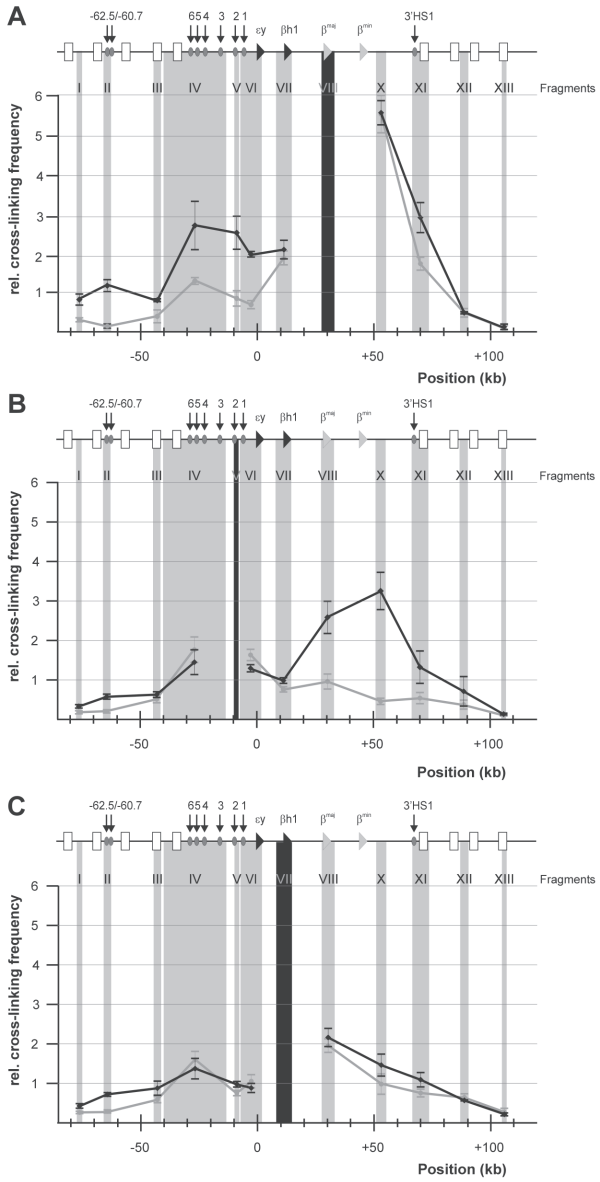


figure 5.3: Erythroid-specific interaction and looping between the LCR and an active β-globin gene.
Relative cross-linking frequencies observed in fetal liver are shown in black. For comparison, data obtained in brain are depicted in grey. Standard-error-of-mean is indicated. Cross-linking frequency with a value of 1 arbitrarily corresponds to the cross-linking frequency between two neighbouring CalR control fragments (with restriction sites analysed being 1.5 kb apart). Scaling on Y-axis (from 0 to 6) allows direct comparison with figures 5.2 and 5.4-5.6.
(A) 'Fixed' *Bgl*II fragment VIII (βmajor) versus the rest of the locus.
(B) 'Fixed' *Bgl*II fragment V (5'HS2) versus the rest of the locus.
(C) 'Fixed' *Bgl*II fragment VII (βh1) versus the rest of the locus.

the region (II) (figure 5.2B). Thus in brain, we observed a direct correlation between spatial proximity and distance along the linear β-globin DNA template. This holds for any 'fixed' fragment in this region, independent of the restriction enzyme used (see data below). Such a correlation between distance in space and distance in kilobases would be expected of a linear structure (Rippe 2001). Hence we conclude that the 200 kb region encompassing the β-globin locus adopts an essentially linear conformation in the nucleus of the non-expressing brain cell.

Spatial interaction and looping between the LCR and the active genes in the expressing fetal liver

Next, we analysed the spatial organisation of the β-globin locus in the expressing 14.5 dpc fetal liver cells. The active globin genes, βmajor and βminor, are 34 and 49 kb away from the 3' side of the LCR, respectively. We first focussed on a *Bgl*II fragment (fragment VIII) containing the active βmajor gene with all the known local regulatory elements, including the promoter and the enhancer ~1 kb downstream of the transcribed sequence. In agreement with the findings presented above, the curve for brain was indicative of a linear conformation (figure 5.3A). In fetal liver, cross-linking frequencies identical to those in brain were observed for fragments closest to fragment VIII. However, when DNA elements more towards the 5' side of the region were analysed, up to 3-fold elevated cross-linking frequencies were found in liver as compared to brain with fragments IV, V and VI. Most interestingly, these are the three *Bgl*II fragments that together cover all six hypersensitive sites of the LCR. Beyond the LCR, even further 5' from the βmajor gene, cross-linking frequencies dropped again to the levels observed in brain (with the exception of fragment II, discussed below). These data indicate that in the nucleus of the expressing fetal liver cell, the active beta major gene comes in close vicinity to the LCR.

This is confirmed when the reciprocal experiment is carried out using an LCR fragment as the fixed fragment. *Bgl*II restriction sites flank HS2 of the LCR, resulting in fragment V. When this fragment was tested versus the others in fetal liver, fragment VIII (βmajor), but also fragment X, containing DNA sequences just 3' of the active β-minor gene, showed highly elevated cross-linking frequencies in fetal liver compared to brain (figure 5.3B). In fact, in fetal liver, but not in brain, the cross-linking frequency between HS2 and the active adult genes is much higher than that between HS2 and the inactive, embryonic genes (εy and βh1, present on fragment VI and VII, respectively). Thus, these data show that in expressing cells, the β-globin LCR and the distal active genes come in physical proximity, whereas the inactive genes appear to be located further away from the LCR fragment V.

In order to determine whether in fetal liver the inactive genes indeed do not come in close proximity to other sequences in the locus, we looked at locus-wide cross-linking frequencies of the βh1 gene (fragment VII). Almost identical cross-linking frequencies between βh1 and the rest of the locus were observed in liver and in brain for both a *Bgl*II (figure 5.3C) and *Hind*III digest (not shown). Similar results were obtained for a *Hind*III fragment close to εy (VII-a, not shown & see figure. 5.4-5.6). This suggests that the inactive genes are not interacting with the LCR. We conclude that the LCR interacts specifically with the active distal β-globin genes with intervening DNA containing the inactive genes looping out.

All hypersensitive sites of the LCR participate in the long-range interactions

Whereas *Bgl*II cuts relatively infrequently in the murine β-globin locus, resulting in the large fragments analysed and described above, digestion by *Hind*III yields smaller DNA fragments, which may allow fine-mapping of the interactions. Most relevant to our studies, *Hind*III cuts in between most of the hypersensitive sites of the LCR (with the exception of HS4 and HS5, which are present on one *Hind*III fragment). Analysis of cross-linking frequencies with a fixed *Hind*III fragment VIII, containing 300 base-pairs of the β-major promoter plus one third of the coding part of this gene, confirmed the fetal liver-specific interaction with the LCR (figure 5.4A). In fact, elevated cross-linking frequencies with the βmajor fragment were

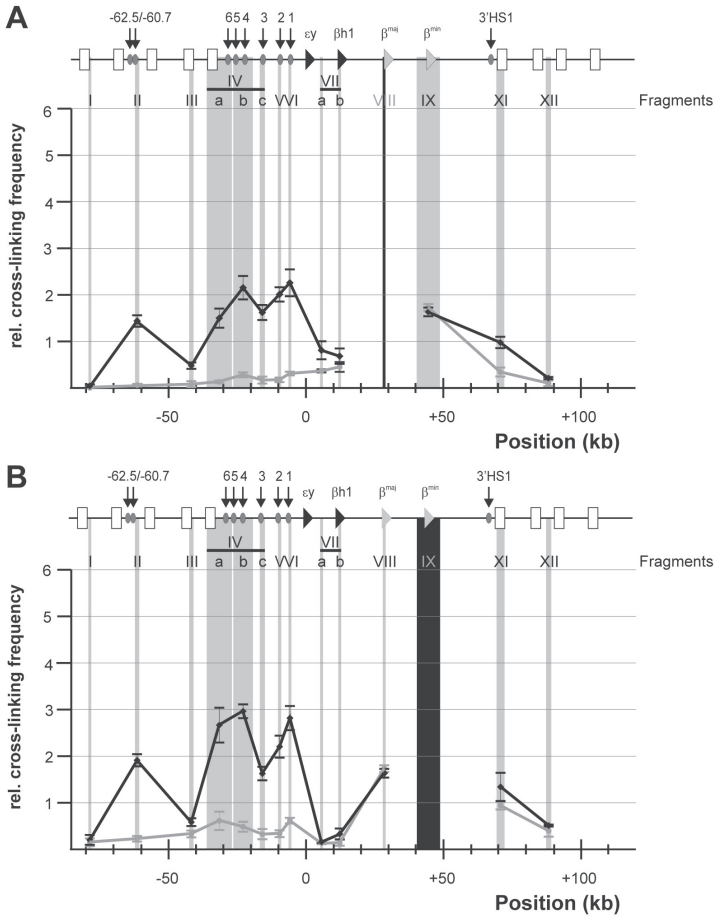


figure 5.4: Erythroid-specific interactions between the active β-globin genes and individual hypersensitive sites in the LCR.

Relative cross-linking frequencies observed in fetal liver (black) and brain (grey) are shown. Standard-error-of-mean is indicated. Cross-linking frequency with a value of 1 arbitrarily corresponds to the cross-linking frequency between two neighbouring CalR control fragments (with restriction sites analysed being 1.5 kb apart). Scaling on Y-axis (from 0 to 6) allows direct comparison with other figures.

(A) 'Fixed' *Hind*III fragment VIII (βmajor) versus the rest of the locus.

(B) 'Fixed' *Hind*III fragment IX (βminor) versus the rest of the locus.

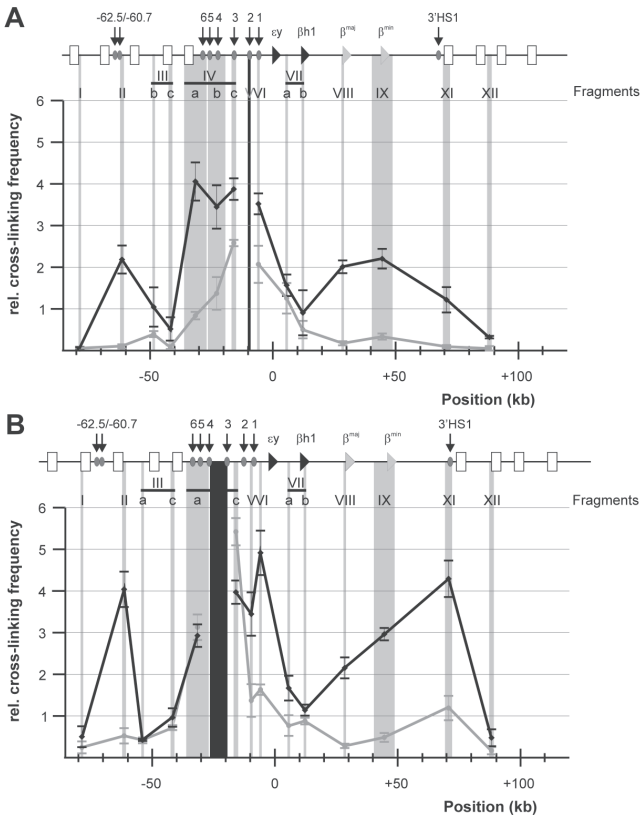


figure 5.5: Erythroid-specific high cross-linking frequencies among the individual hypersensitive sites of the LCR and two distal hypersensitive sites.
Relative cross-linking frequencies observed in fetal liver (black) and brain (grey) are shown. Standard-error-of-mean is indicated. Cross-linking frequency with a value of 1 arbitrarily corresponds to the cross-linking frequency between two neighbouring CalR control fragments (with restriction sites analysed being 1.5 kb apart). Scaling on Y-axis (from 0 to 6) allows direct comparison with other figures.
(A) 'Fixed' *Hind*III fragment V (5'HS2 of the LCR) versus the rest of the locus.
(B) 'Fixed' *Hind*III fragment IV-b (5'HS4-5 of the LCR) versus the rest of the locus.

observed for all fragments containing a hypersensitive site of the LCR (fragments IV-a, -b and -c, and fragment V and VI). As seen in the *Bgl*II experiments, cross-linking frequencies with βmajor dropped for fragments flanking the LCR (again with the exception of fragment II, discussed below). Thus, the *Hind*III data indicate that all individual hypersensitive sites of the LCR (HS1-6) participate in long-range interaction. The same results were obtained with fragment IX, encompassing the active βminor gene (figure 5.4B), although here the data suggest that HS2 (fragment V) and HS3 (fragment IV-c) do not participate as actively in the interaction as the other hypersensitive sites do. This may indeed be the case, but it may equally well reflect a technical problem (see discussion). Nevertheless these data strongly support the hypothesis that the individual hypersensitive sites of the LCR act together to contact distal genes in the fetal liver (Ellis et al. 1993; Wijgerde et al. 1995).

If indeed the LCR forms one spatial entity in expressing cells, tissue-specific high cross-linking frequencies among the individual hypersensitive sites of the LCR would be expected. This is indeed what we observe. For example, taking HS2 (fragment V) as the 'fixed' fragment, we found fetal liver-specific high cross-linking frequencies with all other hypersensitive sites of the LCR (figure 5.5A). Similar results were obtained with fixed fragment IV-b (HS4-5, figure 5.5B), IV-a, IV-c, and VI (HS6, HS3 and HS1 respectively, data

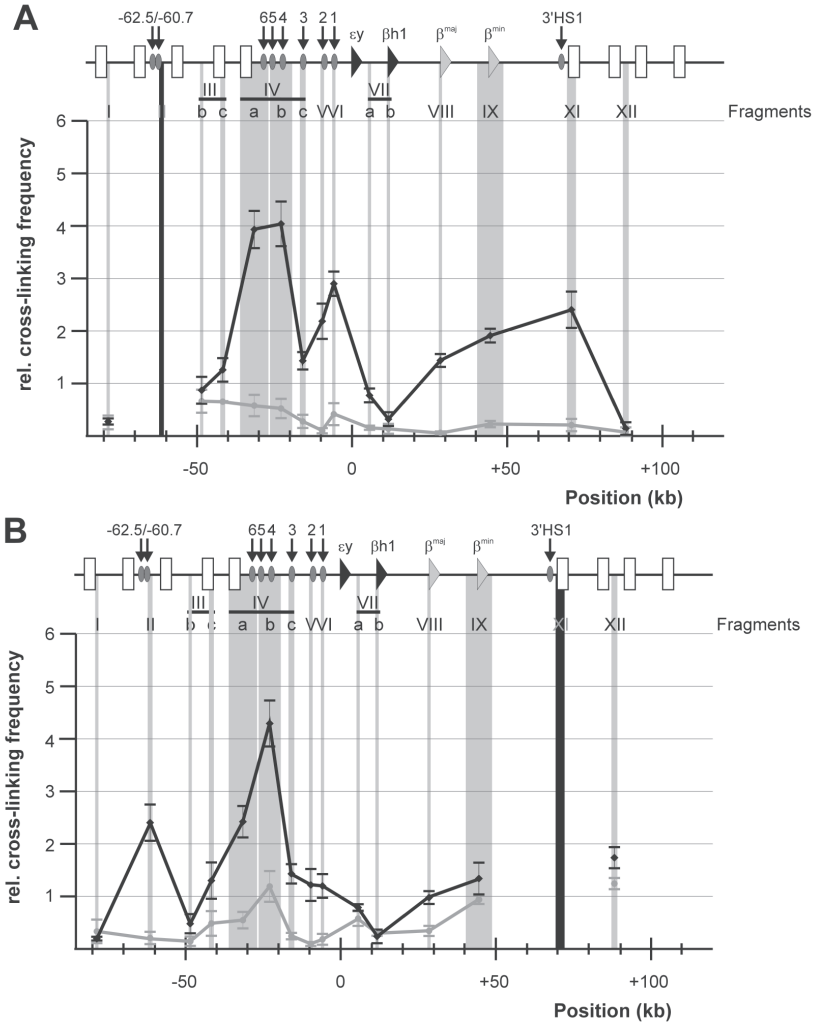


figure 5.6: Two distal hypersensitive sites at each side of the locus cluster with the LCR and the genes.
Relative cross-linking frequencies observed in fetal liver (black) and brain (grey) are shown. Standard-error-of-mean is indicated. Cross-linking frequency with a value of 1 arbitrarily corresponds to the cross-linking frequency between two neighbouring CalR control fragments (with restriction sites analysed being 1.5 kb apart). Scaling on Y-axis (from 0 to 6) allows direct comparison with other figures.
(A) 'Fixed' *Hind*III fragment II (5'HS-62.5/-60.7) versus the rest of the locus.
(B) 'Fixed' *Hind*III fragment XI (3'HS1) versus the rest of the locus.

not shown). Together, these data provide strong support for the LCR acting as a 'holocomplex' in erythroid cells to activate the globin genes.

HSs at both ends of the locus participate in the interactions between the LCR and the active genes

Two other erythroid-specific interactions stand out. In figure 5.5B, for example, high cross-linking frequencies were observed between HS4/5 and the fragments II and XI, at the far 5' and 3' end of the region, respectively. Interestingly, fragment II contains (part of) the recently identified hypersensitive sites –62.5 and –60.7 (Farrell et al. 2000), and fragment XI is located just 3' of another erythroid-specific hypersensitive site, 3'HS1 (Tuan et al. 1985; Grosveld et al. 1987). Interaction with both of the distal hypersensitive sites was seen with all other hypersensitive sites of the LCR, both in the *Hind*III experiments (see figure 5.5A and data not shown) and in the *Bgl*II experiments (see figure 5.3B, and data not shown). Moreover the active βmajor and βminor genes also showed erythroid-specific interactions with 5'HS62.5/-60.7 (figures 5.3A, 5.4A and 5.4B), despite being approximately 100kb away. These data suggest a complex series of interactions between hypersensitive sites in the β-globin locus in expressing tissue.

To further investigate this, we analysed locus-wide interactions with the distal hypersensitive sites. figure 5.6A shows the results for fragment II, which confirm the interaction between 5'HS-62.5/-60.7 and LCR elements in the fetal liver. Fragments I and III, flanking these 5'HS, do not participate in this interaction (both in the *Bgl*II and *Hind*III digestions), suggesting that the intervening DNA loops out. High cross-linking frequencies were also found between 5'HS-62.5/-60.7 and 3'HS1, which is remarkable considering the two sites are 130 kb apart on the linear chromatin template. Comparable interactions were observed using 3'HS1 as the 'fixed' fragment (figure 5.6B). However it should be noted that the data for 3'HS1 are similar to those found for βmajor and βminor and that this region appears to act as one block. The latter may point at some compaction, perhaps caused by the large amount of repetitive DNA present in this region (Bulger et al. 1999). Nevertheless, our data demonstrate that all the hypersensitive sites and the active genes of the β-globin locus cluster together in space in the erythroid nucleus

Discussion

The 'looping model' postulates that regulatory elements and genes/promoters communicate through direct interactions between proteins bound to the DNA, with intervening chromatin looping out. In this paper we have demonstrated that the distal regulatory elements and the active genes, which are linked *in cis* in the murine β-globin locus, interact *in vivo* while the intervening DNA loops out. This looping is only seen in expressing cells and provides direct *in vivo* evidence for the 'looping model'. Previous support for this model has come from several types of studies. *Trans*-activation, i.e. the ability of an enhancer to activate a promoter located on a physically separate DNA molecule, is most easily explained by direct contact between the enhancer and the gene. This has been observed in transvection in *Drosophila* (Bickel and Pirrotta 1990) and in a number of *in vitro* experiments with artificial DNA constructs (Dunaway and Droge 1989; Mueller-Storm et al. 1989; Mahmoudi et al. 2002). Competition between genes for a single regulator (de Villiers et al. 1983; Wasylyk et

al. 1983; Hanscombe et al. 1991) leading to alternate transcription (Wijgerde et al. 1995) is also most easily explained by looping, particularly because the competitive advantage of the enhancer proximal gene is lost when the genes are closely spaced at distances further from the enhancer (Heuchel et al. 1989; Dillon et al. 1997). However, all this evidence is indirect and each can also be explained by other mechanisms. The findings presented here show direct evidence for looping in the active β-globin locus, whereas a linear type of structure is found for the non-expressing locus. In particular, the observation that two hypersensitive sites at the far ends of the region cluster with the LCR and the active genes (i.e. all hypersensitive sites) provides new insights into long-range interactions (see below). However, the limitations of the 3C technique should also be noted in order to avoid overinterpretation of the results.

Interpreting 3C-analysis of the β-globin locus

Some technical and biological aspects of the results by 3C-analysis should be considered. As pointed out originally by Dekker et al. (2002), measuring cross-linking efficiency by the formation of ligation products largely depends on the frequency with which two genomic sites interact. They showed that contributions of other parameters, such as local protein concentrations or a favorable geometry of the cross-linked intermediate, are minor. Our results support this notion. However, we further believe that additional parameters, e.g. the fragment size, notably affect the cross-linking efficiency. Comparison of cross-linking frequencies observed with the large (26kb) *Bgl*II fragment IV (covering HS3-6 of the LCR and 12 kb upstream), to those observed with the much smaller *Hind*III fragments IV-a, -b and –c (containing HS6, 4-5 and 3 as separate entities) reveals an increased background in brain for the large fragment. This can be explained by assuming that the chance of being crosslinked per se increases with fragment size. Also, an increase in ligation to irrelevant fragments will compete with ligation to specifically interacting fragments, causing underestimation of specific interactions in the fetal liver. Thus, to determine whether a specific interaction occurs between two given DNA sequences, it is best to study smaller fragments containing isolated entities.

The accuracy of signals obtained with the control template is crucial for our analysis. Since cross-linking values in brain and in liver are both normalized to the same control value, we were concerned about the fact that *Hind*III fragment IV-c showed a dip in relative cross-linking frequency with every fragment tested, both in brain and in liver. This result was due to high PCR signals in the control rather than low signals in the tissue samples (data not shown). Designing new primers did not solve this problem. Thus, although the observed cross-linking frequencies with *Hind*III fragment IV-c may be real, it is more likely that it reflects an as yet unresolved technical issue.

Purely biological parameters also play a role. For example, in 14.5 dpc fetal liver about 15-20% of the cells is not expressing globin (judged by many RNA FISH experiments). These are likely to adopt a conformation similar to that observed in brain and contribute to the total amount of substrate in the ligation reaction, but not to the specific ligation frequency. Thus the real value of erythroid-specific interactions will be underestimated, which increases the significance of finding these interactions, particularly the ones over large distances. Perhaps most importantly, interactions between distal DNA elements are thought to be dynamic (Wijgerde et al. 1995), while these measurements represent steady-state average levels. For example, a very important, but short-lived interaction for transcription initiation (Wijgerde et al. 1995) may score much lower than a more long-lived interaction that would

only stabilise the complex.

Given these limitations, and the unknown dimensions of the chromatin fiber in the globin locus *in vivo*, the results presented here do not allow a strictly quantitative interpretation or conclusions as to what HS is responsible for a given interaction and/or function. Predictions about the dynamics of the interactions or real nuclear distances are therefore not possible at this stage of development of the technique.

The hypersensitive sites, looping and an open chromatin domain

Despite the limitations of the 3C technique, we can conclude that the 6 hypersensitive sites of the LCR, HS1-6, interact with the active genes, βmajor and βminor in the 14.5 dpc fetal liver, with the inactive εy and βh1 genes on the intervening DNA fiber looping out. The upstream 5'HS-62.5/-60.7 participate in this interaction, again with the intervening DNA looping out. At the other end of the locus the 3'HS1 is also involved in the contacts, but we have no evidence for DNA looping out between the genes and 3'HS1. This region contains a large amount of repetitive DNA and may adopt a compacted structure, as it appears to act in concert. The data also show a subdivision of the interactions, because we consistently

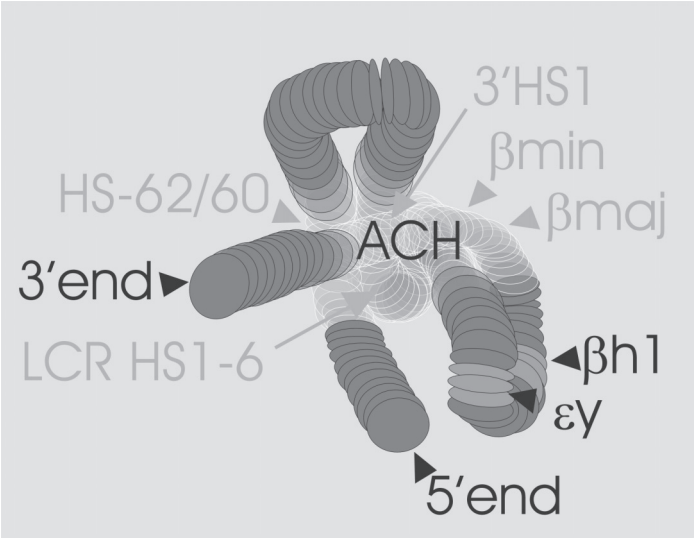


figure 5.7: A 3D model of the ACH.
A hypothetical model of the Active Chromatin Hub (ACH) is shown to illustrate the 3D nature of the ACH (not to scale), not the actual position of the elements relative to each other in vivo. Light grey indicates the active regions (hypersensitive sites and active genes) of the locus forming a hub of hyperaccessible chromatin (ACH). The inactive regions of the locus, having a more compact chromatin structure, are indicated in grey, with the inactive bh1 and ey genes in lighter grey. The olfactory genes are not shown. The interactions in the ACH would be dynamic in nature, in particular with the active genes (βmajor and βminor), which are alternately transcribed.

observe the extreme 5' and 3' HS (5'HS-62.5/-60.7 and 3'HS1, respectively) to be closer to the 5' half of the LCR (HS4-6), which is not observed for the expressed genes.

The clustering of all hypersensitive sites in the β-globin locus is intriguing. Interactions are not confined to the outermost HSs (we cannot exclude the presence of even

more distal erythroid-specific hypersensitive sites), as proposed in boundary models (for review, see Gerasimova and Corces 2001), nor to sequences that have been proposed to act as insulators (Farrell et al. 2002), but include all HSs and the promoters/enhancers of the genes. Thus, rather than being a particular type of transcription element, hypersensitivity appears to be the determining criterion for a DNA element to participate in clustering. We anticipate that this clustering is not confined to the β-globin locus only. We propose to name a 3D clustering of hypersensitive sites an 'active chromatin hub' or ACH (figure 5.7). Its formation is required to initiate transcription in repressive chromatin surroundings. The affinity between distal DNA hypersensitive sites determines whether an ACH is productively formed or not. Affinity depends on the transcription factors bound to these DNA elements and can therefore be modulated (Milot et al. 1996; Wijgerde et al. 1996; Lundgren et al. 2000). Entry of new HSs may stabilize or destabilize existing interactions, which in turn can alter expression levels of genes present in the ACH. The model does not predict how DNA sequences become hypersensitive in the first place (e.g. by mass action)(Locke et al. 1988), but stabilisation/ maintenance of hypersensitivity is proposed to depend on ACH formation. Surrounded by less active chromatin, the ACH would create a biphasic system, ensuring and stabilising a high local concentration of transcription factors and associated chromatin modifying proteins to allow efficient transcription. The hypersensitive regions and promoters of the genes would have very high levels of for example histone acetylation (Bulger et al. 2002; Burgess-Beusse et al. 2002), whereas the chromatin outside the ACH would be less acetylated. An ACH need not occupy a fixed position in the nucleus, but can be a dynamic fluid entity, possibly inside the Interchromatin Domain (ICD) compartment (Cremer et al. 2000). We propose that stable formation of an ACH underlies position-independent expression in transgenic experiments, which indeed can be accomplished by various combinations of HSs. Such a scenario would explain why multicopy inserts may give position independent expression (Ellis et al. 1993; Sabbattini et al. 1999).

Although formation of HSs in the LCR precedes transcription (Groudine and Weintraub 1982; Blom van Assendelft et al. 1989), we presently do not know whether the same holds for ACH formation. However, it is tempting to speculate that the ACH would take shape first, creating the appropriate environment, by modification of the locus, to recruit the actual transcription machinery. The observation that the globin genes are alternately transcribed (Wijgerde et al. 1995; Trimborn et al. 1999) shows that only one of the genes is transcribed at any given moment. This implies that there is only one position of interaction within the ACH that allows initiation of globin gene transcription. In other gene clusters such a ‘productive’ interaction may become stabilised and explain for example single gene expression (Olfactory Receptor genes).

We presently do not know how looping in the β-globin locus is accomplished. Although we like to think that initial contact occurs through random collision between distal elements, we cannot exclude other mechanisms to be involved in loop formation. Also, we do not know whether sequences other than HSs (and cognate factors) participate directly in the ACH or perhaps stabilise its structure. Evidence from both Drosophila (Morcillo et al. 1997; Sipos et al. 1998; Rollins et al. 1999; Zhou et al. 1999; Zhou and Levine 1999) and mammalian systems (Liu et al. 1997; Calzolari et al. 1999; Kmita et al. 2000) strongly suggests that there are elements and protein factors that stabilise long range interactions. It will be interesting to determine whether such sequences are indeed part of the ACH.

Experimental procedures

Chromosome Conformation Capture (3C)

We used the procedure recently developed by Dekker and co-workers (Dekker et al. 2002) with small adaptations to determine the spatial organization of the murine β-globin locus in 14.5 dpc embryos. Per experiment 10-12 fetal livers or fetal brains were resuspended in DMEM supplemented with 10% FCS. The equivalent of 2 fetal livers or 4 fetal brains (approximately $4x10^7$ cells) was diluted to 50 ml with DMEM (10% FCS). Formaldehyde was added to 2%, and the samples were cross-linked for 10 minutes at room temperature. The reaction was quenched by addition of glycine to 0.125M. Nuclei were harvested by lysis of the cells in ice-cold lysis buffer (10mM Tris, 10mM NaCl, 0.2% NP-40, pH=8.0) containing protease inhibitors. Nuclei were resuspended in the appropriate restriction buffer containing 0.3% SDS and incubated for 1 hour at 37°C while shaking. Triton X-100 was added to 1.8% and the nuclei were further incubated for 1 hour at 37°C to sequester the SDS. The cross-linked DNA was digested overnight with the restriction enzyme (*Bgl*II or *Hind*III). Overnight incubation at 37°C did not result in any specific loss of hypersensitive sites due to the action of endogenous nuclease activity (data not shown). The restriction enzyme was inactivated by addition of SDS to 1.6% and incubation at 65°C for 20 minutes. The reaction was diluted (to 2.5ng/µl of genomic DNA) with ligase buffer (30mM Tris-HCl, 10mM $MgCl_2$, 10mM DTT, 1mM ATP, pH 7.8) and Triton X-100 was added to 1% and incubated for 1 hour at 37°C. The DNA was ligated using T4 ligase for 4.5 hours at 16°C followed by 30 minutes at room temperature. Proteinase K was added and samples were incubated overnight at 65°C to reverse the cross-links. The following day samples were incubated for 30 minutes at 37°C with RNAse and the DNA was purified by phenol extraction and ethanol precipitation.

To prepare a control template with detectable amounts of randomly ligated DNA fragments, we had to enrich for ligation products of interest (see also results). PCR fragments spanning the restriction sites of interest were gel purified and the DNA concentration was carefully determined using a Cary 100 Bio spectrophotometer (Varian). Equimolar amounts of the different PCR fragments were mixed and digested with the appropriate restriction enzyme followed by ligation. The mix was purified by phenol extraction and ethanol precipitation. The ligated fragments were diluted to the appropriate concentration (see below) and mixed with ~30 ng digested and ligated genomic DNA.

PCR analysis of the ligation products.

The linear range of amplification was determined for the fetal liver samples and fetal brain samples by serial dilution. An appropriate amount of DNA within the linear range (typically ~30ng of DNA, for both liver and brain) was subsequently used for the experiments. The linear range of the control template was determined with a serial dilution of the random ligation mix made in the same amount (~30ng) of digested and ligated genomic DNA. Standardly, the 5' side of each restriction fragment was used to design primers unless this coincided with repetitive DNA sequences. Primer sequences are available on request. PCR products were run on 2% agarose gels and quantified on a Typhoon 9200 imager (Molecular Dynamics). All data points were generated from an average of five (with a minimum of three) different experiments performed in duplo. PCR products of the ligated fragments were run on agarose gels and quantitated. Cross-linking frequencies were calculated using the equation shown in figure 5.1C. All probes (I-XIII) were tested against all other probes. A selection of the results is presented, and data not shown are in agreement.

As shown before (Dekker et al. 2002) formation of ligation products was strictly dependent on both ligation and cross-linking, i.e. lowering the amount of formaldehyde resulted in loss of PCR product, as did the omission of T4 ligase (data not shown).

Acknowledgements

We would like to thank John Strouboulis, Niels Galjart and Sjaak Philipsen for comments on the manuscript. Also, we thank our collegues in the lab for help with the logistics. This work is supported by a NWO grant to WdL and by NWO and EC grants to FG.

Chapter 6

The β-globin nuclear compartment in development and erythroid differentiation

Based on: Nature Genetics 35, 190-194, (*2003*).

The β-globin nuclear compartment in development and erythroid differentiation

Robert-Jan Palstra[1], Bas Tolhuis[1], Erik Splinter, Rian Nijmeijer, Frank Grosveld and Wouter de Laat[2]

Department of Cell Biology and Genetics Faculty of Medicine Erasmus University, PO Box 1738, 3000DR Rotterdam, The Netherlands, Fax: +31-10-4089468
[1] These authors contributed equally; [2] Corresponding author: w.delaat@erasmusmc.nl

Summary

Efficient transcription of genes requires a high local concentration of the relevant *trans*-acting factors. Nuclear compartmentalization can provide an effective means to locally increase the concentration of rapidly moving *trans*-acting factors, and may be achieved by spatial clustering of chromatin-associated binding sites for such factors (Droge and Muller-Hill 2001; Misteli 2001a; Carmo-Fonseca 2002; Chubb and Bickmore 2003; Isogai and Tjian 2003). Here we analyse the structure of an erythroid-specific spatial cluster of *cis*-regulatory elements and active β-globin genes, the Active Chromatin Hub(ACH) (Tolhuis et al. 2002), at different stages of development and in erythroid progenitors. We show, in mouse and man, that a core ACH is developmentally conserved and consists of the hypersensitive sites (HS1–HS6) of the locus control region (LCR), the upstream 5′ HS–60/–62 and downstream 3′HS1. Globin genes switch their interaction with this cluster during development, correlating with the switch in their transcriptional activity (Stamatoyannopoulos and Grosveld 2001). In murine erythroid progenitors that are committed to, but do not yet express β-globin, only the interactions between 5′ HS–60/–62, 3′ HS1 and HS at the 5′ side of the LCR are stably present. Upon induction of differentiation, these sites cluster with the rest of the LCR and the gene that gets activated. We conclude that during erythroid differentiation, *cis*-regulatory DNA elements create a developmentally conserved nuclear compartment dedicated to RNA polymerase II-transcription of β-globin genes.

Results

The mouse and human β-globin locus contain an upstream LCR and multiple β-like genes arranged from 5′ to 3′ in order of their developmental expression (figure 6.1A). In addition there are several distal hypersensitive sites (HS), including a downstream 3′ HS1 (approx. 20kb 3′ of the β genes) and two upstream HS, ~60 kb (mouse) and ~110 kb (human) away from the genes (Farrell et al. 2000). The loci are embedded in an olfactory receptor gene cluster that is inactive in erythroid cells (Bulger et al. 2000). To investigate the spatial organisation of β-globin gene loci in mouse and man during development and erythroid differentiation, we applied chromosome conformation capture (3C) technology (see

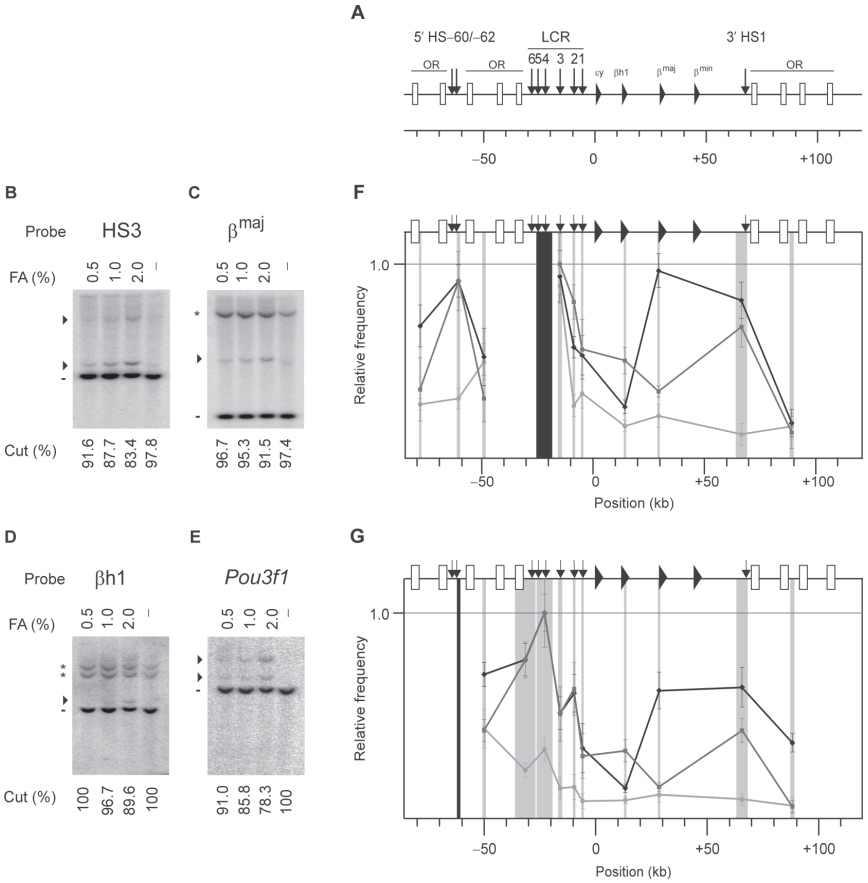


figure 6.1: Spatial organization of the murine β-globin locus.
(A) Schematic presentation of the mouse locus. Arrows depict the individual HSs, globin genes are indicated by triangles, and boxes indicate the olfactory receptor (OR) genes. (B-E) Southern blots show that, in definitive erythrocytes digestion efficiency of cross-linked chromatin depends on formaldehyde concentration and is comparable between a hypersensitive site in the LCR (B), a transcribed gene within the locus (C), a non-expressed gene within the locus (D) and a non-expressed gene on a different chromosome (Chr. 4) (E). Percentage formaldehyde cross-linking is shown at the top of each blot (– depicts genomic DNA not treated with formaldehyde), while the yield of specifically cut fragments is shown (percentages) at the bottom. Arrowheads depict partial digests and asterisks cross-hybridisation signals with other genes (see Methods).(F,G) Erythroid-specific and developmentally stable clustering of cis-regulatory elements. Relative cross-linking frequencies observed in primitive erythrocytes are shown in grey, definitive erythrocytes in black, and non-expressing brain in light grey. Grey shading indicates position and size of the analysed fragments, while black shading represents the 'fixed' fragments HS4–HS5 (F) and 5' HS –60/–62 (G). Within each graph, the highest cross-linking frequency value was set to one. The x-axis shows position in the locus. Standard-error-of-mean is indicated.

Methods and Dekker et al. 2002; Tolhuis et al. 2002). 3C-technology involves quantitative PCR-analysis of cross-linking frequencies between two given DNA restriction fragments, which gives a measure of their proximity in the nuclear space. Local chromatin configuration has no effect on digestion efficiency, implying that the assay is not biased due to preferential

restriction enzyme digestion of one site over the other (figure 6.1B-E; for other controls see Methods and (Tolhuis et al. 2002).

First, we analysed the spatial organisation of the murine β-globin locus in primitive erythroid cells present in 10.5 dpc embryonic blood, that predominantly express the embryonic εy and βh1 globin genes (Trimborn et al. 1999). We determined cross-linking frequencies for 66 pairs of *Hind*III restriction fragments, spread over ~170 kb of DNA encompassing the murine β-globin gene cluster. The 3C-measurements indicate a basic structural organisation in primitive cells very similar to that observed previously in definitive blood cells isolated from 14.5 dpc fetal liver (Tolhuis et al. 2002). This is best illustrated by comparing the locus-wide cross-linking frequencies of a restriction fragment that contains HS4–HS5 of the LCR. Two peaks of high cross-linking frequency with this genomic site stand out in primitive blood cells: one with the upstream HS–60/–62 and another with 3′ HS1 downstream of the genes (figure 6.1F). We found significantly lower cross-linking frequencies with fragments in between, suggesting that the LCR interacts with these distal HS through looping. The same interactions were observed in definitive blood cells that exclusively express the adult βmajor and βminor globin genes (Trimborn et al. 1999), where βmajor is also found in close proximity (figure 6.1F, and see below). In contrast, in non-expressing brain cells HS4–HS5 shows no peaks of interaction with distal DNA fragments, suggesting a linear conformation of the transcriptionally inactive locus (Tolhuis et al. 2002). We obtained similar results when analysing the locus-wide cross-linking frequencies of fragments carrying 5′ HS–60/–62 (figure 6.1G) and other HS (data not shown): interactions among the *cis*-regulatory elements of the β-globin locus were found to be conserved between primitive and definitive erythroid cells. We conclude that the *cis*-regulatory elements of the murine β-globin locus spatially cluster to form a transcription regulatory compartment that is conserved between primitive and definitive erythroid cells, two developmentally different types of cells that express a different subset of β-like globin genes. This core ACH includes the two HS at –60 kb, all HS of the LCR and 3′HS1.

The main differences in conformation between the two expressing cell types appear to be confined to interactions between the globin genes and the regulatory DNA elements. This is confirmed by measuring cross-linking frequencies with HS2 and HS3 of the LCR, two sites previously shown to be the most prominent transcriptional activating elements (Ellis et al. 1993; Fraser et al. 1993; Fiering et al. 1995; Guy et al. 1996; Hug et al. 1996; Carter et al. 2002). We found the embryonic globin genes εy and βh1 to interact frequently with these elements in primitive erythroid cells, whereas in definitive red blood cells interaction frequencies between these sites dropped to levels similar to those observed in the inactive brain (figure 6.2A-D). We found the opposite for the adult βmajor and βminor genes, which interacted most frequently with HS2 and HS3 in definitive erythroid cells. Cross-linking frequencies between these sites in 10.5 dpc embryonic blood were not as low as in brain, probably due the fact that βmajor and βminor are already transcriptionally active at this stage, albeit at less than 10% of the levels observed in definitive cells (Trimborn et al. 1999). Alternatively, it may merely be the result of 3′ HS1 interacting with the LCR and the adult genes being dragged along, as we previously found that the region between βminor and 3′ HS1, which is full of repetitive sequences, acts as a rigid region (Tolhuis et al. 2002). These data demonstrate that there is a developmental switch in contacts between the different globin genes and a core ACH created by regulatory elements that surround the genes in *cis*. This structural change correlates with the developmental switch in expression of the genes.

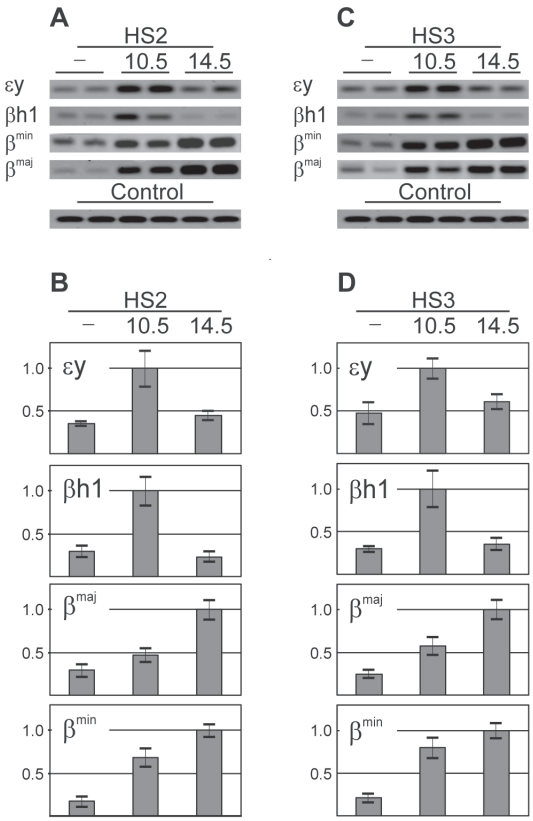


figure 6.2: A developmental switch occurs in contacts between individual β-globin genes and the core ACH of the murine β-globin locus.
(A,B) Cross-linking frequencies of HS2 and the β-globin genes were measured. An example of PCR amplified ligation products is shown on 2% agarose gel (A), as well as the quantified data of all experiments (at least five, in duplo, per primer set) (B). (C,D) Identical to A and B, but now for HS3 and the β-globin genes. Standard-error-of-mean is indicated. Non-expressing brain is depicted by '–', primitive erythrocytes by '10.5', and definitive erythrocytes by '14.5'. Control is PCR-amplified ligation product of two restriction fragments in the *Ercc3* locus (see Methods). Cross-linking frequencies shown in B and D are not corrected for PCR amplification efficiency and therefore only signals obtained with the same primer set can be compared.

To further investigate the significance of our findings, we analysed the conformation of the human β-globin locus at different stages of development. The mouse and human β-globin gene loci show a high degree of nucleotide sequence conservation, particularly at regions implicated in gene regulation (Hardison and Miller 1993; Bulger et al. 2000). We made use of transgenic mice carrying a single copy of a 185 kb PAC (figure 6.3A) spanning the human β-globin locus that displayed a normal expression pattern (Strouboulis et al. 1992; Imam et al. 2000; Patrinos et al. 2004). Although large, this PAC does not include the human equivalent of the murine 5′ HS–60/–62, which is located ~110 kb upstream of the human globin genes (Bulger et al. 2000; Farrell et al. 2000). We analysed the conformation of the transgenic human globin locus in 10.5 dpc embryonic blood, 14.5 dpc fetal liver and 14.5 dpc fetal brain, measuring almost all of the 120 site pairs that can be formed between the 16 *EcoRI* fragments that were selected for analysis. The locus-wide cross-linking frequencies of a fragment corresponding to 3′ HS1 illustrate that also the transgenic human locus forms a core ACH, consisting of the 3′ HS1 and the HS of the LCR, that is conserved in primitive and definitive erythroid cells (figure 6.3B). The structural changes we observed primarily concerned the position of the genes relative to this core ACH, correlating with transcriptional activity. Thus, the embryonic ε- and the two γ-genes most frequently interact with HS2–HS4 (figure 6.3C-E) and 3′ HS1 (figure 6.3B) in primitive erythroid cells and the adult globin gene primarily contacts the ACH in definitive cells (figure 6.3C-E). We found identical results for

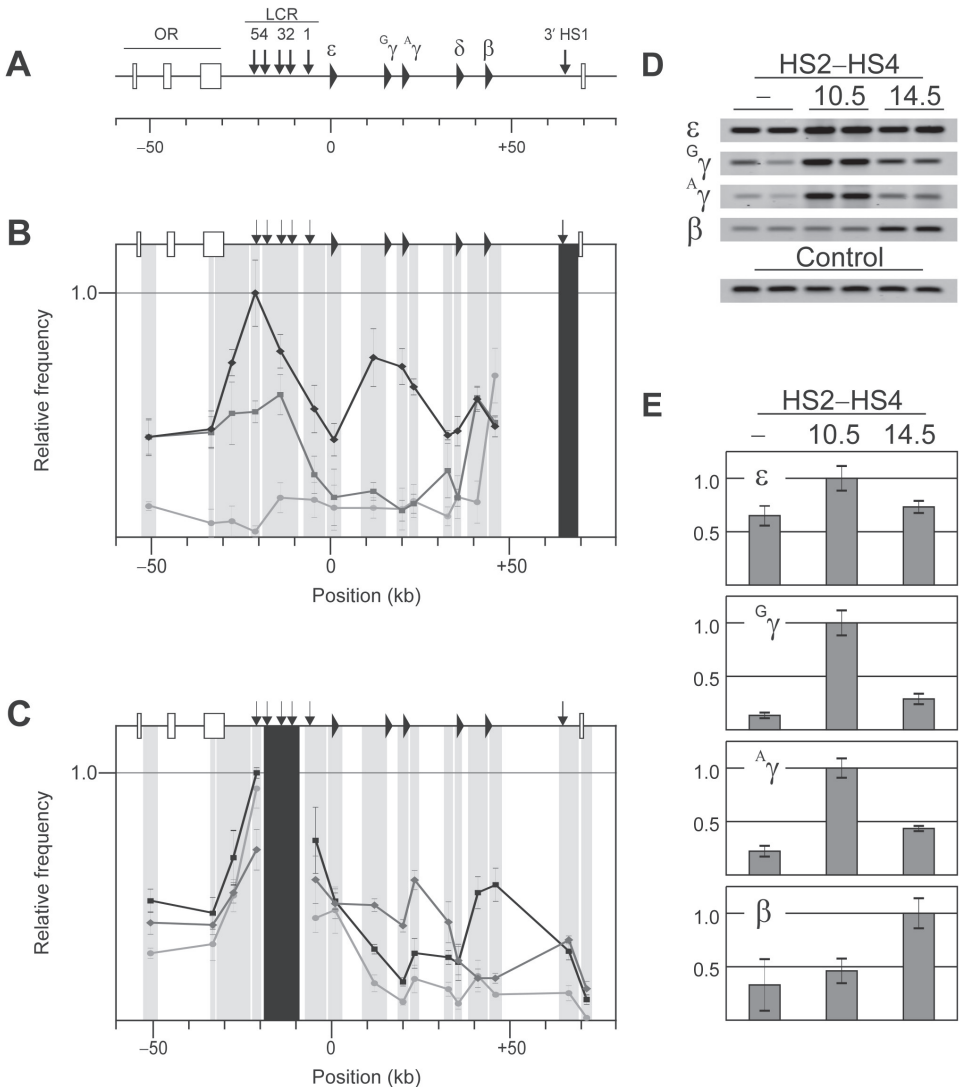


figure 6.3: Spatial organization of the human β-globin locus.
(A) Schematic presentation of the human locus. (B) Locus-wide cross-linking frequencies of a 3' HS1 fragment show erythroid cell specific clustering with the LCR throughout development. (C) Developmental switching in contacts of the LCR between the different β-globin genes as shown by locus-wide cross-linking frequencies of HS2-HS4. (D,E) The contacts between the HS2-HS4 of the LCR and individual β-globin genes alter during development in erythroid cells, as shown (D) by an example on agarose gel and (E) quantified data (at least five experiments in duplo per primer set). Controls, symbols, color patterns and numbering are as in figure 6.1 and figure 6.2.

a *Hind*III digest and for a different transgenic PAC line (data not shown). Results obtained with definitive erythroid cells isolated from adult bone marrow (Ter119^{+}) were identical to those found for 14.5 dpc fetal liver cells (data not shown). It is interesting to also note the decreased cross-linking frequency of HS5 in the definitive cells as we have recently shown

that this element has LCR blocking activity in primitive but not definitive erythroid cells (Wai et al. 2003). We conclude that the overall spatial organisation of the β-globin gene cluster is conserved from mouse to man.

Next we determined β-globin genomic site interactions in I/11 erythroid progenitor cells that are committed to, but do not yet express the β-globin genes. If exposed to physiologically relevant stimuli, I/11 cells synchronously undergo the normal *in vivo* differentiation program to mature terminally into enucleated erythrocytes (Dolznig et al. 2001; von Lindern et al. 2001). As expected, in differentiating I/11 cells that actively transcribe the adult β-like globin genes the locus adopts a spatial organisation very similar to what we observed previously in definitive erythroid cells isolated from fetal livers (Tolhuis et al. 2002) (figure 6.4). However, in uninduced proliferating I/11 cells that do not yet express the β-globin genes we observe a different structure. We found locus-wide cross-linking frequencies of a fragment corresponding to the βmajor gene to be reduced compared to those observed in erythroid cells expressing the gene (figure 6.4A). A similar reduction in locus-wide cross-linking frequencies was found for a fragment containing HS2 and for most other fragments (data not shown). However, the structure of the locus poised for transcription is clearly different from that of the inactive locus in brain cells. This structure is better resolved by looking at the locus-wide cross-linking frequencies of the restriction fragment that contains HS4–HS5 of the LCR. Two peaks of high cross-linking frequency with this fragment stand out in erythroid progenitor cells: one with 5′ HS–60/–62 and another

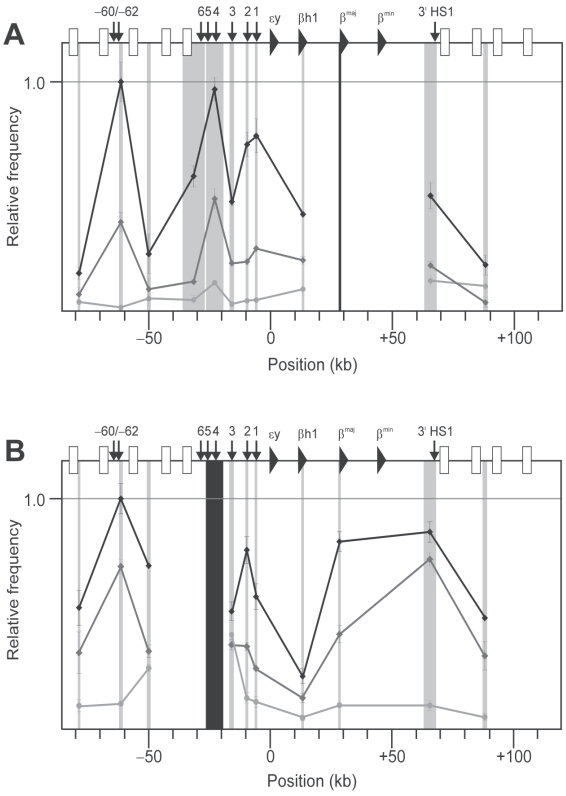


figure 6.4: Spatial organization of the murine β-globin locus in erythroid progenitors. Controls, symbols and numbering are as in figure 6.1, light grey lines represent brain, black lines proliferating I/11 erythroid progenitor cells and grey differentiated I/11 cells. (A) Locus-wide cross-linking frequencies of βmajor. (B) Locus wide cross-linking frequencies of HS4–HS5. Note that only the interactions among HS4–HS5, 5' HS–60/–62 and 3' HS1 are already fully established in non-expressing progenitor cells.

with 3′ HS1 (figure 6.4B). Interactions among these three sites occur almost as frequently in proliferating progenitors as in differentiating erythroid cells, whereas all other interactions examined between globin site pairs are strongly reduced in progenitor cells (figure 6.4A-B, and data not shown). We conclude that the β-globin locus that is poised for transcription in progenitor cells adopts a looped conformation through frequent interactions between the two distal regulatory elements at either end of the locus (5′ HS–60/–62 and 3′ HS1) and HS at the 5′ side of the LCR (HS4, 5 or 6, we currently cannot say which of these HS is responsible for direct interaction). Upon induction of differentiation, clustering with the active genes and the complete LCR is established and the β-globin genes are being expressed (figure 6.5).

Discussion

In summary, our data strongly suggest that regulatory elements surrounding the β-globin genes in *cis* create an erythroid-specific developmentally stable nuclear compartment dedicated to RNA polymerase II transcription (figure 6.5). A sub-structure is already present in erythroid progenitors that do not express globin, and it is worth noting that the three sites involved in this structure all bind CTCF (Bulger et al. 2003). It should be noted that

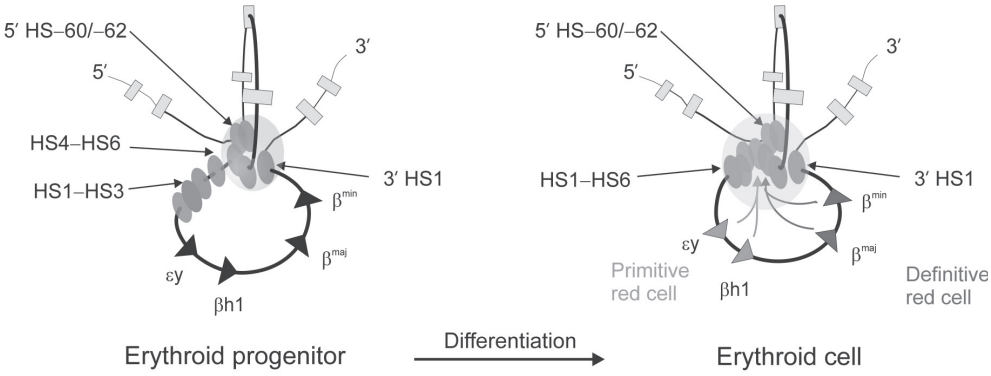


figure 6.5: Cis-regulatory elements of the β-globin locus create a nuclear compartment dedicated to RNApolymerase II transcription: the Active Chromatin Hub.
2D-presentation of 3-dimensional interactions that occur between regulatory DNA elements 130 kb apart (grey ovals) and b-globin genes (active: light or dark grey rectangles; inactive: black) in erythroid progenitors (left) and differentiated primitive and definitive erythroid cells (right). In erythroid progenitors not expressing globin, a sub-structure (sphere) is present which is formed through interactions between the upstream 5¢ HS–60/–62, the downstream 3¢ HS1 and HS at the 5¢ side of the LCR (HS4–HS6; we currently cannot say which of these HS is directly responsible for this interaction). During erythroid differentiation, the b-globin gene that gets activated and the rest of the LCR stably interact with this sub-structure to form a functional ACH (grey sphere); b-globin gene expression is activated. Clustering of binding sites for transcription factors in the ACH causes local accumulation of cognate proteins and associated positive chromatin modifiers, required to drive efficient transcription of the globin genes. The core of the ACH is erythroid-specific and developmentally stable; a developmental switch occurs in globin genes entering this nuclear compartment, as depicted by the arrows. Inactive globin and olfactory receptor genes (grey squares) loop out.

3′ HS1 and 5′ HS–60/–62 are dispensable for normal globin gene expression in transgenic mice (Strouboulis et al. 1992; Wai et al. 2003), suggesting these sites have a more general structural role not related to transcription per se.

Spatial clustering of transcription regulatory elements results in a high local concentration of DNA binding sites for cognate transcription factors, which as a consequence accumulate at the site. Efficiency of transcription is proportional to the concentration of transcription factors involved, and in agreement we found that proximity of β-globin genes to the ACH correlated with transcriptional activity. The paradigm of a chromatin-associated nuclear compartment is the nucleolus, dedicated to RNA polymerase I transcription of ribosomal RNA genes (Carmo-Fonseca 2002). No pol II-dependent gene-specific compartments have been described before, but a precedent for this was provided by electron microscopy studies showing that RNA polymerase II clusters in discrete transcription factories in the nucleus (Jackson et al. 1993; Pombo et al. 1999). The fact that the density of RNA polymerases on active β-globin and ribosomal RNA genes is much higher than on most other active genes (Wijgerde et al. 1995; Cook 1999) suggests that such nuclear compartments formed by numerous chromatin-associated regulatory elements primarily function to increase the efficiency of transcription.

Experimental procedures

Chromosome Conformation Capture (3C)

We performed isolation and formaldehyde fixation of primary cells, restriction enzyme digestion of cross-linked DNA in the nucleus, intramolecular ligation, reversal of cross-links, PCR-analysis of ligation products and calculation of relative cross-linking frequencies as described before (Dekker et al. 2002; Tolhuis et al. 2002), with some modifications. Prior to fixation, we forced cells obtained from embryonic blood (10.5dpc embryos), fetal liver and fetal brain (both 14.5dpc embryos; $4x10^7$ cells per tissue) through a cell-strainer cap (Falcon #352235) to obtain a homogeneous single cell suspension. To correct for differences in quality and quantity of template, we normalized ligation frequencies between globin site pairs to those detected between two restriction fragments (with the sites analysed 8.3 kb apart) in the *Ercc3* locus (instead of the previously used *Calr* locus; Tolhuis et al. 2002). *Ercc3* encodes a subunit of the basal transcription factor TFIIH, and we assumed that expression levels and spatial conformation of this gene are similar in all analysed tissues. To be able to compare signal intensities obtained with different primer sets in a quantitative manner, we included a control template containing all possible ligation products in equimolar amounts to correct for the PCR amplification efficiency of each set. For this purpose we used BAC and PAC clones spanning the complete loci (instead of the previously used PCR fragments that span the restriction sites of interest (Tolhuis et al. 2002)). For the mouse β-globin locus we used a 214 kb BAC (#RP23-370E12, Ensembl Genome Browser), and for the human β-globin locus we used a 185 kb PAC (Imam et al. 2000). In addition, we used a 60-70 kb PAC containing the mouse *Ercc3* locus (PAC Clone #443-C18, MRC gene service). We mixed either the mouse globin BAC or the human globin PAC with the *Ercc3* PAC at equimolar amounts. Subsequently, we digested and ligated the mixes as described (Tolhuis et al. 2002). We could not obtain control PCR products with primers designed to analyse fragments containing εy and βminor, due to polymorphisms in the BAC clone #RP23-370E12. As a consequence, these fragments were not included in the locus-wide cross-linking frequency analysis (see figure 6.1). We performed all animal experiments according to institutional and national guidelines.

Southern blotting

We treated fetal liver cells (14.5dpc embryos) as above (with indicated formaldehyde concentrations), but we omitted ligation and analysed 10 µg of purified DNA by southern blotting. We used the following probes: βh1, a 255 bp *Hinf*I fragment, hybridises to a 2.7 kb *Hind*III βh1fragment and βh0 (5.5 kb) and βh0 (6.4 kb) psuedogene fragments; βmajor, a 700 bp *Hind*III/*Nco*I fragment, hybridises to a 1.0 kb *Hind*III fragment and a βminor (8.6 kb) fragment; HS3, 300 bp PCR fragment, hybridises to a 2.0 kb *Hind*III fragment; *Pou3f1*, 100 bp PCR fragment, hybridises to a 4.0 kb *Hind*III fragment.

Cell culture

We cultured I/11 cells as described previously (Dolznig et al. 2001; von Lindern et al. 2001). Briefly, we maintained proliferating I/11 cells in StemPro-34™ containing 2 units/ml human recombinant erythropoietin, 100 ng/ml murine recombinant SCF, 10^{-6} M dexamethasone and 40 ng/ml insulin-like growth factor. We expanded cells by daily partial medium changes and addition of fresh factors, keeping cell density between 1.5–4 x10^6 cells/ml. For induction of differentiation, we removed continuously proliferating I/11 cells from the culture, washed them twice in PBS, and seeded them at 2–3 x10^6 cells/ml in differentiation medium containing 10 units/ml Epo, 4 x10^{-4} IE/ml Insulin, the Dex-antagonist ZK-112993 (3 x10^{-6} M), and 1 mg/ml iron-saturated human transferrin. We maintained differentiating erythroblasts at densities between 2-6 x10^6 cells/ml. For 3C-analysis of differentiating I/11 cells, we fixed cells with formaldehyde 40 hours after induction and processed them as described above.

URLs

Ensembl Genome Browser	http://www.ensembl.org
MRC gene service	http://www.hgmp.mrc.ac.uk

Chapter 7

The active spatial organisation of the β-globin locus requires the transcription factor EKLF

Based on: Genes & Development 18, 2485-2490, (*2004*).

The active spatial organisation of the β-globin locus requires the transcription factor EKLF

Roy Drissen[1], Robert-Jan Palstra[1], Nynke Gillemans, Erik Splinter, Frank Grosveld, Sjaak Philipsen[2] and Wouter de Laat

Department of Cell Biology and Genetics Faculty of Medicine Erasmus University, PO Box 1738, 3000DR Rotterdam, The Netherlands, Fax: +31-10-4089468
[1] These authors contributed equally; [2] Corresponding author: j.philipsen@erasmusmc.nl

Summary

Three-dimensional organisation of a gene locus is important for its regulation, as recently demonstrated for the β-globin locus. When actively expressed, the *cis*-regulatory elements of the β-globin locus are in proximity in the nuclear space, forming a compartment termed the Active Chromatin Hub (ACH). However, it is unknown which proteins are involved in ACH formation. Here we show that EKLF, an erythroid transcription factor required for adult β-globin gene transcription, is also required for ACH formation. We conclude that transcription factors can play an essential role in the three-dimensional organisation of gene loci.

Introduction

The mouse β-globin locus contains multiple β-like globin genes, arranged from 5' to 3' in order of their developmental expression (figure 7.1A). The adult-type βmajor-gene is transcribed at a very low level during primitive erythropoiesis in the embryonic yolk sac, but becomes expressed at high levels around day 11 of gestation (E11) when definitive erythropoiesis commences in the fetal liver (Trimborn et al. 1999). The β-globin locus control region (LCR) is essential for efficient globin transcription (Grosveld et al. 1987; Bender et al. 2000a). It consists of a series of DNaseI hypersensitive sites (HS) located ~50 kb upstream of the βmajor promoter (figure 7.1A). We have shown that the β-globin locus forms an Active Chromatin Hub (ACH) in erythroid cells (Tolhuis et al. 2002). The ACH is a nuclear compartment dedicated to RNA polymerase II transcription, formed by the *cis*-regulatory elements of the β-globin locus with the intervening DNA looping out. The ACH consists of the HS of the LCR, two HS located ~60 kb upstream of the embryonic εy-globin gene (5'HS-62/-60) and 3'HS1 downstream of the genes. In addition, the actively expressed globin genes are part of the ACH (Carter et al. 2002; Tolhuis et al. 2002). In erythroid precursors which do not express the globin genes yet, a substructure of the ACH, called a chromatin hub (CH) (Patrinos et al. 2004) is found, which excludes the genes and the HS at the 3' site of the LCR (Palstra et al. 2003).

Expression of the βmajor-gene requires the presence of the erythroid Krüppel-like transcription factor EKLF, the erythroid-specific member of the Sp/XKLF-family (Miller

and Bieker 1993). EKLF$^{-/-}$ mice die of anaemia around E14, because of a deficit in β-globin expression (Nuez et al. 1995; Perkins et al. 1995). The β-globin locus contains a number of EKLF binding sites, in particular in the LCR and the βmajor-globin promoter (Perkins 1999; Bieker 2001). Since βmajor-globin expression depends on the presence of EKLF, we were interested to determine whether EKLF is involved in the formation of the ACH.

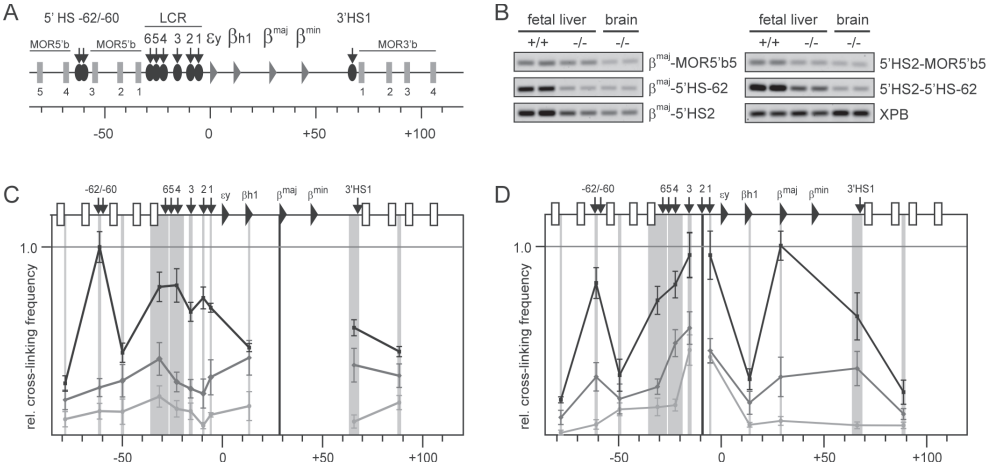


figure 7.1: EKLF influences the spatial organisation of the β-globin locus.
(A) Schematic presentation of the mouse β-globin locus. Globin genes are indicated by triangles. Olfactory receptor genes (MOR5'b and MOR3'b) are indicated by grey rectangles and numbered. DNaseI HS are shown as black ovals with arrows. The scale is in kb. (B) Examples of PCR-amplified ligation products run on a 2% agarose gel. Primer combinations are shown on the right. *XPB* is used to standardise the amount of template (Palstra et al. 2003). +/+ = wildtype; -/- = EKLF knockout. (C-D) Locus-wide relative crosslinking frequencies in E12.5 fetal livers. Results obtained with wild type livers are shown in black; EKLF-/- livers in grey, non-expressing brains in light grey. The x-axis shows position in the locus. Grey shading indicates the positions and sizes of the *Hind*III fragments containing primers used in the PCR analysis. Black shading represents the position of the fragment containing the "fixed" primer in the *Hind*III fragment of the βmajor-gene (C) or 5'HS2 (D). Within each graph, the highest crosslinking frequency value is set to 1. Error bars indicate the standard-error-of-mean.

Results

We used chromatin conformation capture (3C) technology (Dekker et al. 2002) to investigate the three-dimensional conformation of the mouse β-globin locus in the absence of EKLF. Cells from E12.5 EKLF$^{-/-}$ and wild type fetal livers were crosslinked with formaldehyde, followed by restriction enzyme digestion of the DNA. The samples were ligated under conditions that favour the ligation of DNA fragments that are physically connected through the crosslinks. Quantitative PCR across the junctions is used to determine the relative crosslinking frequencies between restriction fragments in the locus. This provides an indication of the nuclear proximity of DNA fragments *in vivo* (Dekker et al. 2002; Tolhuis et al. 2002; Palstra et al. 2003). Crosslinking frequencies were determined for a total of 66 junctions that can be formed between 12 selected *Hind*III fragments spread over ~170 kb of DNA encompassing the β-globin gene cluster (figure 7.1).

Examples of quantitative PCR reactions with some of the primer combinations are shown in figure 7.1B. An overview of the locus-wide crosslinking frequencies of a restriction fragment that contains the βmajor promoter is shown in figure 7.1C. The brain serves as a non-expressing control tissue (light grey curve), in which the β-globin locus appears to adopt a linear conformation (Tolhuis et al. 2002). In wild type E12.5 fetal liver cells, high crosslinking frequencies are found with the LCR and 5'HS-62, indicating their proximity to the βmajor promoter *in vivo* (black curve). In the absence of EKLF however, these crosslinking frequencies are much lower and no interaction with a distal site stands out clearly (grey curve), showing that the βmajor promoter does not participate stably in a spatial clustering of chromatin. A comparable pattern is observed with locus-wide crosslinking frequencies of a fragment containing 5'HS2 (figure 7.1D). Together with 5'HS3, 5'HS2 is the most prominent transcriptional activating element of the LCR (Ellis et al. 1993; Fraser et al. 1993; Fiering et al. 1995; Ellis et al. 1996; Hug et al. 1996). Interactions with 5'HS-62, βmajor, 3'HS1, and the other HS of the LCR are strongly reduced in the absence of EKLF, indicating that 5'HS2 requires the presence of EKLF to participate in the ACH.

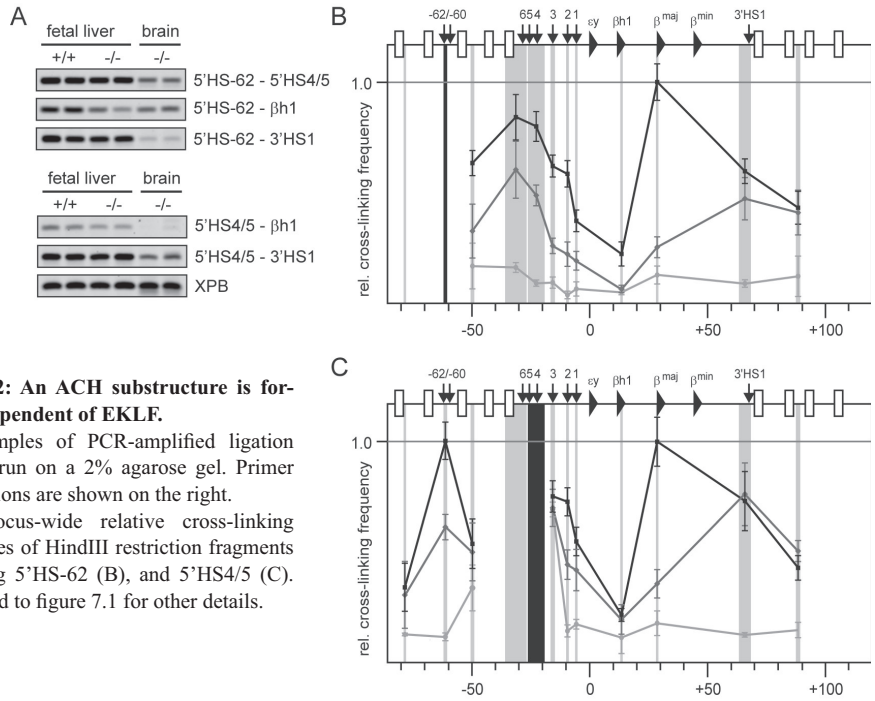


figure 7.2: An ACH substructure is formed independent of EKLF.
(A) Examples of PCR-amplified ligation products run on a 2% agarose gel. Primer combinations are shown on the right.
(B,C) Locus-wide relative cross-linking frequencies of HindIII restriction fragments containing 5'HS-62 (B), and 5'HS4/5 (C). See legend to figure 7.1 for other details.

The results shown in figure 7.1 demonstrate that the complete ACH is not formed in the absence of EKLF. However, the observed crosslinking frequencies in EKLF$^{-/-}$ fetal liver cells are still higher than those found in non-expressing brain cells indicating a different, non-linear, structure. To investigate this, we compared the locus-wide crosslinking frequencies of restriction fragments containing 5'HS-62 and 5'HS4/5 (figure 7.2). These sites participate in the CH present in erythroid progenitor cells before the globin genes are transcribed (Palstra et al. 2003). Examples of quantitative PCR reactions with some of the primer combinations

are shown in figure 7.2A. The graphs in figure 7.2B show that in wild type fetal liver cells, 5'HS-62 is in proximity to the LCR, βmajor and 3'HS1. In EKLF$^{-/-}$ fetal liver cells, 5'HS-62 interactions with the HS at the 5' side of the LCR and with the distal 3'HS1 still stand out, whereas all other crosslinking frequencies are strongly reduced. This indicates the presence of a globin CH, containing 5'HS-62/-60, the HS at the 5' side of the LCR, and 3'HS1. The same structure is apparent when analysing locus-wide crosslinking frequencies of a restriction fragment containing 5'HS4/5 at the 5' side of the LCR (figure 7.2C).

There are remarkable similarities between the structure of the β-globin locus in EKLF$^{-/-}$ fetal liver cells and that observed in I/11 erythroid progenitor cells which do not yet express globin (Palstra et al. 2003) (figure 7.3). This suggests that EKLF is required for progression from the chromatin hub present in erythroid precursors to a fully active ACH.

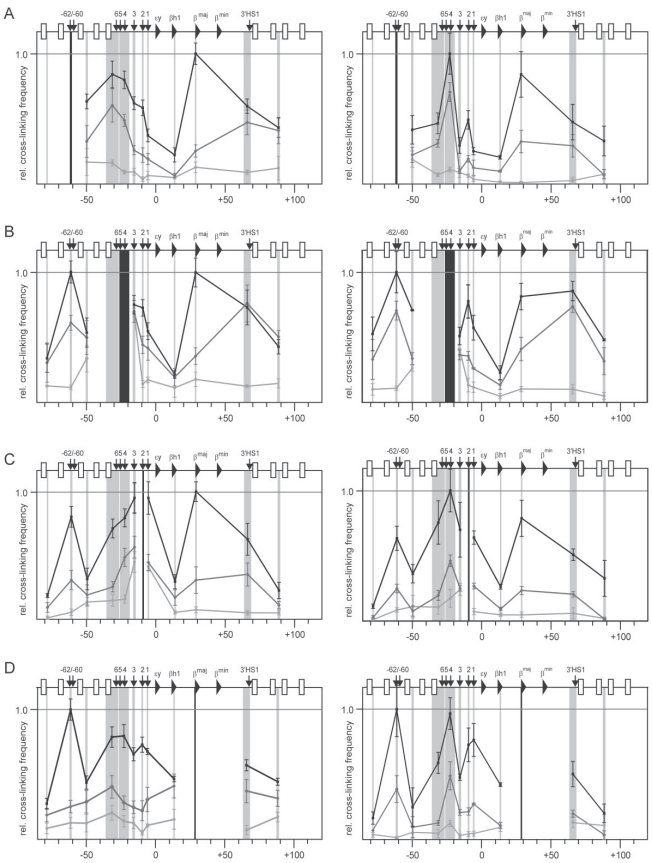


figure 7.3: The spatial organisation of the β-globin locus in EKLF-/- fetal liver cells is similar to that observed in erythroid progenitor cells.
(A,B,C,D) Comparison of locus-wide crosslinking frequencies of the *Hind*III restriction fragment containing 5'HS-62 (A), 5'HS4/5 (B), 5'HS2 (C), and βmajor (D), in fetal livers (graphs on the left) and I/11 erythroid progenitor cells (Palstra et al. 2003) (graphs on the right). Grey curves in graphs on the right: proliferating I/11 erythroid progenitor cells, not expressing globins. Black curves in graphs on the right: I/11 cells after induction of globin expression (Dolznig et al. 2001). Other details are as for figure 7.1.

To investigate if this β-globin structure in EKLF *null* cells is a direct consequence of EKLF-deficiency or caused by a general differentiation failure, we analysed expression of another erythroid-specific, but EKLF-independent, globin locus, the α-globin gene locus. Consistent with previous observations, primary transcript *in situ* hybridisation experiments (Van de Corput and Grosveld 2001) show that α-globin expressing cells are abundantly present in the EKLF$^{-/-}$ fetal liver, demonstrating that in the absence of EKLF cells are progressing to the stage of active globin expression (figure 7.4A) (Nuez et al. 1995; Perkins et al. 1995; Wijgerde et al. 1996; Lim et al. 1997). We do observe that EKLF$^{-/-}$ fetal livers contain approximately 20% less α-globin expressing cells than wild type fetal livers (~55% *versus* 70% of the total number of cells in the fetal liver). Whereas this may explain the small reduction in crosslinking frequencies observed between some of the β-globin elements, it cannot account for the strongly reduced locus-wide crosslinking frequencies seen with, for example, βmajor and 5'HS2 (figure 7.1C,D). We conclude that the dramatically altered chromatin organisation of the β-globin locus in EKLF *null* erythroid cells is not due to a general differentiation problem.

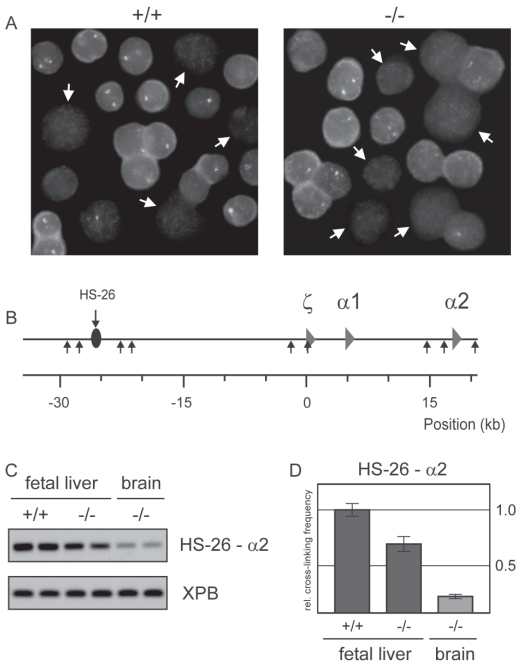


figure 7.4: HS-26-promoter interactions in the α-globin locus are not affected by EKLF.
(A) In situ hybridisation of E12.5 fetal liver cells of wildtype and EKLF-/- foetuses, detecting α-globin mRNA (light halos) and primary transcripts (bright dots). DAPI staining (dark halos) is used to show nuclear DNA. White arrows indicate cells that were scored negative for α-globin expression. (B) Schematic drawing of the mouse α-globin locus. The black oval with arrow depicts the position of the HS-26 distal regulatory element. The α-like globin genes are indicated by grey triangles. Small arrows = *Hind*III restriction sites. (C) Example of PCR-amplified ligation products of *Hind*III restriction fragments containing HS-26 and α2 in E12.5 day fetal liver and brain cells of wildtype and EKLF-/- foetuses. The *XPB* PCR product is used as template control. (D) Quantified data of PCR-amplified ligation products. Dark bars = fetal liver; light bar = brain. Error bars indicate standard-error-of-mean. The crosslinking frequency in wildtype fetal liver cells is set to 1.

To substantiate the specificity of the changes in the three-dimensional structure of the β-globin locus in the absence of EKLF we investigated interactions between the promoter and remote regulatory element of the erythroid-specific, EKLF-independent, α-globin locus. The mouse α-globin locus has two active genes in the fetal liver, α1 and α2, and contains a HS 26 kb upstream of the ζ-globin promoter that is similar to the human α-globin enhancer HS-40 (figure 7.4B) (Flint et al. 2001). It is likely that, analogous to the LCR, this element will interact with the α-like globin promoters to enhance expression. The crosslinking frequencies of the restriction fragments containing HS-26 and α2-globin are shown in figure 7.4C and 7.4D. In wild type and EKLF$^{-/-}$ fetal liver cells, the crosslinking frequencies are clearly higher than those observed in non-expressing brain tissue, indicating that HS-26 and the α2-globin gene are in close proximity in both types of erythroid cells. The slightly reduced interaction frequencies observed in EKLF knockout compared to wild-type fetal liver can be explained by the 20% reduction of α-globin expressing cells (see above). We conclude that major alterations in spatial organisation are restricted to the EKLF-dependent β-globin locus.

To further investigate if changes in the spatial organisation of the β-globin locus are a direct effect of the activity of EKLF, we wished to induce EKLF activation and simultanously prevent it from activating secondary pathways. For this, we used a fusion between EKLF and a modified estrogen receptor ligand binding domain (EKLF-lbd protein), that can be activated by 4-hydroxy-tamoxifen (4-OHT) (Littlewood et al. 1995). We wanted to test if, in an EKLF *null* background, activated EKLF-lbd protein restores ACH formation in the presence of the *de novo* protein synthesis inhibitor cycloheximide (CHX). In such a set up, genes activated by EKLF cannot be translated into protein and therefore any structural changes would have to be attributed to EKLF acting directly on the β-globin locus.

Transgenic mice carrying an expression construct of an EKLF-lbd fusion protein were generated. To ensure expression of the fusion protein in EKLF *null* erythroid cells, we used the erythroid-specific pEV3 expression vector (Needham et al. 1992) and replaced the β-globin promoter by the α-globin promoter. Western blot analysis demonstrates the presence of the HA-tagged EKLF-lbd fusion protein (figure 7.5A). We have previously shown that an EKLF-pEV3 transgene rescues the EKLF *null* mutation (Tewari et al. 1998). To test if uninduced EKLF-lbd fusion protein is inactive, we crossed the EKLF-lbd transgenics with the EKLF knockout mice. No EKLF *null*::EKLF-lbd transgene pups were born. When we dissected the foetuses resulting from this cross at E12.5, we found that the EKLF *null*::EKLF-lbd transgenic foetuses were indistinguishable from EKLF *null* foetuses, e.g. displaying signs of severe anemia and having very pale fetal livers (data not shown). We conclude that the EKLF-lbd fusion protein is inactive and does not rescue the EKLF *null* mutation.

To test the ability of activated EKLF-lbd fusion protein to rescue β-globin gene transcription, we cultured EKLF *null*::EKLF-lbd fetal liver cells in the presence of 4-OHT (figure 7.5B). After 16 hours of culturing, a subset of the cells was used to check for the activation of β-globin gene expression. Real-time RT-PCR analysis of steady-state mRNA levels shows that the β-globin gene is activated in EKLF *null*::EKLF-lbd cells in the presence of 4-OHT (figure 7.5C). The amount of β-globin transcripts in the tamoxifen-rescued cells is much lower than in wild type cells, which is is not surprising since the former cells just start to accumulate β-globin mRNA levels. We conclude that the EKLF-lbd fusion protein can be induced with 4-OHT to activate β-globin gene expression. Moreover, β-globin gene activation by 4-OHT-induced EKLF-lbd also occurs in the presence of CHX (figure 7.5C).

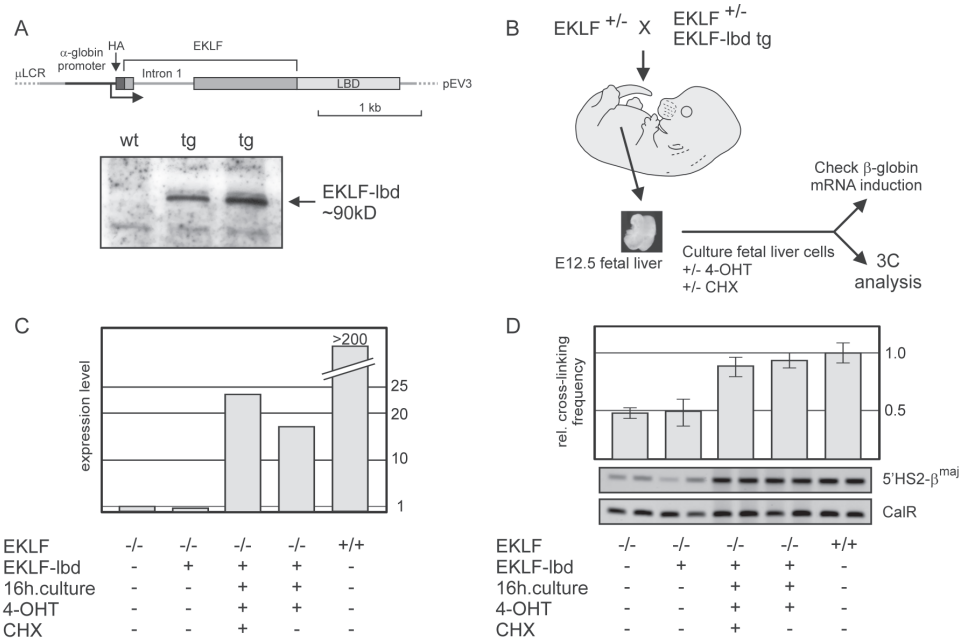


figure 7.5: EKLF is directly involved in the spatial organisation of the β-globin locus.
(A) Schematic drawing of the EKLF-lbd expression construct used to generate transgenic mice. The Western blot shows expression of the EKLF-lbd fusion protein in the fetal livers of transgenic mice detected by an antibody recognizing the HA tag. (B) Flow chart of the experimental design. Fetal livers are isolated from E12.5 control- and EKLF null::EKLF-lbd tg foetuses, disrupted and the erythroid cells are cultured in the presence of 4-OHT with or without CHX for 16 hours. Cells are then harvested, crosslinked with formaldehyde and subjected to 3C analysis. From a portion of the cells, RNA is isolated to check β-globin gene expression. (C) Expression of β-globin analysed by real-time RT-PCR. Expression of Hprt was used to standardize the β-globin expression levels. Representative experiment is shown. (D) 3C analysis of the interactions between 5'HS2 and the β-globin promoter. Representative examples of the PCR reactions are shown. Error bars indicate the standard-error-of-mean. Calreticulin was used as template control.

The remaining cells were subjected to 3C analysis using a procedure modified for use with small numbers of cells. Since the amount of material was limiting, we focussed on the analysis of interactions between 5'HS2, one of the most prominent activating elements of the LCR, and the promoter of the βmajor gene (Carter et al. 2002; Tolhuis et al. 2002). In untreated EKLF *null* fetal liver cells, we found similarly low crosslinking frequencies between 5'HS2 and βmajor regardless of the presence of the (uninduced) EKLF-lbd protein (figure 7.5D). In contrast, this interaction is restored in EKLF *null*::EKLF-lbd cells after culturing for 16 hours in the presence of 4-OHT. Importantly, the same effect is also observed when CHX and 4-OHT are present simultaneously (figure 7.5D). These data indicate that ACH interactions are restored in EKLF *null*::EKLF-lbd cells when the EKLF-lbd fusion protein is activated by 4-OHT. Since this also occurs when protein synthesis is inhibited through the addition of CHX, we conclude that EKLF is directly involved in the completion of ACH formation.

Discussion

In conclusion, our data show that a chromatin hub is formed independent of EKLF during erythropoiesis, consisting of the 5'HS-62/-60, the HS at the 5' side of the LCR, and 3'HS1. EKLF is required for the progression to, or stabilisation of, a fully functional ACH, which includes the remaining HS of the LCR and the actively transcribed βmajor globin gene (figure 7.6). The βminor gene which is also expressed in definitive erythroid cells is known to alternate with the βmajor gene in the ACH in a dynamic "flip-flop" mechanism (Wijgerde et al. 1995; Trimborn et al. 1999). The EKLF-independent chromatin hub is structurally similar to that present in erythroid precursor cells, which were previously found to already contain EKLF mRNA (Dolznig et al. 2001) and protein (data not shown). This suggests that modifications of the EKLF protein or other protein factors are required to collaborate with EKLF in organizing a fully active β-globin ACH.

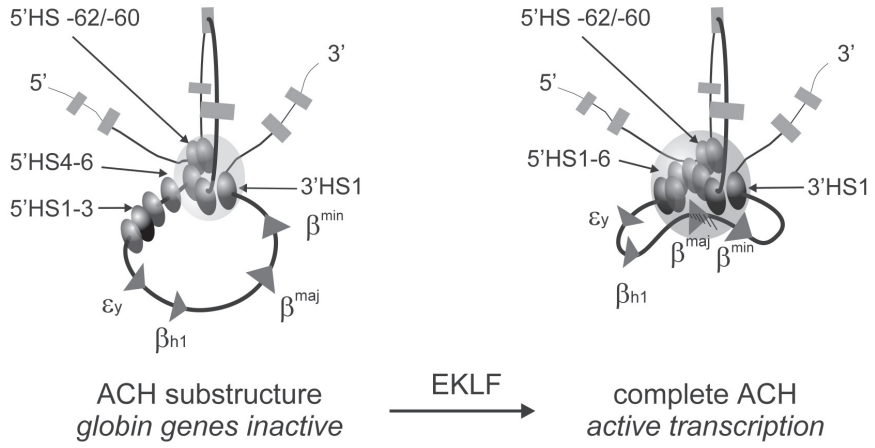


figure 7.6: The formation of the complete ACH requires the presence of EKLF.
A two-dimensional representation of the proposed three-dimensional structure of the ACH is shown. The ACH is a nuclear compartment dedicated to RNA polymerase II transcription, formed by cis-regulatory elements of the β-globin locus (Palstra et al. 2003). In erythroid cells, a substructure of the ACH, consisting of 5'HS-62/-60, 3'HS1 and HS at the 5' side of the LCR, is formed independently of EKLF. Progression of this substructure to a fully functional ACH, including the HS at the 3' side of the LCR and the active β-globin gene, is dependent on the presence of EKLF. Grey sphere on the left: CH structure. Grey sphere on the right: ACH. RNA transcripts are indicated as curved lines. See legend to figure 7.1a for other details.

Recent work has shown that deletion of the promoter of the adult β-globin gene in the human β-globin locus mildly affects ACH formation, suggesting that in addition to the β-globin promoter other *cis*-regulatory elements in the human β-globin locus are involved in these interactions (Patrinos et al. 2004). EKLF binding sites are also present in the LCR, in particular in 5'HS3, and in the 3' enhancer of the β-globin gene (Wall et al. 1988; Gillemans et al. 1998). Together these data suggest that the EKLF-dependent interactions of the adult β-globin genes with the ACH are not a mere consequence of the activation of gene transcription, but occur in a step preceding transcriptional activation and involve multiple *cis*-regulatory elements.

It is also interesting to note that in the EKLF knockout absence of spatial interactions coincides with loss of chromatin accessibility at 5'HS3 and the βmajor-promoter (Wijgerde et al. 1996; De Laat and Grosveld 2003). We conclude that EKLF is necessary for hypersensitive formation and the participation of the LCR and the β-globin promoter in the ACH, probably through interactions with a SWI/SNF-related chromatin remodelling complex (Armstrong et al. 1998). Thus, EKLF is the first example of a transcription factor that is required for the proper spatial organisation of a mammalian gene locus.

Experimental procedures

Chromosome Conformation Capture

EKLF$^{+/-}$ mice (Nuez et al. 1995) were crossed and E12.5 fetal livers and brains were isolated. 3C analysis was performed as described (Splinter et al. 2004), with minor adjustments. Individual liver and brain samples were subjected to formaldehyde crosslinking. *Hind*III restriction enzyme digestion of crosslinked DNA, intra-molecular ligation, reversal of crosslinks, PCR analysis of ligation products and calculation of relative crosslinking frequencies was done with 15 pooled wildtype fetal livers, 15 EKLF$^{-/-}$ fetal livers and cells of 3 pooled EKLF$^{-/-}$ brains. Two independent samples were prepared for the analysis. Each PCR reaction was performed in duplicate and repeated at least 3 times.

α-globin

HS-26 - α2 promoter crosslinking frequencies were determined with the DNA samples described above and primers recognising the *Hind*III restriction fragment containing HS-26 (5'-GAATCTCCATCTCCAAGGG-3') and the α2 promoter (5'-AAGAGGTGCAGGTGTATTACTG-3'). *In situ* hybridisation of E12.5 fetal liver cells was performed as described before (Van de Corput and Grosveld 2001). Cells were scored positive if α-globin mRNA, primary transcript, or both, was detected. >300 cells were counted to determine the percentage of α-globin-positive cells in each sample.

Generation of EKLF-lbd transgenic mice

A DNA fragment containing EKLF cDNA and the first intron was linked in frame with the HA tag sequence at the 5' side and the lbd coding sequence at the 3' side. This construct was cloned into the pEV3 vector (Needham et al. 1992) and the β-promoter was replaced by a fragment containing the α-globin promoter. The vector was linearised by *Aat*II and transgenic mice were generated as described (Kollias et al. 1986).

Culture of primary fetal liver cells

Livers were isolated from E12.5 control- and EKLF null::EKLF-lbd tg foetuses. The genotype of the foetuses was confirmed by PCR. Single cell suspensions of individual fetal livers were cultured for 16h in StemPro-34™ containing 1% BSA, 1% glutamine and 10 units/ml epo, but without serum supplement. The EKLF-lbd was activated by supplementing the medium with either 250nM 4-hydroxy-tamoxifen (4-OHT) alone or with 250nM 4-OHT and 20 μg/ml cycloheximide (CHX). After 16h of culture, cells were harvested and a small aliquot was taken for RNA isolation. The 16h period was chosen because it allowed detection of 4-OHT-induced β-globin gene transcription without CHX causing toxic effects. Fixation of the remainder of the cells with formaldehyde and subsequent isolation of nuclei was performed as described before (Tolhuis et al. 2002).

Preparation of cDNA and Real-time PCR

RNA was isolated using Trizol, according to the manufacturers guidelines (Invitrogen). The Super-script™ reverse transcriptase Kit (Invitrogen) was used for preparation of oligo-dT primed cDNA. Expression levels were determined on the Bio-Rad I-Cycler using the qPCR™ Core kit for Sybr

Green 1 (Eurogentec). Expression levels of Hprt were used for normalization of β-globin expression levels.

Primers used were as follows: Hprt-s, AGCCTAAGATGAGCGCAAGT; Hprt-as, ATGGCCACAGGACTAGAACA; βmajor-s, ATGCCAAAGTGAAGGCCCAT; βmajor-as, CCCAGCACAATCACGATCAT.

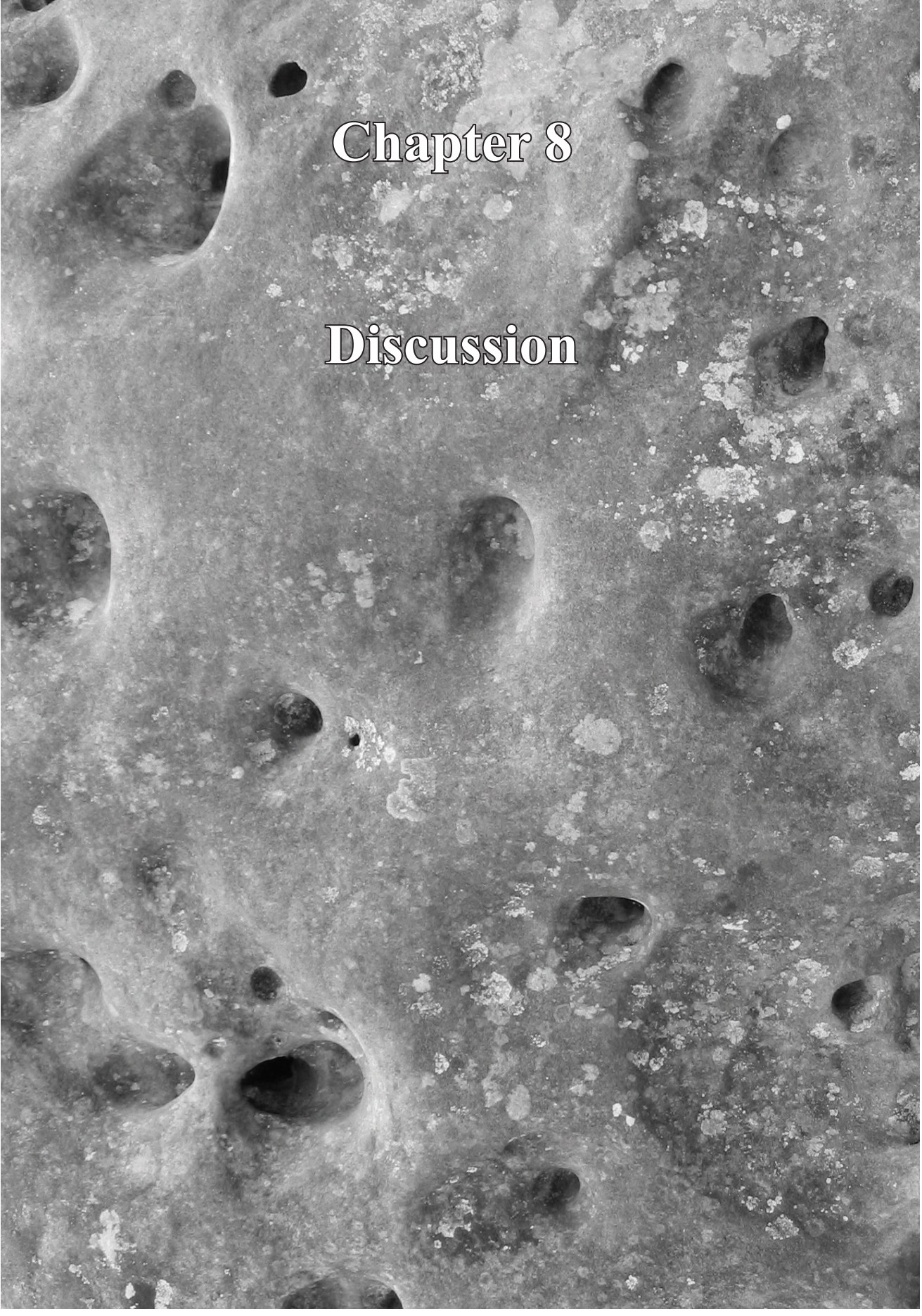
Preparation of 3C templates

For the limiting number of cells (approximately 1.10^6) obtained from the individual EKLF null::EKLF-lbd tg fetal livers, we adapted the previously described protocol (Tolhuis et al. 2002). Cross-linked nuclei of E12.5 fetal livers were re-suspended in 50 µl digestion buffer containing 0.1% SDS and incubated for 1 hour at 37^0C with agitation, Triton X-100 was added to 2.6% and the nuclei were further incubated for 1 hour at 37^0C.

The cross-linked chromatin was digested overnight at 37^0C with 10 units of *Hind*III. The restriction enzyme was heat-inactivated (25 minutes at 65^0C). After addition of 200 µl of 1.25x ligase buffer and 40 U of T4 ligase the chromatin was ligated for 4.5 hours at 16^0C followed by 30 minutes at room temperature. Proteinase K was added and samples were incubated overnight at 65^0C to reverse the cross-links. The following day samples were incubated for 30 minutes with RNAse and the DNA was purified by phenol extraction and ethanol precipitation using glycogen as a carrier. Locus wide cross-linking frequencies of wild type fetal livers treated with this adapted protocol were similar as those found previously (data not shown). PCR analysis of the ligation products was performed as described before (Tolhuis et al. 2002; Palstra et al. 2003).

Acknowledgements

This work was supported by the Dutch Organisation for Scientific Research NWO (FG, WdL and SP), the EU (FG) and the Center for Biomedical Genetics (CBG). We thank Ton de Wit for micro-injection.



Chapter 8

Discussion

Discussion

Transcriptional activation in higher eukaryotes frequently involves the long-range action of a number of regulatory DNA elements. One of the main questions in transcriptional regulation is how the *cis*-regulatory elements communicate over large distances with the promoter of a gene. There has been a lively debate about whether the mode of action of distant *cis*-regulatory elements is via a non-contact mechanism (linking, tracking) or via a contact mechanism (looping).

The spatial organisation of the mouse β-globin gene locus during development and erythroid differentiation was investigated in order to study how these regulatory elements communicate with the genes. The mouse β-globin locus contains multiple β-like globin genes, arranged from 5' to 3' in order of their developmental expression, the locus control region (LCR), which consists of a series of DNaseI hypersensitive sites (HS), ~50 kb upstream of the β-major promoter, and several distal upstream and downstream hypersensitive sites with unknown function.

The spatial organisation of a 200 kb region spanning the mouse β-globin locus was analysed, at an unprecedented level of resolution, in expressing erythroid and non-expressing brain tissue with chromosome conformation capture (3C) technology. 3C-technology involves quantitative PCR-analysis of cross-linking frequencies between two given DNA restriction fragments, which gives a measure of their proximity in the nuclear space. Originally developed to analyse the conformation of chromosomes in yeast (Dekker et al. 2002), this procedure was optimised to study the conformation of chromosomal regions in (much more complex) mammalian cells (Tolhuis et al. 2002).

In brain, the β-globin cluster adopts a seemingly linear conformation. However, in erythroid cells, the DNaseI hypersensitive sites (HSs) of the β-globin LCR, located 40-60kb away from the active genes, come in close spatial proximity with these genes. The intervening chromatin with inactive globin genes loops out. Thus the work presented here shows proximity between the β-globin LCR, a collection of distant *cis*-regulatory elements, and the β-globin promoter which strongly suggests that these elements physically interact to control gene activity, as predicted by the original looping models (Tolhuis et al. 2002). Surprisingly, two distant hypersensitive regions, being 130 kb apart from each other, also participate in these interactions. The spatial clustering of transcriptional regulatory elements is referred to as an Active Chromatin Hub (ACH) (Tolhuis et al. 2002).

Next, the 3D-structure of the β-globin locus in primitive erythroid cells was analysed, a developmental stage when the embryonic globin genes are active and the adult genes are silent. A core ACH containing the β-globin regulatory elements is developmentally conserved, while globin genes switch their interaction with this cluster during development, and this correlates with the switch in their transcriptional activity. The human β-globin locus, as a transgene in mice, adopts a similar configuration as was observed for the mouse β-globin locus (Palstra et al. 2003).

In erythroid progenitors that are committed to, but do not yet express globin, only a subset of regulatory sites cluster and form a structure called a Chromatin Hub (CH). Upon erythroid differentiation, a fully functional ACH is formed, containing all regulatory elements and the gene that becomes activated (Palstra et al. 2003).

Finally, the role of transcription factors in β-globin ACH formation was addressed. EKLF is

an erythroid-specific transcription factor essential for expression of the adult β-like genes. In fetal livers obtained from EKLF knockout mice a conformation of the locus very similar to that observed in erythroid progenitors was found, indicating that EKLF is needed for the progression to and/or stabilisation of a fully functional ACH. This shows that transcription factors can play an essential role in the three-dimensional organisation of gene loci (Drissen et al. 2004).

It is tempting to propose that ACH formation is key to establishing gene expression in repressive chromatin and to maintain hypersensitivity at genomic sites (Tolhuis et al. 2002). Clustering of regulatory DNA elements, mediated by the affinity between proteins bound to these sites, may explain how genes on overlapping loci establish independent expression patterns (de Laat and Grosveld 2003). Also, the work suggests that clustering of *cis*-regulatory elements and active genes is essential for the high transcription rate of the β-globin genes.

Transcription factors move rapidly through the nucleus, but efficient transcription requires a high local concentration of these factors at the genomic site. Local accumulation of *trans*-acting factors can be obtained by spatial clustering of cognate binding sites. Thus, the β-globin ACH can be seen as a nuclear compartment dedicated to efficient (RNA polymerase II-mediated) transcription of the β-globin genes, analogous to the nucleolus being a compartment dedicated to RNA polymerase I-mediated transcription of rRNA genes (Palstra et al. 2003).

The pros and cons of 3C

3C technology has turned out to be a powerful technique and it has several important advantages over other methods used to determine the spatial organisation of a locus; cross-linking frequencies can be determined at many positions within a locus (e.g. active or inactive genes, intergenic sequences or regulatory elements) with the position of restriction sites as the only limitation and it can measure cross-linking frequencies independent of the transcriptional status of a given locus. However, it is important to keep the current limitations of 3C technology in mind. Cross-linking frequencies represent steady-state average levels of interactions between DNA elements. Interactions between *cis*-regulatory elements are thought to be dynamic *in vivo*, therefore short-lived but important interactions may give lower cross-linking frequencies than more long-lived but less crucial interactions. Furthermore, cell populations analysed are never homogeneous and cells that differ from the cells of interest still contribute to the measured cross-linking frequencies. With the current state of the 3C technology, interpretation of 3C data in a strictly quantitative manner is in my view not possible. Cross-linking frequencies are influenced by several parameters like the fragment size generated by the restriction enzymes used and changes in cross-linking conditions. How chromatin elements are cross-linked to each other and how the proteins present on a chromatin fibre influence this is totally unknown. Formaldehyde is a tight (2 Å) cross-linking agent that efficiently produces both protein–nucleic acid and protein–protein cross-links *in vivo* and primarily interacts with amino and imino groups of amino acids (lysines, arginines and histidines) and of DNA (primarily adenines and cytosines). Therefore, unusual occurrence of these amino and imino groups at certain sites could in principle bias cross-

linking of certain chromatin regions. The size of the restriction fragments has a major impact on the assay given the inability to separate *cis*-regulatory elements of interest and a high 'background' cross-linking. Linearity of PCR amplification for each amplicon in a locus is also not guaranteed under the PCR conditions used and therefore some interactions might be under- or overestimated. The use of frequent cutters, milder fixation conditions and real time PCR to quantify amplicons will significantly enhance the resolution and better allow quantitative interpretation of 3C data.

At present it remains impossible to determine the compaction ratio of a locus and thus determine wether chromatin in a locus is folded as a 10nm fibre, a 30nm fibre or wether it adopts yet a different "higher" conformation. Moreover, chromatin folding is certainly not uniform along the entire β-globin locus, given the presence of domains of chromatin modifications and Dnase I hypersensitive sites. Therefore, we are still a long way from describing the data obtained by 3C technology in quantitative terms, and interpreting the relationship between cross-linking frequency and genomic site separation in terms of real distances. Measured interactions are particularly meaningful if they can be correlated to a phenotype, e.g. if they occur only in a transcriptional active locus and only take place between certain genomic elements like *cis*-regulatory elements. From this follows that the interpretation of data obtained with 3C is another concern. Conclusions about the conformation of a locus can only be made when cross-linking frequencies between enough fragments are analysed from different viewpoints, i.e. diffrent genomic sites. Obtaining a PCR signal between two given restriction fragments is not enough to conclude that specific interactions take place between elements and a looped structure is formed. For this several elements in the locus have to be analysed and it has to be shown that other elements interact less frequently.

Several groups have recently published papers in which 3C technology was applied. It was used for the detection of interactions between boundary elements in *drosophila* (Blanton et al. 2003), for the detection of parent of origin specific interactions between DMRs in the *Igf2/H19* locus (Murrell et al. 2004), to determine conformational organisation of single genes (Eivazova and Aune 2004; O'Sullivan et al. 2004), to show co-localisation of multiple active genes to a shared site of ongoing transcription (Osborne et al. 2004) and in a nicely executed analysis of long-range intra-chromosomal interactions within the T helper 2 (T_H2) cytokine locus (Spilianakis and Flavell 2004). Most of these studies used the 3C technology as an independent technique to confirm observations made using other techniques such as FISH and ChIP. In these studies the proper controls were often omitted, linearity of amplification was not determined and extreme PCR conditions were used to obtain amplification products. If the use of nested PCR, involving many PCR cycles, is required to demonstrate the existence of a given ligation product as is done in 'qualitative 3C', one can question if this reflects nuclear proximity or a rare random ligation event between these two fragments (Murrell et al. 2004; Osborne et al. 2004). In another example, Eivazova and Aune analysed re-ligation efficiency of a restriction site to normalise their PCR signals between non-homologous primer pairs (Eivazova and Aune 2004). However, since PCR across an existing restriction site analyses both the non-digested and re-ligated site, which both will be present in vast excess over any ligation product between non-related restriction sites, linearity of amplification for all amplicons is impossible to maintain, and therefore the results will be skewed. The authors propose a spatial model of the *ifng* gene in different T cell subsets based on the relative positions of interacting restriction sites, however it is not the restriction sites

but the restriction fragments that are cross-linked together and when taken into account this would change their model dramatically. Only recently the first locus wide analysis of cross-linking frequencies of a locus different from the β-globin locus was published (Spilianakis and Flavell 2004). The T_H2 cytokine locus was analysed in a large panel of different cells and a great deal of attention was spend on validating the 3C conditions and the results these authors obtained will be discussed later.

The active chromatin hub; establishment of LCR-promoter contacts

The work presented here clearly demonstrates that distant *cis*-regulatory elements, the β-globin LCR and promoter, can physically interact to control gene activity, as predicted by the original looping models (Tolhuis et al. 2002). How this looped structure is formed and maintained remains elusive though. Random collision (by Brownian motion) between distal elements in the active β-globin locus currently seems to be the most likely hypothesis although theoretical calculations (Rippe 2001) and experimental measurements of site-specific recombination between sites separated by several kilobases in mammalian cells (Ringrose et al. 1999) suggest that random diffusion alone is not enough to establish contacts between two sites separated by several kilobases on a chromatin fibre. As discussed in chapter 1 additional mechanisms might exist that facilitate enhancer-promoter contact over a distance.

Nevertheless, processes like V(D)J recombination are thought to involve DNA looping and this occurs over vast distances. Moreover lox/Cre based recombination techniques work efficiently over distance, as exemplified in an experiment that inverts the order of the five human beta-globin genes over a distance similar to the distance of LCR-promoter contact (Tanimoto et al. 1999). This shows that establishment of loops in chromatin might be less problematic than suggested. However, recombination events only require short lived one time interactions while transcription of, for example the β-globin genes requires longer contact times estimated to be in the range from 45 to 80 minutes for the β-LCR interaction in definitive erythrocytes (Wijgerde et al. 1995). Stabilisation of the interaction might therefore be a more important determinant of successful loop formation/maintenance. Self-organising principles could guide loop formation and ACH maintenance. Self-organisation involves the physical interaction of molecules in a steady-state structure, which is open for exchange of energy and matter, and is a principle used for building most cellular structures (Misteli 2001a; Cook 2002). These structures are intrinsically unstable and persist only by exchanging components with their surroundings and maintenance of the structure depends on continuing function, which is transcription in the case of the ACH. DNase I HS formation precedes transcription but hypersensitivity of the LCR depends on the presence of active promoters (Blom van Assendelft et al. 1989; Reitman et al. 1993; Guy et al. 1996; Tewari et al. 1996). This suggests that the stability and maintenance of these sites and therefore ACH formation relies on ongoing transcription. Furthermore, deletions of LCR elements result in variegated expression of transgenes and importantly, the affected loci are insensitive to DNase I in the non-transcribing portion of the cells (Festenstein et al. 1996; Milot et al. 1996).

In the absence of the erythroid specific transcription factor EKLF, contacts the LCR and genes are lost (chapter 7; Drissen et al. 2004). EKLF binds DNA but is not known to homomultimerise (like e.g. Bach1 or GAGA); therefore EKLF is not likely to be directly

responsible for bringing the distant DNA elements together. Rather, the ACH must be thought of as a multi-component entity containing many different proteins and DNA elements. Removal of one of the crucial components like e.g. EKLF will destabilise the structure and the ACH will collapses into the non-transcriptionally active, EKLF independent, CH. This hypothesis predicts that the ability to attract the right trans acting factors at the right time will determine the affinity between the *cis*-regulatory DNA elements and indeed in the absence of GATA-1 the LCR-promoter interactions also appear to be lost (G.A. Blobel, personal communication).

A group of ubiquitously expressed proteins, facilitator factors, might further stabilize the ACH formation. Facilitator factors were originally identified in *Drosophila* because they facilitate enhancer-promoter communication (Dorsett 1999). These facilitators have mammalian homologues like LIM-domain binding protein 1 (Ldb1) and Idn3 (Morcillo et al. 1997; Rollins et al. 1999). Ldb-1 was found in a large complex with erythroid cell specific transcription factors, including GATA-1 (Wadman et al. 1997). Thus, these facilitators act through interactions with other *trans*-acting factors and may link the interacting *cis*-regulatory DNA elements to stabilize the ACH.

In progenitor erythrocytes the distal HSs and the HS4/5 of the LCR are spatially clustered in a structure called the CH. None of the β-globin genes are expressed at this stage and the β-globin genes themselves are not yet associated with the CH suggesting a more structural role for the CH. The presence of a CH formed by the distal HS and the 5'HS of the LCR might facilitate the formation of productive interactions between the LCR and the β-globin genes by restricting movement of the DNA elements to a confined space. It has long been proposed that chromatin in the nucleus is organised into loops, which might be anchored at their bases to a nuclear substructure: the elusive nuclear matrix. An interesting possibility is that those chromatin loops (presumably found throughout the genome) and the specific loops of the CH are related. All the sites involved in forming the CH bind the ubiquitous zinc finger protein CTCF (Bulger et al. 2003), a protein involved in enhancer blocking (Bell et al. 1999), shown to interact with the nuclear matrix (Yusufzai and Felsenfeld 2004) and possibly involved in setting up looped higher order chromatin structures (Yusufzai et al. 2004). If the CH is part of the genome wide higher order chromatin loops, and perhaps interacts with the nuclear matrix, it appears to be erythroid-specific and therefore other, probably erythroid-specific, proteins besides CTCF must be involved. Recently it was shown that Poly(ADP-ribosyl)ation modification of CTCF is essential for its enhancer blocking activity (Yu et al. 2004) and possibly an erythroid specific mechanism exists to target this modification to CTCF bound to the β-globin locus.

The active chromatin hub; activation of a gene by an enhancer/LCR

The ability of LCRs and other combinations of *cis*-regulatory elements to form an ACH may underlie the position independent, copy number dependent expression levels that are obtained when transgenes are linked to these elements. This view is supported by experiments with transgenic constructs that suggest a mass action model in which a critical number of regulatory elements are necessary to form an ACH in repressive chromatin. A mass action model strongly depends on a critical number of interactions to increase the likelihood

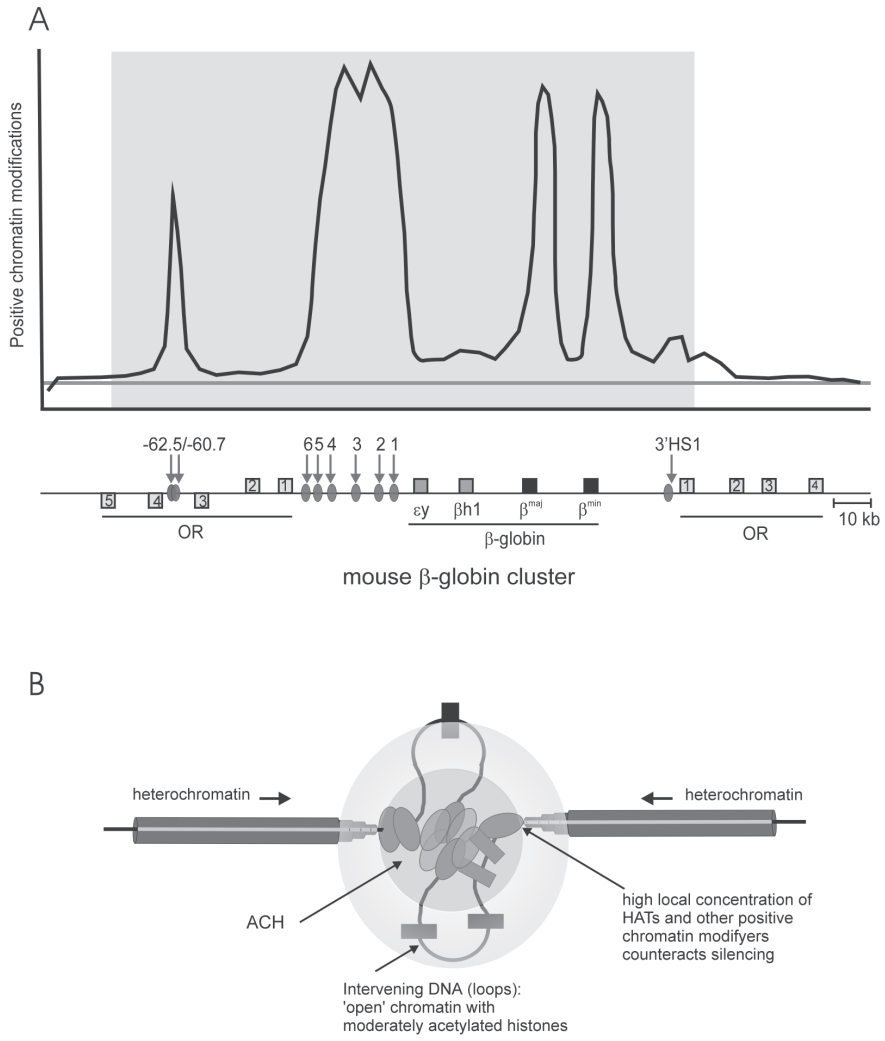


figure 8.1: Positive chromatin modifications in the β-globin locus and the ACH.
(A). The approximate positions of positive chromatin modifications (e.g. acetylated H3) in the mouse β-globin locus. The horizontal grey line indicates background level and the grey box indicates the Dnase I sensitive domain.
(B). The ACH and transcription in repressive chromatin: 2D presentation of a 3D gene cluster embedded in heterochromatin (dark grey), with active genes and cognate cis-regulatory DNA elements (grey ovals) and inactive genes (grey boxes). The cluster of regulatory sequences and active genes forms an ACH, which contains a high local concentration (indicated by grey inner circle) of transcription factors, HATs and other positive chromatin modifying factors. Spreading of heterochromatic silencing (dark grey) is blocked at the ACH, mainly due to the local accumulation of these factors. Concentration of positive factors decreases with increasing distance from the ACH (indicated by the outer light-grey circle). Loops of intervening DNA (with inactive genes) protrude from the ACH; the nearby hub of positive chromatin modifiers affects the chromatin status of these loops. This figure was adapted from (de Laat and Grosveld 2003)

of establishing a stable ACH structure. In agreement with this model is the observation that position-independent expression of an enhancer driven construct is only observed with multi-copy integrations (Ellis et al. 1993). Moreover, deletion of single hypersensitive sites from the LCR in transgenic constructs results in position effects and this is enhanced when levels of erythroid transcription factors are reduced (reviewed in de Laat and Grosveld 2003).

The clustering of regulatory elements in nuclear space is proposed to create a high local concentration of transcription factors and associated 'positive' chromatin-modifying enzymes and 'insulates' the locus from the encroachment of 'silent chromatin'. The DNase I sensitivity and the acetylation levels of the β-globin locus demonstrate an unexpected complexity associated with the active locus (Bulger et al. 2003). The complete active β-globin locus has a low level of acetylation on its histones, although this is still higher than observed in inactive regions. Superimposed on this, regions highly enriched in acetylated histones H3 and H4 are found over the regulatory elements (figure 8.1A) (Forsberg and Bresnick 2001; Bulger et al. 2003). The high peaks of positive chromatin modifications, like acetylated histone H3 and H4, could be caused by the high local concentration of positive chromatin modifiers located at the clustered *cis*-regulatory elements (figure 8.1B). Interestingly, the observed pattern of modestly increased acetylation levels of intergenic regions, often taken as evidence for tracking mechanisms, can also be readily explained by the formation of an ACH. The lower level of modifications spread throughout the locus are in this situation the consequence of the relative proximity of these regions to the ACH, and can be regarded as being by-products of activities concentrated at the ACH without necessarily having functional significance (figure 8.1B) (de Laat and Grosveld 2003).

Similarly, intergenic transcripts detected in the human β-globin locus (Ashe et al. 1997; Gribnau et al. 2000) can be explained by promiscuous initiation of transcription as a consequence of established open chromatin domains. The apparent specific initiation sites found for these transcripts could be due to the presence of sequences with a higher affinity for RNA polymerase II. Deletion of the proposed intergenic promoter down stream of the ^{A}γ-gene results in position effects and a decrease in the efficiency of transcription of the beta-globin gene (Calzolari et al. 1999) or only very mild changes in the stage-specific sequential expression of the ε-, γ-, and β-globin genes (Gaensler et al. 2003). Moreover, substitution of the ^{A}γ-promoter by a human β-spectrin promoter in a 40-kb cosmid construct, containing a micro-LCR and all genes downstream of and including the ^{A}γ-gene, showed no developmental switching and expressed both human γ- and β-globin mRNAs in erythroid cells of all stages throughout development (Sabatino et al. 1998). These observations argue against any putative regulatory sequences downstream of the ^{A}γ-gene necessary for the intergenic transcript dependent opening of chromatin domains.

Deletions of HS from the endogenous mouse β-globin LCR have surprisingly mild phenotypes and even when the complete LCR is deleted the locus still remains in an 'open' acetylated state while the rate of transcription is reduced to ~4% of that observed in the wild type and acetylation at the promoter is maintained (Bender et al. 2000a; Schubeler et al. 2001). The presence of distal HS could still maintain an 'open' chromatin domain and the promoters are in this situation still able to attract chromatin modifying transcription factors. Elucidation of the structure of the mouse β-globin locus with a LCR deletion would therefore be of considerable interest.

The active chromatin hub; protein concentrations at promoters and transcription factories

The overall concentration of many proteins in the nucleus is estimated to be in the range of, or lower than, the dissociation constant for many protein–protein or protein–DNA interactions (Chambeyron and Bickmore 2004b). The formation of productive transcription complexes on DNA is expected to take place via a 'stop and go' mechanism (Misteli 2001b). In this mechanism a factor binds to its binding site and will drop off if another factor doesn't bind within its residence time, which is the basic principle of mass activation (Locke et al. 1988). This makes protein concentration a very sensitive limiting factor in the efficiency of transcription. Moreover, transcription factors move rapidly through the nucleus, so how is a high local concentration of these factors at the genomic site required for efficient transcription accomplished? A local accumulation of *trans*-acting factors can be obtained by a spatial clustering of their cognate binding sites (Droge and Muller-Hill 2001).

Recent analysis of the RNA polymerase II transcription cycle has shown that most transcription units are unoccupied by RNA polymerases for three quarters of their time (Kimura et al. 2002). Initiation of transcription is very rapid (seconds) and elongation takes minutes to hours. On the other hand the assembly of the pre-initiation complex, by the stochastic exchange of freely mobile components, takes up one half to five sixths of the duration of a transcription cycle. Transcription reinitiation, the cyclic process of RNA synthesis from active genes, bypasses several protein-DNA association steps and the rate of new transcription cycles is increased with respect to the first transcription round (Dieci and Sentenac 2003). A reinitiation intermediate that includes transcription factors TFIID, TFIIA, TFIIH, TFIIE and Mediator has been identified and this intermediate is stabilized in the presence of an activator (Yudkovsky et al. 2000). Several studies demonstrate that RNA Polymerase II and other basal transcription factors are recruited to LCR core HS elements and then transferred to the β-globin promoter (Johnson et al. 2002; Johnson et al. 2003; Vieira et al. 2004). Concentration of both RNA polymerases and transcription units in specific clusters or compartments is expected to significantly facilitate RNA polymerase recycling and high-frequency re-utilisation of stable preinitiation complexes, thus enabling the production of extremely high levels of transcript. Thus, the β-globin ACH may function as a nuclear compartment dedicated to efficient RNA polymerase II-mediated transcription, in particular reinitiation, of the β-globin genes (Palstra et al. 2003).

Estimates of the total number of nascent transcripts or active polymerase molecules and the number of transcription sites within a cell led to the conclusion that transcription by RNA polymerase II is carried out in transcription factories (Jackson et al. 1993). The general notion is that each transcription factory would contain several active polymerases and associated transcripts. It is tempting to speculate that the β-globin ACH is such a RNA polymerase II transcription factory. A recent study indicates that transcriptional active distal genes co-localise at high frequencies to the active β-globin gene (Osborne et al. 2004). The authors suggest that, rather than recruiting and assembling transcription complexes; active genes migrate to pre-assembled transcription sites. Given the high mobility of RNA polymerase II in the nucleus, it is very unlikely that preformed factories exist. The observed

'flip-flop' mechanism of transcription of the β-globin genes and the estimated load of 8-10 RNA polymerase II molecules on a β-globin open reading frame, which is close to the amount estimated to occupy a transcription factory, also argue against a direct participation of these genes in the postulated β-globin factory. Therefore, a more plausible explanation is that these genes make promiscuous use of the high concentration of transcription factors accumulating near the (dynamic) β-globin ACH. This process might be driven by the initiation of large-scale heterochromatisation of erythroid chromatin upon differentiation.

The active chromatin hub; organisation of multi gene domains

The discovery of cis-regulatory elements able to promote transcription of genes over large distances led to the postulate that elements, termed insulators, should exist that would limit the action of enhancers, LCRs and silencers to defined domains. Insulator and boundary models postulate that chromosomes are subdivided into physically distinct expression domains containing a gene or gene cluster and all its *cis*-regulatory elements. Specialised elements, called insulators or boundary elements, are proposed to be located at the borders of such domains to counteract the inappropriate effects of nearby heterochromatin and/or distal enhancers (Labrador and Corces 2002; West et al. 2002). Indeed, several elements with these properties have been identified (reviewed in Geyer and Clark 2002). An important implication of the insulator and boundary model is that genes are supposed to reside in functionally distinct domains. However, unlike housekeeping genes, tissue-specific genes generally do not cluster on chromosomes, can be found in gene-dense regions of chromosomes as well as in regions of repressive chromatin and these genes with their associated *cis*-acting sequences often do not occupy physically distinct domains. Several gene loci, such as the chicken lysozyme locus (Bonifer et al. 1997; Chong et al. 2002; Lee et al. 2003), the T_H2 cytokine locus (Lee et al. 2003), the α- and β-globin loci (Kielman et al. 1993; Bulger et al. 1999), the human growth hormone locus (Bennani-Baiti et al. 1995; Bennani-Baiti et al. 1998) and the CD3 locus (Lerner et al. 1993) overlap with unrelated genes and their *cis*-acting sequences are spread out over large domains and can be found tens or hundreds of kilobases away from the target gene (figure 8.2).

The observation that independently regulated loci can overlap partially or completely raises questions about the functional requirements for physically isolated domain structures. Alternatively a domain, or functional expression module, could be determined by the distribution of enhancers containing binding sites for positively acting factors. Specifity of interactions between these enhancers and promoters would play a major role in maintaining the functional autonomy. Sequences that interfere with these interactions, characterised as insulators but often associated with other functions like transcription, would be selected against if they occurred within the domain but not at the edges (Dillon and Sabbattini 2000). The formation of an ACH might in such instances underlie correct gene expression, and proximity and specificity determines which *cis*-acting sequences and promoter(s) form an ACH, and thus which gene will be expressed (de Laat and Grosveld 2003). ACH formation therefore provides a mechanistic framework to understand how overlapping gene loci can set-up independent tissue specific expression profiles. Depending on the presence of tissue specific proteins with the appropriate affinities for the cognate DNA sequences, a tissue

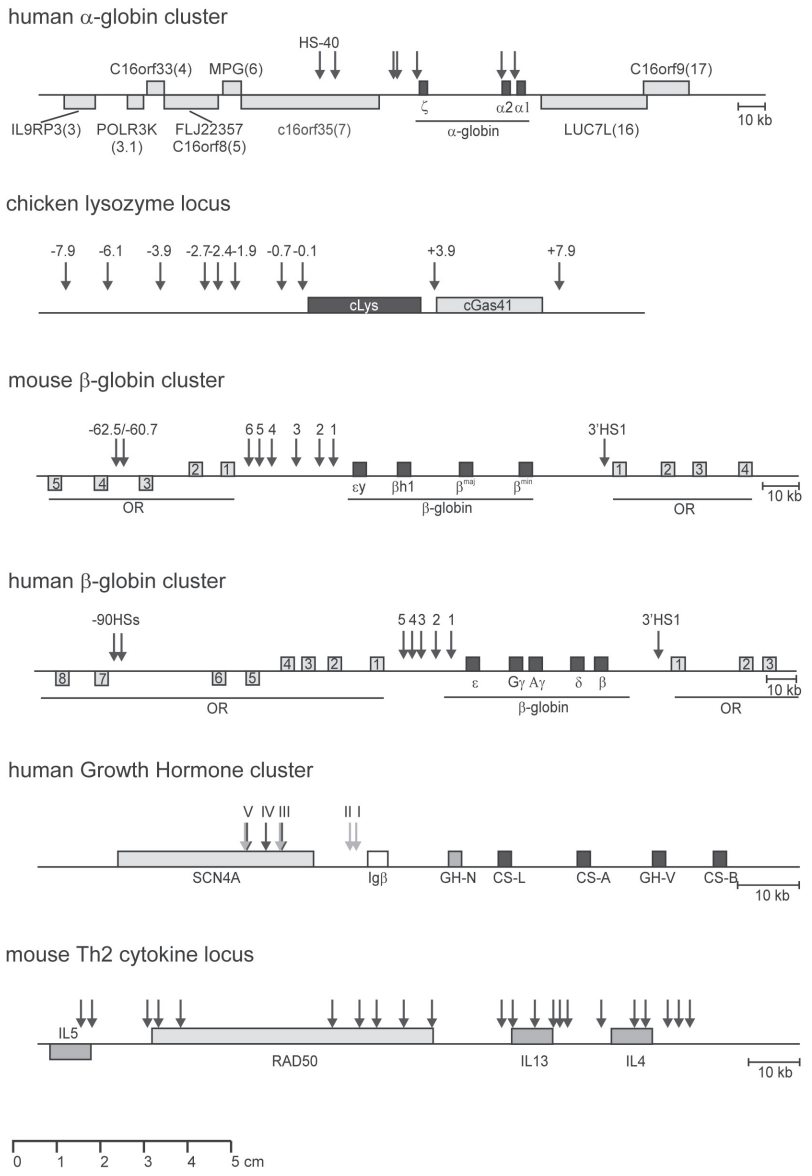


figure 8.2: Genomic organization of a variety of gene loci.
Boxes above the horizontal lines represent genes transcribing from left to right, boxes below represent genes transcribed in the opposite direction. Differentially expressed genes are indicated by different shades of grey and occur within one locus, but not separated by insulator or boundary elements. Arrows indicate known DNase I HSs. The HSs are located throughout the depicted gene loci and often separated by a few kb to tens of kb. Note that the scale in kb is indicated below each genomic fragment, except for the chicken lysozyme locus, where the numbering of HSs corresponds to the distance in kb from the transcriptional start site of the lysozyme gene. This figure was adapted from (de Laat and Grosveld 2003).

specific ACH in one cell-type will be set-up, while a different ACH activating a different gene will be set up in another cell-type.

It remains to be determined how common the mechanism of ACH formation is in different gene loci. Recently an extensive analysis of the spatial organisation of the T_H2 cytokine locus was conducted (Spilianakis and Flavell 2004). This study identified two different long-range intra-chromosomal interactions in the locus, forming a structure similar to the β-globin ACH, but with some intriguing differences. In fibroblast, B cell, NK cells and $CD4^+$ T cells the promoters of all the cytokine genes, *Il5*, *Il13* and *Il4*, interact and form a 'pre-poised' initial core chromatin formation analogous to the β-globin CH. In cells of the T cell lineage that do not express the T_H2 cytokines, the T_H2 LCR interacts with this initial core forming a 'poised' chromatin conformation, which T_H2 cell-specific transcription factors can rapidly activate, by binding to the LCR. In this way the T_H2 LCR can rapidly co-activate the three cytokine-encoding genes upon stimulation without the need to induce additional conformational changes in the locus. The intervening, constitutively expressed *Rad50* gene, which is looped out of the ACH, is not influenced. The study also shows that two transcription factors highly expressed in T_H2 cells, GATA3 and STAT6, are responsible for generating and/or maintaining this structure (Spilianakis and Flavell 2004).

It is clear that more ACHs will be found in different loci, each with its own peculiar organisation and dynamics. This however, does not necessarily mean that all genes will form an ACH and it is more likely that ACH formation is confined to loci that have either high transcription levels, are located in highly restrictive chromatin and/or are under stringent transcriptional control. It can furthermore be envisioned that clustering of chromosomal regions plays a role in establishing and/or maintaining inactive chromatin states e.g. polycomb mediated silencing (Bantignies et al. 2003; Chambeyron and Bickmore 2004b).

Future directions

The analysis of the spatial organisation of the mouse β-globin locus has led to exciting new insights into its transcriptional regulation. However, many key questions about the regulation of the β-globin locus and chromatin folding in the spatial context of the nucleus remain.

Given the limitations in relating the 3C interaction data to the exact spatial conformation, it is currently not clear how close in space the interacting elements really are. Therefore, statements about direct interactions between the β-globin gene and elements within the LCR are currently not possible. 3C technology as it is currently applied is not quantitative enough and lacks the resolution that is needed to discriminate between direct and indirect interactions. Increasing the resolution of 3C technology by using restriction enzymes, which cut more frequent and adjustment of crosslink conditions can potentially further improve the resolution of the 3C technique. Quantification of ligation products by real-time PCR would allow for a more quantitative interpretation of the data and would mark an important improvement of the technology. This improvement in resolution is necessary to investigate the chronology of events that occur in the establishment of an ACH from the CH during erythroid maturation and induction of β-globin expression. Determination of the β-globin ACH structure after treatment of the cells with different transcription blockers like α-amanitin

or DRB will show if maintenance of the ACH is dependent on ongoing transcription. Analysis of the conformational changes in the β-globin locus during differentiation of hematopoeitic stem cells into erythrocytes and correlated to differences in DNaseI (hyper)-sensitivity and histone modifications may reveal an order of events at the chromatin level that underlie gene activation and silencing.

The nature of long-range chromatin interactions observed in non-expressing progenitors and the identification of distal DNA elements that loop to this structure merits further investigation and would give unprecedented insight into chromosomal organisation. This could be achieved by cloning the DNA-fragments that are ligated to the distal HS followed by sequencing to identify the interacting DNA-elements. Not only *cis*-regulatory elements but also non-globin genes (Osborne et al. 2004) interacting with the ACH can possibly be detected via this approach. A candidate approach based on conserved sequences identified by bioinformatics could also be a feasible option.

Visualisation of chromatin looping and ACH formation by DNA-FISH methods would be extremely valuable in confirming and extending the data obtained with 3C. The data generated with 3C technology provide a steady state population average of the structure of the β-globin locus and visualisation of the ACH could determine wether the distal *cis*-regulatory elements are in close proximity to the LCR in every expressing cell or wether there are subsets of cells that have different structures. Dynamics of the interactions can be studied by tagging the most distal hypersensitive sites, (-85HS (unpublished results) and 3'HS1) which participate in the hub (150 kb apart), with repeats of binding sites for bacterial repressor proteins. Both ends of the mouse β-globin locus can in this way be labelled differentially with GFP-repressor protein fusions and the dynamics of the interactions between the *cis*-regulatory elements of the endogenous mouse β-globin locus can then be visualised in living interphase nuclei. The concept of transcription factories and the dynamics of transcription factors on active transcription units can conceivably be investigated in more detail by using a tagged β-globin locus to locate an active locus in the nucleus in combination with the bleaching of GFP-tagged transcription factors in its neighbourhood to measure the dynamics of these factors.

The detailed knowledge of the β-globin locus makes it a perfect model system to study the spatial organisation of multi gene loci and its dynamics in expressing vs. non-expressing cells. Future analysis of the β-globin locus will keep leading the way in revealing new mechanisms of gene regulation and will provide answers to questions such as: can we reconstruct the three-dimensional structure of a genomic locus, can we visualise chromatin hub formation, what is the functional consequence of chromatin hubs in the nucleus, what are the dynamics of chromatin hubs in the nucleus? Answers to these questions will have a high impact on our understanding of transcriptional regulation and the functional significance of chromatin folding *in vivo*.

References

Abranches R., Beven A.F., Aragon-Alcaide L., and Shaw P.J. 1998. Transcription sites are not correlated with chromosome territories in wheat nuclei. *J Cell Biol* 143: 5-12.

Adkins N.L., Watts M., and Georgel P.T. 2004. To the 30-nm chromatin fiber and beyond. *Biochim Biophys Acta* 1677: 12-23.

Agalioti T., Lomvardas S., Parekh B., Yie J., Maniatis T., and Thanos D. 2000. Ordered recruitment of chromatin modifying and general transcription factors to the IFN-beta promoter. *Cell* 103: 667-678.

Allfrey V.G., Faulkner R., and Mirsky A.E. 1964. Acetylation and Methylation of Histones and Their Possible Role in the Regulation of Rna Synthesis. *Proc Natl Acad Sci U S A* 51: 786-794.

Alvarez M., Rhodes S.J., and Bidwell J.P. 2003. Context-dependent transcription: all politics is local. *Gene* 313: 43-57.

Andrews N.C., Erdjument-Bromage H., Davidson M.B., Tempst P., and Orkin S.H. 1993a. Erythroid transcription factor NF-E2 is a haematopoietic-specific basic-leucine zipper protein. *Nature* 362: 722-728.

Andrews N.C., Kotkow K.J., Ney P.A., Erdjument-Bromage H., Tempst P., and Orkin S.H. 1993b. The ubiquitous subunit of erythroid transcription factor NF-E2 is a small basic-leucine zipper protein related to the v-maf oncogene. *Proc Natl Acad Sci U S A* 90: 11488-11492.

Armstrong J.A., Bieker J.J., and Emerson B.M. 1998. A SWI/SNF-related chromatin remodeling complex, E-RC1, is required for tissue-specific transcriptional regulation by EKLF in vitro. *Cell* 95: 93-104.

Ashe H.L., Monks J., Wijgerde M., Fraser P., and Proudfoot N.J. 1997. Intergenic transcription and transinduction of the human beta-globin locus. *Genes Dev* 11: 2494-2509.

Banerji J., Rusconi S., and Schaffner W. 1981. Expression of a beta-globin gene is enhanced by remote SV40 DNA sequences. *Cell* 27: 299-308.

Bannister A.J., Schneider R., and Kouzarides T. 2002. Histone methylation: dynamic or static? *Cell* 109: 801-806.

Bannister A.J., Zegerman P., Partridge J.F., Miska E.A., Thomas J.O., Allshire R.C., and Kouzarides T. 2001. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410: 120-124.

Bantignies F., Grimaud C., Lavrov S., Gabut M., and Cavalli G. 2003. Inheritance of Polycomb-dependent chromosomal interactions in Drosophila. *Genes Dev* 17: 2406-2420.

Bazett-Jones D.P., Cote J., Landel C.C., Peterson C.L., and Workman J.L. 1999. The SWI/SNF complex creates loop domains in DNA and polynucleosome arrays and can disrupt DNA-histone contacts within these domains. *Mol Cell Biol* 19: 1470-1478.

Becker M., Baumann C., John S., Walker D.A., Vigneron M., McNally J.G., and Hager G.L. 2002. Dynamic behavior of transcription factors on a natural promoter in living cells. *EMBO Rep* 3: 1188-1194.

Behringer R.R., Hammer R.E., Brinster R.L., Palmiter R.D., and Townes T.M. 1987. Two 3' sequences direct adult erythroid-specific expression of human beta-globin genes in transgenic mice. *Proc Natl Acad Sci U S A* 84: 7056-7060.

Bell A.C. and Felsenfeld G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the Igf2 gene. *Nature* 405: 482-485.

Bell A.C., West A.G., and Felsenfeld G. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* 98: 387-396.

Belmont A. 2003. Dynamics of chromatin, proteins, and bodies within the cell nucleus. *Curr Opin Cell Biol* 15: 304-310.

Belmont A.S. 2001. Visualizing chromosome dynamics with GFP. *Trends Cell Biol* 11: 250-257.

Belmont A.S., Braunfeld M.B., Sedat J.W., and Agard D.A. 1989. Large-scale chromatin structural domains within mitotic and interphase chromosomes in vivo and in vitro. *Chromosoma* 98: 129-143.

Bender M.A., Bulger M., Close J., and Groudine M. 2000a. Beta-globin gene switching and DNase I sensitivity of the endogenous beta-globin locus in mice do not require the locus control region. *Mol Cell* 5: 387-393.

Bender M.A., Mehaffey M.G., Telling A., Hug B., Ley T.J., Groudine M., and Fiering S. 2000b. Independent formation of DnaseI hypersensitive sites in the murine beta-globin locus control region. *Blood* 95: 3600-3604.

Bender M.A., Reik A., Close J., Telling A., Epner E., Fiering S., Hardison R., and Groudine M. 1998. Description and targeted deletion of 5' hypersensitive site 5 and 6 of the mouse beta-globin locus control region. *Blood* 92: 4394-4403.

Bender M.A., Roach J.N., Halow J., Close J., Alami R., Bouhassira E.E., Groudine M., and Fiering S.N. 2001. Targeted deletion of 5'HS1 and 5'HS4 of the beta-globin locus control region reveals additive activity of the DNaseI hypersensitive sites. *Blood* 98: 2022-2027.

Bennani-Baiti I.M., Cooke N.E., and Liebhaber S.A. 1998. Physical linkage of the human growth hormone gene cluster and the CD79b (Ig beta/B29) gene. *Genomics* 48: 258-264.

Bennani-Baiti I.M., Jones B.K., Liebhaber S.A., and Cooke N.E. 1995. Physical linkage of the human growth hormone gene cluster and the skeletal muscle sodium channel alpha-subunit gene (SCN4A) on chromosome 17. *Genomics* 29: 647-652.

Benoist C. and Chambon P. 1981. In vivo sequence requirements of the SV40 early promotor region. *Nature* 290: 304-310.

Berg O.G., Winter R.B., and von Hippel P.H. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* 20: 6929-6948.

Bernstein B.E., Liu C.L., Humphrey E.L., Perlstein E.O., and Schreiber S.L. 2004. Global nucleosome occupancy in yeast. *Genome Biol* 5: R62.

Bi X., Yu Q., Sandmeier J.J., and Zou Y. 2004. Formation of boundaries of transcriptionally silent chromatin by nucleosome-excluding structures. *Mol Cell Biol* 24: 2118-2131.

Bickel S. and Pirrotta V. 1990. Self-association of the Drosophila zeste protein is responsible for transvection effects. *Embo J* 9: 2959-2967.

Bieker J.J. 2001. Kruppel-like factors: three fingers in many pies. *J Biol Chem* 276: 34355-34358.

Blackwood E.M. and Kadonaga J.T. 1998. Going the distance: a current view of enhancer action. *Science* 281: 61-63.

Blanton J., Gaszner M., and Schedl P. 2003. Protein:protein interactions and the pairing of boundary elements in vivo. *Genes Dev*

17: 664-675.
Blessing C.A., Ugrinova G.T., and Goodson H.V. 2004. Actin and ARPs: action in the nucleus. *Trends Cell Biol* 14: 435-442.
Blobel G.A., Nakajima T., Eckner R., Montminy M., and Orkin S.H. 1998. CREB-binding protein cooperates with transcription factor GATA-1 and is required for erythroid differentiation. *Proc Natl Acad Sci U S A* 95: 2061-2066.
Blom van Assendelft G., Hanscombe O., Grosveld F., and Greaves D.R. 1989. The beta-globin dominant control region activates homologous and heterologous promoters in a tissue-specific manner. *Cell* 56: 969-977.
Bode J., Benham C., Knopp A., and Mielke C. 2000. Transcriptional augmentation: modulation of gene expression by scaffold/matrix-attached regions (S/MAR elements). *Crit Rev Eukaryot Gene Expr* 10: 73-90.
Bodine D.M. and Ley T.J. 1987. An enhancer element lies 3' to the human A gamma globin gene. *Embo J* 6: 2997-3004.
Bondarenko V.A., Jiang Y.I., and Studitsky V.M. 2003a. Rationally designed insulator-like elements can block enhancer action in vitro. *Embo J* 22: 4728-4737.
Bondarenko V.A., Liu Y.V., Jiang Y.I., and Studitsky V.M. 2003b. Communication over a large distance: enhancers and insulators. *Biochem Cell Biol* 81: 241-251.
Bonifer C., Jagle U., and Huber M.C. 1997. The chicken lysozyme locus as a paradigm for the complex developmental regulation of eukaryotic gene loci. *J Biol Chem* 272: 26075-26078.
Bottardi S., Aumont A., Grosveld F., and Milot E. 2003. Developmental stage-specific epigenetic control of human beta-globin gene expression is potentiated in hematopoietic progenitor cells prior to their transcriptional activation. *Blood* 102: 3989-3997.
Boyes J., Omichinski J., Clark D., Pikaart M., and Felsenfeld G. 1998. Perturbation of nucleosome structure by the erythroid transcription factor GATA-1. *J Mol Biol* 279: 529-544.
Brand M., Ranish J.A., Kummer N.T., Hamilton J., Igarashi K., Francastel C., Chi T.H., Crabtree G.R., Aebersold R., and Groudine M. 2004. Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. *Nat Struct Mol Biol* 11: 73-80.
Brownell J.E. and Allis C.D. 1996. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev* 6: 176-184.
Brownell J.E., Zhou J., Ranalli T., Kobayashi R., Edmondson D.G., Roth S.Y., and Allis C.D. 1996. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84: 843-851.
Buck M., Gallegos M.T., Studholme D.J., Guo Y., and Gralla J.D. 2000. The bacterial enhancer-dependent sigma(54) (sigma(N)) transcription factor. *J Bacteriol* 182: 4129-4136.
Bulger M., Bender M.A., van Doorninck J.H., Wertman B., Farrell C.M., Felsenfeld G., Groudine M., and Hardison R. 2000. Comparative structural and functional analysis of the olfactory receptor genes flanking the human and mouse beta-globin gene clusters. *Proc Natl Acad Sci U S A* 97: 14560-14565.
Bulger M. and Groudine M. 1999. Looping versus linking: toward a model for long-distance gene activation. *Genes Dev* 13: 2465-2477.
Bulger M. and Groudine M. 2002. TRAPping enhancer function. *Nat Genet* 32: 555-556.
Bulger M., Sawado T., Schubeler D., and Groudine M. 2002. ChIPs of the beta-globin locus: unraveling gene regulation within an active domain. *Curr Opin Genet Dev* 12: 170-177.
Bulger M., Schubeler D., Bender M.A., Hamilton J., Farrell C.M., Hardison R.C., and Groudine M. 2003. A complex chromatin landscape revealed by patterns of nuclease sensitivity and histone modification within the mouse beta-globin locus. *Mol Cell Biol* 23: 5234-5244.
Bulger M., van Doorninck J.H., Saitoh N., Telling A., Farrell C., Bender M.A., Felsenfeld G., Axel R., Groudine M., and von Doorninck J.H. 1999. Conservation of sequence and structure flanking the mouse and human beta-globin loci: the beta-globin genes are embedded within an array of odorant receptor genes. *Proc Natl Acad Sci U S A* 96: 5129-5134.
Bungert J., Dave U., Lim K.C., Lieuw K.H., Shavit J.A., Liu Q., and Engel J.D. 1995. Synergistic regulation of human beta-globin gene switching by locus control region elements HS3 and HS4. *Genes Dev* 9: 3083-3096.
Burgess-Beusse B., Farrell C., Gaszner M., Litt M., Mutskov V., Recillas-Targa F., Simpson M., West A., and Felsenfeld G. 2002. The insulation of genes from external enhancers and silencing chromatin. *Proc Natl Acad Sci U S A* 99 Suppl 4: 16433-16437.
Burke T.W., Willy P.J., Kutach A.K., Butler J.E., and Kadonaga J.T. 1998. The DPE, a conserved downstream core promoter element that is functionally analogous to the TATA box. *Cold Spring Harb Symp Quant Biol* 63: 75-82.
Butler J.E. and Kadonaga J.T. 2001. Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. *Genes Dev* 15: 2515-2519.
Byrd K. and Corces V.G. 2003. Visualization of chromatin domains created by the gypsy insulator of Drosophila. *J Cell Biol* 162: 565-574.
Cai H.N. and Shen P. 2001. Effects of cis arrangement of chromatin insulators on enhancer-blocking activity. *Science* 291: 493-495.
Cai H.N., Zhang Z., Adams J.R., and Shen P. 2001. Genomic context modulates insulator activity through promoter competition. *Development* 128: 4339-4347.
Calhoun V.C., Stathopoulos A., and Levine M. 2002. Promoter-proximal tethering elements regulate enhancer-promoter specificity in the Drosophila Antennapedia complex. *Proc Natl Acad Sci U S A* 99: 9243-9247.
Calzolari R., McMorrow T., Yannoutsos N., Langeveld A., and Grosveld F. 1999. Deletion of a region that is a candidate for the difference between the deletion forms of hereditary persistence of fetal hemoglobin and deltabeta-thalassemia affects beta- but not gamma-globin gene expression. *Embo J* 18: 949-958.
Cao S.X., Gutman P.D., Dave H.P., and Schechter A.N. 1989a. Identification of a transcriptional silencer in the 5'-flanking region of the human epsilon-globin gene. *Proc Natl Acad Sci U S A* 86: 5306-5309.
Cao S.X., Gutman P.D., Dave H.P., and Schechter A.N. 1989b. Negative control of the human epsilon-globin gene. *Prog Clin Biol*

Res 316A: 279-289.
Carmo-Fonseca M. 2002. The contribution of nuclear compartmentalization to gene regulation. *Cell* 108: 513-521.
Carmo-Fonseca M., Mendes-Soares L., and Campos I. 2000. To be or not to be in the nucleolus. *Nat Cell Biol* 2: E107-112.
Carter D., Chakalova L., Osborne C.S., Dai Y.F., and Fraser P. 2002. Long-range chromatin regulatory interactions in vivo. *Nat Genet* 32: 623-626.
Caterina J.J., Ciavatta D.J., Donze D., Behringer R.R., and Townes T.M. 1994. Multiple elements in human beta-globin locus control region 5' HS 2 are involved in enhancer activity and position-independent, transgene expression. *Nucleic Acids Res* 22: 1006-1011.
Chada K., Magram J., and Costantini F. 1986. An embryonic pattern of expression of a human fetal globin gene in transgenic mice. *Nature* 319: 685-689.
Chambeyron S. and Bickmore W.A. 2004a. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. *Genes Dev* 18: 1119-1130.
Chambeyron S. and Bickmore W.A. 2004b. Does looping and clustering in the nucleus regulate gene expression? *Curr Opin Cell Biol* 16: 256-262.
Cheutin T., McNairn A.J., Jenuwein T., Gilbert D.M., Singh P.B., and Misteli T. 2003. Maintenance of stable heterochromatin domains by dynamic HP1 binding. *Science* 299: 721-725.
Chong S., Riggs A.D., and Bonifer C. 2002. The chicken lysozyme chromatin domain contains a second, widely expressed gene. *Nucleic Acids Res* 30: 463-467.
Chubb J.R. and Bickmore W.A. 2003. Considering nuclear compartmentalization in the light of nuclear dynamics. *Cell* 112: 403-406.
Chubb J.R., Boyle S., Perry P., and Bickmore W.A. 2002. Chromatin motion is constrained by association with nuclear compartments in human cells. *Curr Biol* 12: 439-445.
Clapier C.R., Langst G., Corona D.F., Becker P.B., and Nightingale K.P. 2001. Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. *Mol Cell Biol* 21: 875-883.
Collins F.S., Metherall J.E., Yamakawa M., Pan J., Weissman S.M., and Forget B.G. 1985. A point mutation in the A gamma-globin gene promoter in Greek hereditary persistence of fetal haemoglobin. *Nature* 313: 325-326.
Collis P., Antoniou M., and Grosveld F. 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.R. 1999. The organization of replication and transcription. *Science* 284: 1790-1795.
Cook P.R. 2002. Predicting three-dimensional genome structure from transcriptional activity. *Nat Genet* 32: 347-352.
Cook P.R. 2003. Nongenic transcription, gene regulation and action at a distance. *J Cell Sci* 116: 4483-4491.
Copenhaver G.P., Nickel K., Kuromori T., Benito M.I., Kaul S., Lin X., Bevan M., Murphy G., Harris B., Parnell L.D., McCombie W.R., Martienssen R.A., Marra M., and Preuss D. 1999. Genetic definition and sequence analysis of Arabidopsis centromeres. *Science* 286: 2468-2474.
Cremer T. and Cremer C. 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nat Rev Genet* 2: 292-301.
Cremer T., Kreth G., Koester H., Fink R.H., Heintzmann R., Cremer M., Solovei I., Zink D., and Cremer C. 2000. Chromosome territories, interchromatin domain compartment, and nuclear matrix: an integrated view of the functional nuclear architecture. *Crit Rev Eukaryot Gene Expr* 10: 179-212.
Crossley M., Tsang A.P., Bieker J.J., and Orkin S.H. 1994. Regulation of the erythroid Kruppel-like factor (EKLF) gene promoter by the erythroid transcription factor GATA-1. *J Biol Chem* 269: 15440-15444.
Crotta S., Nicolis S., Ronchi A., Ottolenghi S., Ruzzi L., Shimada Y., Migliaccio A.R., and Migliaccio G. 1990. Progressive inactivation of the expression of an erythroid transcriptional factor in GM- and G-CSF-dependent myeloid cell lines. *Nucleic Acids Res* 18: 6863-6869.
de Bruin D., Zaman Z., Liberatore R.A., and Ptashne M. 2001. Telomere looping permits gene activation by a downstream UAS in yeast. *Nature* 409: 109-113.
de Laat W. and Grosveld F. 2003. Spatial organization of gene expression: the Active Chromatin Hub. *Chromosome Res* 5: 447-459.
de Villiers J., Olson L., Banerji J., and Schaffner W. 1983. Analysis of the transcriptional enhancer effect. *Cold Spring Harb Symp Quant Biol* 47 Pt 2: 911-919.
Dekker J., Rippe K., Dekker M., and Kleckner N. 2002. Capturing chromosome conformation. *Science* 295: 1306-1311.
Delassus S., Titley I., and Enver T. 1999. Functional and molecular analysis of hematopoietic progenitors derived from the aorta-gonad-mesonephros region of the mouse embryo. *Blood* 94: 1495-1503.
Devlin R.H., Bingham B., and Wakimoto B.T. 1990. The organization and expression of the light gene, a heterochromatic gene of Drosophila melanogaster. *Genetics* 125: 129-140.
Dickinson L.A., Joh T., Kohwi Y., and Kohwi-Shigematsu T. 1992. A tissue-specific MAR/SAR DNA-binding protein with unusual binding site recognition. *Cell* 70: 631-645.
Dieci G. and Sentenac A. 2003. Detours and shortcuts to transcription reinitiation. *Trends Biochem Sci* 28: 202-209.
Dietzel S., Schiebel K., Little G., Edelmann P., Rappold G.A., Eils R., Cremer C., and Cremer T. 1999. The 3D positioning of ANT2 and ANT3 genes within female X chromosome territories correlates with gene activity. *Exp Cell Res* 252: 363-375.
Dillon N. and Grosveld F. 1991. Human gamma-globin genes silenced independently of other genes in the beta-globin locus. *Nature* 350: 252-254.
Dillon N. and Sabbattini P. 2000. Functional gene expression domains: defining the functional unit of eukaryotic gene regulation. *Bioessays* 22: 657-665.
Dillon N., Trimborn T., Strouboulis J., Fraser P., and Grosveld F. 1997. The effect of distance on long-range chromatin interactions. *Mol Cell* 1: 131-139.

Dolznig H., Boulme F., Stangl K., Deiner E.M., Mikulits W., Beug H., and Mullner E.W. 2001. Establishment of normal, terminally differentiating mouse erythroid progenitors: molecular characterization by cDNA arrays. *Faseb J* 15: 1442-1444.

Dorsett D. 1999. Distant liaisons: long-range enhancer-promoter interactions in Drosophila. *Curr Opin Genet Dev* 9: 505-514.

Driscoll M.C., Dobkin C.S., and Alter B.P. 1989. Gamma delta beta-thalassemia due to a de novo mutation deleting the 5' beta-globin gene activation-region hypersensitive sites. *Proc Natl Acad Sci U S A* 86: 7470-7474.

Drissen R., Palstra R.-J., Gillemans N., Splinter E., Grosveld F., Philipsen S., and de Laat W. 2004. The active spatial organization of the {beta}-globin locus requires the transcription factor EKLF. *Genes Dev.* 18: 2485-2490.

Droge P. and Muller-Hill B. 2001. High local protein concentrations at promoters: strategies in prokaryotic and eukaryotic cells. *Bioessays* 23: 179-183.

Dunaway M. and Droge P. 1989. Transactivation of the Xenopus rRNA gene promoter by its enhancer. *Nature* 341: 657-659.

Dundr M. and Misteli T. 2001. Functional architecture in the cell nucleus. *Biochem J* 356: 297-310.

Eberl D.F., Duyf B.J., and Hilliker A.J. 1993. The role of heterochromatin in the expression of a heterochromatic gene, the rolled locus of Drosophila melanogaster. *Genetics* 134: 277-292.

Eivazova E.R. and Aune T.M. 2004. Dynamic alterations in the conformation of the Ifng gene region during T helper cell differentiation. *Proc Natl Acad Sci U S A* 101: 251-256.

Ellis J., Talbot D., Dillon N., and Grosveld F. 1993. Synthetic human beta-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 Grosveld F. 1996. A dominant chromatin-opening activity in 5' hypersensitive site 3 of the human beta-globin locus control region. *Embo J* 15: 562-568.

Enver T., Raich N., Ebens A.J., Papayannopoulou T., Costantini F., and Stamatoyannopoulos G. 1990. Developmental regulation of human fetal-to-adult globin gene switching in transgenic mice. *Nature* 344: 309-313.

Epner E., Reik A., Cimbora D., Telling A., Bender M.A., Fiering S., Enver T., Martin D.I., Kennedy M., Keller G., and Groudine M. 1998. The beta-globin LCR is not necessary for an open chromatin structure or developmentally regulated transcription of the native mouse beta-globin locus. *Mol Cell* 2: 447-455.

Farrell C.M., Grinberg A., Huang S.P., Chen D., Pichel J.G., Westphal H., and Felsenfeld G. 2000. A large upstream region is not necessary for gene expression or hypersensitive site formation at the mouse beta -globin locus. *Proc Natl Acad Sci U S A* 97: 14554-14559.

Farrell C.M., West A.G., and Felsenfeld G. 2002. Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol Cell Biol* 22: 3820-3831.

Feng W.C., Southwood C.M., and Bieker J.J. 1994. Analyses of beta-thalassemia mutant DNA interactions with erythroid Kruppel-like factor (EKLF), an erythroid cell-specific transcription factor. *J Biol Chem* 269: 1493-1500.

Festenstein R., Tolaini M., Corbella P., Mamalaki C., Parrington J., Fox M., Miliou A., Jones M., and Kioussis D. 1996. Locus control region function and heterochromatin-induced position effect variegation. *Science* 271: 1123-1125.

Fiering S., Epner E., Robinson K., Zhuang Y., Telling A., Hu M., Martin D.I., Enver T., Ley T.J., and Groudine M. 1995. Targeted deletion of 5'HS2 of the murine beta-globin LCR reveals that it is not essential for proper regulation of the beta-globin locus. *Genes Dev* 9: 2203-2213.

Flint J., Tufarelli C., Peden J., Clark K., Daniels R.J., Hardison R., Miller W., Philipsen S., Tan-Un K.C., McMorrow T., Frampton J., Alter B.P., Frischauf A.M., and Higgs D.R. 2001. Comparative genome analysis delimits a chromosomal domain and identifies key regulatory elements in the alpha globin cluster. *Hum Mol Genet* 10: 371-382.

Forrester W.C., Epner E., Driscoll M.C., Enver T., Brice M., Papayannopoulou T., and Groudine M. 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. *Genes Dev* 4: 1637-1649.

Forrester W.C., Thompson C., Elder J.T., and Groudine M. 1986. A developmentally stable chromatin structure in the human beta-globin gene cluster. *Proc Natl Acad Sci U S A* 83: 1359-1363.

Forsberg E.C. and Bresnick E.H. 2001. Histone acetylation beyond promoters: long-range acetylation patterns in the chromatin world. *Bioessays* 23: 820-830.

Forsberg E.C., Downs K.M., Christensen H.M., Im H., Nuzzi P.A., and Bresnick E.H. 2000. Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. *Proc Natl Acad Sci U S A* 97: 14494-14499.

Francastel C., Magis W., and Groudine M. 2001. Nuclear relocation of a transactivator subunit precedes target gene activation. *Proc Natl Acad Sci U S A* 98: 12120-12125.

Francastel C., Walters M.C., Groudine M., and Martin D.I. 1999. A functional enhancer suppresses silencing of a transgene and prevents its localization close to centrometric heterochromatin. *Cell* 99: 259-269.

Fraser P. and Grosveld F. 1998. Locus control regions, chromatin activation and transcription. *Curr Opin Cell Biol* 10: 361-365.

Fraser P., Hurst J., Collis P., and Grosveld F. 1990. DNaseI hypersensitive sites 1, 2 and 3 of the human beta-globin dominant control region direct position-independent expression. *Nucleic Acids Res* 18: 3503-3508.

Fraser P., Pruzina S., Antoniou M., and Grosveld F. 1993. Each hypersensitive site of the human beta-globin locus control region confers a different developmental pattern of expression on the globin genes. *Genes Dev* 7: 106-113.

Fujiwara Y., Browne C.P., Cunniff K., Goff S.C., and Orkin S.H. 1996. Arrested development of embryonic red cell precursors in mouse embryos lacking transcription factor GATA-1. *Proc Natl Acad Sci U S A* 93: 12355-12358.

Fuxa M., Skok J., Souabni A., Salvagiotto G., Roldan E., and Busslinger M. 2004. Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev* 18: 411-422.

Gaensler K.M., Zhang Z., Lin C., Yang S., Hardt K., and Flebbe-Rehwaldt L. 2003. Sequences in the (A)gamma-delta intergenic region are not required for stage-specific regulation of the human beta-globin gene locus. *Proc Natl Acad Sci U S A* 100: 3374-3379.

Gall J.G. 2000. Cajal bodies: the first 100 years. *Annu Rev Cell Dev Biol* 16: 273-300.

Gaszner M., Vazquez J., and Schedl P. 1999. The Zw5 protein, a component of the scs chromatin domain boundary, is able to block

enhancer-promoter interaction. *Genes Dev* 13: 2098-2107.
Gause M., Morcillo P., and Dorsett D. 2001. Insulation of enhancer-promoter communication by a gypsy transposon insert in the Drosophila cut gene: cooperation between suppressor of hairy-wing and modifier of mdg4 proteins. *Mol Cell Biol* 21: 4807-4817.
Gelles J. and Landick R. 1998. RNA polymerase as a molecular motor. *Cell* 93: 13-16.
Gerasimova T.I., Byrd K., and Corces V.G. 2000. A chromatin insulator determines the nuclear localization of DNA. *Mol Cell* 6: 1025-1035.
Gerasimova T.I. and Corces V.G. 2001. Chromatin insulators and boundaries: effects on transcription and nuclear organization. *Annu Rev Genet* 35: 193-208.
Geyer P.K. 1997. The role of insulator elements in defining domains of gene expression. *Curr Opin Genet Dev* 7: 242-248.
Geyer P.K. and Clark I. 2002. Protecting against promiscuity: the regulatory role of insulators. *Cell Mol Life Sci* 59: 2112-2127.
Giglioni B., Casini C., Mantovani R., Merli S., Comi P., Ottolenghi S., Saglio G., Camaschella C., and Mazza U. 1984. A molecular study of a family with Greek hereditary persistence of fetal hemoglobin and beta-thalassemia. *Embo J* 3: 2641-2645.
Gilbert N., Boyle S., Fiegler H., Woodfine K., Carter N.P., and Bickmore W.A. 2004. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. *Cell* 118: 555-566.
Gillemans N., Tewari R., Lindeboom F., Rottier R., de Wit T., Wijgerde M., Grosveld F., and Philipsen S. 1998. Altered DNA-binding specificity mutants of EKLF and Sp1 show that EKLF is an activator of the beta-globin locus control region in vivo. *Genes Dev* 12: 2863-2873.
Goodwin A.J., McInerney J.M., Glander M.A., Pomerantz O., and Lowrey C.H. 2001. In vivo formation of a human beta-globin locus control region core element requires binding sites for multiple factors including GATA-1, NF-E2, erythroid Kruppel-like factor, and Sp1. *J Biol Chem* 276: 26883-26892.
Gribnau J., de Boer E., Trimborn T., Wijgerde M., Milot E., Grosveld F., and Fraser P. 1998. Chromatin interaction mechanism of transcriptional control in vivo. *Embo J* 17: 6020-6027.
Gribnau J., Diderich K., Pruzina S., Calzolari R., and Fraser P. 2000. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. *Mol Cell* 5: 377-386.
Grosveld F. 1999. Activation by locus control regions? *Curr Opin Genet Dev* 9: 152-157.
Grosveld F., van Assendelft G.B., Greaves D.R., and Kollias G. 1987. Position-independent, high-level expression of the human beta-globin gene in transgenic mice. *Cell* 51: 975-985.
Groudine M., Eisenman R., Gelinas R., and Weintraub H. 1983. Developmental aspects of chromatin structure and gene expression. *Prog Clin Biol Res* 134: 159-182.
Groudine M. and Weintraub H. 1982. Propagation of globin DNAase I-hypersensitive sites in absence of factors required for induction: a possible mechanism for determination. *Cell* 30: 131-139.
Guarente L. 1988. UASs and enhancers: common mechanism of transcriptional activation in yeast and mammals. *Cell* 52: 303-305.
Guy L.G., Kothary R., DeRepentigny Y., Delvoye N., Ellis J., and Wall L. 1996. The beta-globin locus control region enhances transcription of but does not confer position-independent expression onto the lacZ gene in transgenic mice. *Embo J* 15: 3713-3721.
Hancock R. 2000. A new look at the nuclear matrix. *Chromosoma* 109: 219-225.
Hanscombe O., Whyatt D., Fraser P., Yannoutsos N., Greaves D., Dillon N., and Grosveld F. 1991. Importance of globin gene order for correct developmental expression. *Genes Dev* 5: 1387-1394.
Hansen J.C. 2002. Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. *Annu Rev Biophys Biomol Struct* 31: 361-392.
Hardison R. and Miller W. 1993. Use of long sequence alignments to study the evolution and regulation of mammalian globin gene clusters. *Mol Biol Evol* 10: 73-102.
Hardison R., Slightom J.L., Gumucio D.L., Goodman M., Stojanovic N., and Miller W. 1997. Locus control regions of mammalian beta-globin gene clusters: combining phylogenetic analyses and experimental results to gain functional insights. *Gene* 205: 73-94.
Harju S., McQueen K.J., and Peterson K.R. 2002. Chromatin structure and control of beta-like globin gene switching. *Exp Biol Med (Maywood)* 227: 683-700.
Hark A.T., Schoenherr C.J., Katz D.J., Ingram R.S., Levorse J.M., and Tilghman S.M. 2000. CTCF mediates methylation-sensitive enhancer-blocking activity at the H19/Igf2 locus. *Nature* 405: 486-489.
Hatzis P. and Talianidis I. 2002. Dynamics of enhancer-promoter communication during differentiation-induced gene activation. *Mol Cell* 10: 1467-1477.
Henikoff S., Furuyama T., and Ahmad K. 2004. Histone variants, nucleosome assembly and epigenetic inheritance. *Trends Genet* 20: 320-326.
Herendeen D.R., Kassavetis G.A., and Geiduschek E.P. 1992. A transcriptional enhancer whose function imposes a requirement that proteins track along DNA. *Science* 256: 1298-1303.
Heuchel R., Matthias P., and Schaffner W. 1989. Two closely spaced promoters are equally activated by a remote enhancer: evidence against a scanning model for enhancer action. *Nucleic Acids Res* 17: 8931-8947.
Higgs D.R. 1998. Do LCRs open chromatin domains? *Cell* 95: 299-302.
Hirasawa R., Shimizu R., Takahashi S., Osawa M., Takayanagi S., Kato Y., Onodera M., Minegishi N., Yamamoto M., Fukao K., Taniguchi H., Nakauchi H., and Iwama A. 2002. Essential and instructive roles of GATA factors in eosinophil development. *J Exp Med* 195: 1379-1386.
Holaska J.M., Kowalski A.K., and Wilson K.L. 2004. Emerin caps the pointed end of actin filaments: evidence for an actin cortical network at the nuclear inner membrane. *PLoS Biol* 2: E231.
Hoogstraten D., Nigg A.L., Heath H., Mullenders L.H., van Driel R., Hoeijmakers J.H., Vermeulen W., and Houtsmuller A.B. 2002. Rapid switching of TFIIH between RNA polymerase I and II transcription and DNA repair in vivo. *Mol Cell* 10: 1163-

1174.
Horn P.J. and Peterson C.L. 2002. Molecular biology. Chromatin higher order folding--wrapping up transcription. *Science* 297: 1824-1827.
Hu M., Krause D., Greaves M., Sharkis S., Dexter M., Heyworth C., and Enver T. 1997. Multilineage gene expression precedes commitment in the hemopoietic system. *Genes Dev* 11: 774-785.
Hug B.A., Wesselschmidt R.L., Fiering S., Bender M.A., Epner E., Groudine M., and Ley T.J. 1996. Analysis of mice containing a targeted deletion of beta-globin locus control region 5' hypersensitive site 3. *Mol Cell Biol* 16: 2906-2912.
Iarovaia O.V., Bystritskiy A., Ravcheev D., Hancock R., and Razin S.V. 2004. Visualization of individual DNA loops and a map of loop domains in the human dystrophin gene. *Nucleic Acids Res* 32: 2079-2086.
Iborra F.J. and Cook P.R. 1998. The size of sites containing SR proteins in human nuclei. Problems associated with characterizing small structures by immunogold labeling. *J Histochem Cytochem* 46: 985-992.
Im H., Park C., Feng Q., Johnson K.D., Kiekhaefer C.M., Choi K., Zhang Y., and Bresnick E.H. 2003. Dynamic regulation of histone H3 methylated at lysine 79 within a tissue-specific chromatin domain. *J Biol Chem* 278: 18346-18352.
Imam A.M., Patrinos G.P., de Krom M., Bottardi S., Janssens R.J., Katsantoni E., Wai A.W., Sherratt D.J., and Grosveld F.G. 2000. Modification of human beta-globin locus PAC clones by homologous recombination in Escherichia coli. *Nucleic Acids Res* 28: E65.
Ishii K., Arib G., Lin C., Van Houwe G., and Laemmli U.K. 2002. Chromatin boundaries in budding yeast: the nuclear pore connection. *Cell* 109: 551-562.
Ishii K. and Laemmli U.K. 2003. Structural and dynamic functions establish chromatin domains. *Mol Cell* 11: 237-248.
Isogai Y. and Tjian R. 2003. Targeting genes and transcription factors to segregated nuclear compartments. *Curr Opin Cell Biol* 15: 296-303.
Jackson D.A. 2003. The anatomy of transcription sites. *Curr Opin Cell Biol* 15: 311-317.
Jackson D.A., Hassan A.B., Errington R.J., and Cook P.R. 1993. Visualization of focal sites of transcription within human nuclei. *Embo J* 12: 1059-1065.
Jackson D.A., Iborra F.J., Manders E.M., and Cook P.R. 1998. Numbers and organization of RNA polymerases, nascent transcripts, and transcription units in HeLa nuclei. *Mol Biol Cell* 9: 1523-1536.
Janicki S.M. and Spector D.L. 2003. Nuclear choreography: interpretations from living cells. *Curr Opin Cell Biol* 15: 149-157.
Janicki S.M., Tsukamoto T., Salghetti S.E., Tansey W.P., Sachidanandam R., Prasanth K.V., Ried T., Shav-Tal Y., Bertrand E., Singer R.H., and Spector D.L. 2004. From silencing to gene expression: real-time analysis in single cells. *Cell* 116: 683-698.
Jeffreys A.J. and Flavell R.A. 1977. The rabbit beta-globin gene contains a large large insert in the coding sequence. *Cell* 12: 1097-1108.
Jenuwein T. and Allis C.D. 2001. Translating the histone code. *Science* 293: 1074-1080.
Johnson K.D., Christensen H.M., Zhao B., and Bresnick E.H. 2001. Distinct mechanisms control RNA polymerase II recruitment to a tissue-specific locus control region and a downstream promoter. *Mol Cell* 8: 465-471.
Johnson K.D., Grass J.A., Boyer M.E., Kiekhaefer C.M., Blobel G.A., Weiss M.J., and Bresnick E.H. 2002. Cooperative activities of hematopoietic regulators recruit RNA polymerase II to a tissue-specific chromatin domain. *Proc Natl Acad Sci U S A* 99: 11760-11765.
Johnson K.D., Grass J.A., Park C., Im H., Choi K., and Bresnick E.H. 2003. Highly restricted localization of RNA polymerase II within a locus control region of a tissue-specific chromatin domain. *Mol Cell Biol* 23: 6484-6493.
Kellum R. and Schedl P. 1991. A position-effect assay for boundaries of higher order chromosomal domains. *Cell* 64: 941-950.
Kiekhaefer C.M., Grass J.A., Johnson K.D., Boyer M.E., and Bresnick E.H. 2002. Hematopoietic-specific activators establish an overlapping pattern of histone acetylation and methylation within a mammalian chromatin domain. *Proc Natl Acad Sci U S A* 99: 14309-14314.
Kielman M.F., Smits R., Devi T.S., Fodde R., and Bernini L.F. 1993. Homology of a 130-kb region enclosing the alpha-globin gene cluster, the alpha-locus controlling region, and two non-globin genes in human and mouse. *Mamm Genome* 4: 314-323.
Kim J., Kollhoff A., Bergmann A., and Stubbs L. 2003. Methylation-sensitive binding of transcription factor YY1 to an insulator sequence within the paternally expressed imprinted gene, Peg3. *Hum Mol Genet* 12: 233-245.
Kimura H. and Cook P.R. 2001. Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. *J Cell Biol* 153: 1341-1353.
Kimura H., Sugaya K., and Cook P.R. 2002. The transcription cycle of RNA polymerase II in living cells. *J Cell Biol* 159: 777-782.
Kioussis D., Vanin E., deLange T., Flavell R.A., and Grosveld F.G. 1983. Beta-globin gene inactivation by DNA translocation in gamma beta-thalassaemia. *Nature* 306: 662-666.
Kmita M., Kondo T., and Duboule D. 2000. Targeted inversion of a polar silencer within the HoxD complex re-allocates domains of enhancer sharing. *Nat Genet* 26: 451-454.
Kolesky S.E., Ouhammouch M., and Geiduschek E.P. 2002. The mechanism of transcriptional activation by the topologically DNA-linked sliding clamp of bacteriophage T4. *J Mol Biol* 321: 767-784.
Kollias G., Hurst J., deBoer E., and Grosveld F. 1987. The human beta-globin gene contains a downstream developmental specific enhancer. *Nucleic Acids Res* 15: 5739-5747.
Kollias G., Wrighton N., Hurst J., and Grosveld F. 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.
Kong S., Bohl D., Li C., and Tuan D. 1997. Transcription of the HS2 enhancer toward a cis-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. *Mol Cell Biol* 17: 3955-3965.
Kosak S.T. and Groudine M. 2002. The undiscovered country: chromosome territories and the organization of transcription. *Dev Cell* 2: 690-692.
Kosak S.T., Skok J.A., Medina K.L., Riblet R., Le Beau M.M., Fisher A.G., and Singh H. 2002. Subnuclear compartmentalization of

immunoglobulin loci during lymphocyte development. *Science* 296: 158-162.
Kulozik A.E., Bellan-Koch A., Bail S., Kohne E., and Kleihauer E. 1991. Thalassemia intermedia: moderate reduction of beta globin gene transcriptional activity by a novel mutation of the proximal CACCC promoter element. *Blood* 77: 2054-2058.
Kuras L., Borggrefe T., and Kornberg R.D. 2003. Association of the Mediator complex with enhancers of active genes. *Proc Natl Acad Sci U S A* 100: 13887-13891.
Kurz A., Lampel S., Nickolenko J.E., Bradl J., Benner A., Zirbel R.M., Cremer T., and Lichter P. 1996. Active and inactive genes localize preferentially in the periphery of chromosome territories. *J Cell Biol* 135: 1195-1205.
Kutach A.K. and Kadonaga J.T. 2000. The downstream promoter element DPE appears to be as widely used as the TATA box in Drosophila core promoters. *Mol Cell Biol* 20: 4754-4764.
Labrador M. and Corces V.G. 2002. Setting the boundaries of chromatin domains and nuclear organization. *Cell* 111: 151-154.
Laloraya S., Guacci V., and Koshland D. 2000. Chromosomal addresses of the cohesin component Mcd1p. *J Cell Biol* 151: 1047-1056.
Langst G. and Becker P.B. 2001. Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors. *J Cell Sci* 114: 2561-2568.
Leach K.M., Nightingale K., Igarashi K., Levings P.P., Engel J.D., Becker P.B., and Bungert J. 2001. Reconstitution of human beta-globin locus control region hypersensitive sites in the absence of chromatin assembly. *Mol Cell Biol* 21: 2629-2640.
Leach K.M., Vieira K.F., Kang S.H., Aslanian A., Teichmann M., Roeder R.G., and Bungert J. 2003. Characterization of the human beta-globin downstream promoter region. *Nucleic Acids Res* 31: 1292-1301.
Lee C.K., Shibata Y., Rao B., Strahl B.D., and Lieb J.D. 2004. Evidence for nucleosome depletion at active regulatory regions genome-wide. *Nat Genet* 36: 900-905.
Lee D.H. and Schleif R.F. 1989. In vivo DNA loops in araCBAD: size limits and helical repeat. *Proc Natl Acad Sci U S A* 86: 476-480.
Lee G.R., Fields P.E., Griffin T.J., and Flavell R.A. 2003. Regulation of the Th2 cytokine locus by a locus control region. *Immunity* 19: 145-153.
Lerner A., D'Adamio L., Diener A.C., Clayton L.K., and Reinherz E.L. 1993. CD3 zeta/eta/theta locus is colinear with and transcribed antisense to the gene encoding the transcription factor Oct-1. *J Immunol* 151: 3152-3162.
Li Q., Peterson K.R., Fang X., and Stamatoyannopoulos G. 2002. Locus control regions. *Blood* 100: 3077-3086.
Li Q. and Stamatoyannopoulos G. 1994. Hypersensitive site 5 of the human beta locus control region functions as a chromatin insulator. *Blood* 84: 1399-1401.
Lim S.K., Bieker J.J., Lin C.S., and Costantini F. 1997. A shortened life span of EKLF-/- adult erythrocytes, due to a deficiency of beta-globin chains, is ameliorated by human gamma-globin chains. *Blood* 90: 1291-1299.
Ling K.W. and Dzierzak E. 2002. Ontogeny and genetics of the hemato/lymphopoietic system. *Curr Opin Immunol* 14: 186-191.
Lippincott-Schwartz J., Snapp E., and Kenworthy A. 2001. Studying protein dynamics in living cells. *Nat Rev Mol Cell Biol* 2: 444-456.
Litt M.D., Simpson M., Recillas-Targa F., Prioleau M.N., and Felsenfeld G. 2001. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. *Embo J* 20: 2224-2235.
Littlewood T.D., Hancock D.C., Danielian P.S., Parker M.G., and Evan G.I. 1995. A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. *Nucleic Acids Res* 23: 1686-1690.
Liu Q., Bungert J., and Engel J.D. 1997. Mutation of gene-proximal regulatory elements disrupts human epsilon-, gamma-, and beta-globin expression in yeast artificial chromosome transgenic mice. *Proc Natl Acad Sci U S A* 94: 169-174.
Locke J., Kotarski M.A., and Tartof K.D. 1988. Dosage-dependent modifiers of position effect variegation in Drosophila and a mass action model that explains their effect. *Genetics* 120: 181-198.
Lu S.J., Rowan S., Bani M.R., and Ben-David Y. 1994. Retroviral integration within the Fli-2 locus results in inactivation of the erythroid transcription factor NF-E2 in Friend erythroleukemias: evidence that NF-E2 is essential for globin expression. *Proc Natl Acad Sci U S A* 91: 8398-8402.
Lundgren M., Chow C.M., Sabbattini P., Georgiou A., Minaee S., and Dillon N. 2000. Transcription factor dosage affects changes in higher order chromatin structure associated with activation of a heterochromatic gene. *Cell* 103: 733-743.
Magram J., Chada K., and Costantini F. 1985. Developmental regulation of a cloned adult beta-globin gene in transgenic mice. *Nature* 315: 338-340.
Mahmoudi T., Katsani K.R., and Verrijzer C.P. 2002. GAGA can mediate enhancer function in trans by linking two separate DNA molecules. *Embo J* 21: 1775-1781.
Mahy N.L., Perry P.E., and Bickmore W.A. 2002a. Gene density and transcription influence the localization of chromatin outside of chromosome territories detectable by FISH. *J Cell Biol* 159: 753-763.
Mahy N.L., Perry P.E., Gilchrist S., Baldock R.A., and Bickmore W.A. 2002b. Spatial organization of active and inactive genes and noncoding DNA within chromosome territories. *J Cell Biol* 157: 579-589.
Marshall W.F., Straight A., Marko J.F., Swedlow J., Dernburg A., Belmont A., Murray A.W., Agard D.A., and Sedat J.W. 1997. Interphase chromosomes undergo constrained diffusional motion in living cells. *Curr Biol* 7: 930-939.
Martin D.I. and Orkin S.H. 1990. Transcriptional activation and DNA binding by the erythroid factor GF-1/NF-E1/Eryf 1. *Genes Dev* 4: 1886-1898.
Mastrangelo I.A., Courey A.J., Wall J.S., Jackson S.P., and Hough P.V. 1991. DNA looping and Sp1 multimer links: a mechanism for transcriptional synergism and enhancement. *Proc Natl Acad Sci U S A* 88: 5670-5674.
Matthews K.S. and Nichols J.C. 1998. Lactose repressor protein: functional properties and structure. *Prog Nucleic Acid Res Mol Biol* 58: 127-164.
Maya-Mendoza A., Hernandez-Munoz R., Gariglio P., and Aranda-Anzaldo A. 2003. Gene positional changes relative to the nuclear substructure correlate with the proliferating status of hepatocytes during liver regeneration. *Nucleic Acids Res* 31: 6168-6179.

McKittrick E., Gafken P.R., Ahmad K., and Henikoff S. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proc Natl Acad Sci U S A* 101: 1525-1530.
McNagny K. and Graf T. 2002. Making eosinophils through subtle shifts in transcription factor expression. *J Exp Med* 195: F43-47.
McNally J.G., Muller W.G., Walker D., Wolford R., and Hager G.L. 2000. The glucocorticoid receptor: rapid exchange with regulatory sites in living cells. *Science* 287: 1262-1265.
Merika M. and Orkin S.H. 1995. Functional synergy and physical interactions of the erythroid transcription factor GATA-1 with the Kruppel family proteins Sp1 and EKLF. *Mol Cell Biol* 15: 2437-2447.
Mignotte V., Wall L., deBoer E., Grosveld F., and Romeo P.H. 1989. Two tissue-specific factors bind the erythroid promoter of the human porphobilinogen deaminase gene. *Nucleic Acids Res* 17: 37-54.
Miller I.J. and Bieker J.J. 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. *Mol Cell Biol* 13: 2776-2786.
Miller T., Krogan N.J., Dover J., Erdjument-Bromage H., Tempst P., Johnston M., Greenblatt J.F., and Shilatifard A. 2001. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A* 98: 12902-12907.
Milot E., Strouboulis J., Trimborn T., Wijgerde M., de Boer E., Langeveld A., Tan-Un K., Vergeer W., Yannoutsos N., Grosveld F., and Fraser P. 1996. Heterochromatin effects on the frequency and duration of LCR-mediated gene transcription. *Cell* 87: 105-114.
Misteli T. 2001a. The concept of self-organization in cellular architecture. *J Cell Biol* 155: 181-185.
Misteli T. 2001b. Protein dynamics: implications for nuclear architecture and gene expression. *Science* 291: 843-847.
Misteli T., Caceres J.F., and Spector D.L. 1997. The dynamics of a pre-mRNA splicing factor in living cells. *Nature* 387: 523-527.
Misteli T., Gunjan A., Hock R., Bustin M., and Brown D.T. 2000. Dynamic binding of histone H1 to chromatin in living cells. *Nature* 408: 877-881.
Mizzen C.A., Yang X.J., Kokubo T., Brownell J.E., Bannister A.J., Owen-Hughes T., Workman J., Wang L., Berger S.L., Kouzarides T., Nakatani Y., and Allis C.D. 1996. The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* 87: 1261-1270.
Moon A.M. and Ley T.J. 1990. Conservation of the primary structure, organization, and function of the human and mouse beta-globin locus-activating regions. *Proc Natl Acad Sci U S A* 87: 7693-7697.
Moon A.M. and Ley T.J. 1991. Functional properties of the beta-globin locus control region in K562 erythroleukemia cells. *Blood* 77: 2272-2284.
Morcillo P., Rosen C., Baylies M.K., and Dorsett D. 1997. Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in Drosophila. *Genes Dev* 11: 2729-2740.
Morris J.R., Geyer P.K., and Wu C.T. 1999. Core promoter elements can regulate transcription on a separate chromosome in trans. *Genes Dev* 13: 253-258.
Mueller-Storm H.P., Sogo J.M., and Schaffner W. 1989. An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. *Cell* 58: 767-777.
Muller A.M., Medvinsky A., Strouboulis J., Grosveld F., and Dzierzak E. 1994. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity* 1: 291-301.
Muller F. and Tora L. 2004. The multicoloured world of promoter recognition complexes. *Embo J* 23: 2-8.
Muller H.P., Matthias P., and Schaffner W. 1990. A transcriptional terminator between enhancer and promoter does not affect remote transcriptional control. *Somat Cell Mol Genet* 16: 351-360.
Muller H.P. and Schaffner W. 1990. Transcriptional enhancers can act in trans. *Trends Genet* 6: 300-304.
Muller W.G., Walker D., Hager G.L., and McNally J.G. 2001. Large-scale chromatin decondensation and recondensation regulated by transcription from a natural promoter. *J Cell Biol* 154: 33-48.
Muravyova E., Golovnin A., Gracheva E., Parshikov A., Belenkaya T., Pirrotta V., and Georgiev P. 2001. Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science* 291: 495-498.
Murrell A., Heeson S., and Reik W. 2004. Interaction between differentially methylated regions partitions the imprinted genes Igf2 and H19 into parent-specific chromatin loops. *Nat Genet* 36: 889-893.
Narlikar G.J., Fan H.Y., and Kingston R.E. 2002. Cooperation between complexes that regulate chromatin structure and transcription. *Cell* 108: 475-487.
Navas P.A., Peterson K.R., Li Q., Skarpidi E., Rohde A., Shaw S.E., Clegg C.H., Asano H., and Stamatoyannopoulos G. 1998. Developmental specificity of the interaction between the locus control region and embryonic or fetal globin genes in transgenic mice with an HS3 core deletion. *Mol Cell Biol* 18: 4188-4196.
Needham M., Gooding C., Hudson K., Antoniou M., Grosveld F., and Hollis M. 1992. LCR/MEL: a versatile system for high-level expression of heterologous proteins in erythroid cells. *Nucleic Acids Res* 20: 997-1003.
Ney P.A., Andrews N.C., Jane S.M., Safer B., Purucker M.E., Weremowicz S., Morton C.C., Goff S.C., Orkin S.H., and Nienhuis A.W. 1993. Purification of the human NF-E2 complex: cDNA cloning of the hematopoietic cell-specific subunit and evidence for an associated partner. *Mol Cell Biol* 13: 5604-5612.
Ney P.A., Sorrentino B.P., Lowrey C.H., and Nienhuis A.W. 1990a. Inducibility of the HS II enhancer depends on binding of an erythroid specific nuclear protein. *Nucleic Acids Res* 18: 6011-6017.
Ney P.A., Sorrentino B.P., McDonagh K.T., and Nienhuis A.W. 1990b. Tandem AP-1-binding sites within the human beta-globin dominant control region function as an inducible enhancer in erythroid cells. *Genes Dev* 4: 993-1006.
Nickerson J. 2001. Experimental observations of a nuclear matrix. *J Cell Sci* 114: 463-474.
Nuez B., Michalovich D., Bygrave A., Ploemacher R., and Grosveld F. 1995. Defective haematopoiesis in fetal liver resulting from inactivation of the EKLF gene. *Nature* 375: 316-318.
Ogbourne S. and Antalis T.M. 1998. Transcriptional control and the role of silencers in transcriptional regulation in eukaryotes.

Biochem J 331 (Pt 1): 1-14.
Ohtsuki S., Levine M., and Cai H.N. 1998. Different core promoters possess distinct regulatory activities in the Drosophila embryo. *Genes Dev* 12: 547-556.
Orkin S.H. 1992. GATA-binding transcription factors in hematopoietic cells. *Blood* 80: 575-581.
Ortiz B.D., Cado D., and Winoto A. 1999. A new element within the T-cell receptor alpha locus required for tissue-specific locus control region activity. *Mol Cell Biol* 19: 1901-1909.
Ortiz B.D., Harrow F., Cado D., Santoso B., and Winoto A. 2001. Function and factor interactions of a locus control region element in the mouse T cell receptor-alpha/Dad1 gene locus. *J Immunol* 167: 3836-3845.
Osborne C.S., Chakalova L., Brown K.E., Carter D., Horton A., Debrand E., Goyenechea B., Mitchell J.A., Lopes S., Reik W., and Fraser P. 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet* 36: 1065-1071.
O'Sullivan J.M., Tan-Wong S.M., Morillon A., Lee B., Coles J., Mellor J., and Proudfoot N.J. 2004. Gene loops juxtapose promoters and terminators in yeast. *Nat Genet* 36: 1014-1018.
Palstra R.J., Tolhuis B., Splinter E., Nijmeijer R., Grosveld F., and de Laat W. 2003. The beta-globin nuclear compartment in development and erythroid differentiation. *Nat Genet* 35: 190-194.
Parkhurst S.M., Harrison D.A., Remington M.P., Spana C., Kelley R.L., Coyne R.S., and Corces V.G. 1988. The Drosophila su(Hw) gene, which controls the phenotypic effect of the gypsy transposable element, encodes a putative DNA-binding protein. *Genes Dev* 2: 1205-1215.
Patrinos G.P., de Krom M., de Boer E., Langeveld A., Imam A.M., Strouboulis J., de Laat W., and Grosveld F.G. 2004. Multiple interactions between regulatory regions are required to stabilize an active chromatin hub. *Genes Dev* 18: 1495-1509.
Pederson T. 2000. Half a century of "the nuclear matrix". *Mol Biol Cell* 11: 799-805.
Perkins A. 1999. Erythroid Kruppel like factor: from fishing expedition to gourmet meal. *Int J Biochem Cell Biol* 31: 1175-1192.
Perkins A.C., Gaensler K.M., and Orkin S.H. 1996. Silencing of human fetal globin expression is impaired in the absence of the adult beta-globin gene activator protein EKLF. *Proc Natl Acad Sci U S A* 93: 12267-12271.
Perkins A.C., Sharpe A.H., and Orkin S.H. 1995. Lethal beta-thalassaemia in mice lacking the erythroid CACCC-transcription factor EKLF. *Nature* 375: 318-322.
Perry C. and Soreq H. 2002. Transcriptional regulation of erythropoiesis. Fine tuning of combinatorial multi-domain elements. *Eur J Biochem* 269: 3607-3618.
Peterson C.L. and Laniel M.A. 2004. Histones and histone modifications. *Curr Biol* 14: R546-551.
Peterson K.R. and Stamatoyannopoulos G. 1993. Role of gene order in developmental control of human gamma- and beta-globin gene expression. *Mol Cell Biol* 13: 4836-4843.
Pevny L., Simon M.C., Robertson E., Klein W.H., Tsai S.F., D'Agati V., Orkin S.H., and Costantini F. 1991. Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1. *Nature* 349: 257-260.
Phair R.D., Scaffidi P., Elbi C., Vecerova J., Dey A., Ozato K., Brown D.T., Hager G., Bustin M., and Misteli T. 2004. Global nature of dynamic protein-chromatin interactions in vivo: three-dimensional genome scanning and dynamic interaction networks of chromatin proteins. *Mol Cell Biol* 24: 6393-6402.
Plant K.E., Routledge S.J., and Proudfoot N.J. 2001. Intergenic transcription in the human beta-globin gene cluster. *Mol Cell Biol* 21: 6507-6514.
Pombo A., Jackson D.A., Hollinshead M., Wang Z., Roeder R.G., and Cook P.R. 1999. Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *Embo J* 18: 2241-2253.
Ptashne M. 1986. Gene regulation by proteins acting nearby and at a distance. *Nature* 322: 697-701.
Ptashne M. and Gann A. 1997. Transcriptional activation by recruitment. *Nature* 386: 569-577.
Ragoczy T., Telling A., Sawado T., Groudine M., and Kosak S.T. 2003. A genetic analysis of chromosome territory looping: diverse roles for distal regulatory elements. *Chromosome Res* 11: 513-525.
Raich N., Papayannopoulou T., Stamatoyannopoulos G., and Enver T. 1992. Demonstration of a human epsilon-globin gene silencer with studies in transgenic mice. *Blood* 79: 861-864.
Rea S., Eisenhaber F., O'Carroll D., Strahl B.D., Sun Z.W., Schmid M., Opravil S., Mechtler K., Ponting C.P., Allis C.D., and Jenuwein T. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406: 593-599.
Recillas-Targa F., Pikaart M.J., Burgess-Beusse B., Bell A.C., Litt M.D., West A.G., Gaszner M., and Felsenfeld G. 2002. Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc Natl Acad Sci U S A* 99: 6883-6888.
Reese J.C. 2003. Basal transcription factors. *Curr Opin Genet Dev* 13: 114-118.
Reik A., Telling A., Zitnik G., Cimbora D., Epner E., and Groudine M. 1998. The locus control region is necessary for gene expression in the human beta-globin locus but not the maintenance of an open chromatin structure in erythroid cells. *Mol Cell Biol* 18: 5992-6000.
Reitman M., Lee E., Westphal H., and Felsenfeld G. 1993. An enhancer/locus control region is not sufficient to open chromatin. *Mol Cell Biol* 13: 3990-3998.
Richards E.J. and Elgin S.C. 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108: 489-500.
Richmond T.J. and Davey C.A. 2003. The structure of DNA in the nucleosome core. *Nature* 423: 145-150.
Ringrose L., Chabanis S., Angrand P.O., Woodroofe C., and Stewart A.F. 1999. Quantitative comparison of DNA looping in vitro and in vivo: chromatin increases effective DNA flexibility at short distances. *Embo J* 18: 6630-6641.
Rippe K. 2001. Making contacts on a nucleic acid polymer. *Trends Biochem Sci* 26: 733-740.
Rollins R.A., Korom M., Aulner N., Martens A., and Dorsett D. 2004. Drosophila nipped-B protein supports sister chromatid cohesion and opposes the stromalin/Scc3 cohesion factor to facilitate long-range activation of the cut gene. *Mol Cell*

Biol 24: 3100-3111.
Rollins R.A., Morcillo P., and Dorsett D. 1999. Nipped-B, a Drosophila homologue of chromosomal adherins, participates in activation by remote enhancers in the cut and Ultrabithorax genes. *Genetics* 152: 577-593.
Ronchi A., Berry M., Raguz S., Imam A., Yannoutsos N., Ottolenghi S., Grosveld F., and Dillon N. 1996. Role of the duplicated CCAAT box region in gamma-globin gene regulation and hereditary persistence of fetal haemoglobin. *Embo J* 15: 143-149.
Rupley J.A. 1964. Comparison of Protein Structure in the Crystal and in Solution. I. The Tyrosyl Ionization of Crystalline Methemoglobin. *Biochemistry* 20: 1524-1529.
Sabatino D.E., Cline A.P., Gallagher P.G., Garrett L.J., Stamatoyannopoulos G., Forget B.G., and Bodine D.M. 1998. Substitution of the human beta-spectrin promoter for the human agamma-globin promoter prevents silencing of a linked human beta-globin gene in transgenic mice. *Mol Cell Biol* 18: 6634-6640.
Sabbattini P., Georgiou A., Sinclair C., and Dillon N. 1999. Analysis of mice with single and multiple copies of transgenes reveals a novel arrangement for the lambda5-VpreB1 locus control region. *Mol Cell Biol* 19: 671-679.
Santoso B., Ortiz B.D., and Winoto A. 2000. Control of organ-specific demethylation by an element of the T-cell receptor-alpha locus control region. *J Biol Chem* 275: 1952-1958.
Sawado T., Halow J., Bender M.A., and Groudine M. 2003. The beta -globin locus control region (LCR) functions primarily by enhancing the transition from transcription initiation to elongation. *Genes Dev* 17: 1009-1018.
Sawado T., Igarashi K., and Groudine M. 2001. Activation of beta-major globin gene transcription is associated with recruitment of NF-E2 to the beta-globin LCR and gene promoter. *Proc Natl Acad Sci U S A* 98: 10226-10231.
Schubeler D., Francastel C., Cimbora D.M., Reik A., Martin D.I., and Groudine M. 2000. Nuclear localization and histone acetylation: a pathway for chromatin opening and transcriptional activation of the human beta-globin locus. *Genes Dev* 14: 940-950.
Schubeler D., Groudine M., and Bender M.A. 2001. The murine beta-globin locus control region regulates the rate of transcription but not the hyperacetylation of histones at the active genes. *Proc Natl Acad Sci U S A* 98: 11432-11437.
Scott K.C., Taubman A.D., and Geyer P.K. 1999. Enhancer blocking by the Drosophila gypsy insulator depends upon insulator anatomy and enhancer strength. *Genetics* 153: 787-798.
Shang Y., Hu X., DiRenzo J., Lazar M.A., and Brown M. 2000. Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription. *Cell* 103: 843-852.
Shilatifard A. 2004. Transcriptional elongation control by RNA polymerase II: a new frontier. *Biochim Biophys Acta* 1677: 79-86.
Shivdasani R.A. and Orkin S.H. 1995. Erythropoiesis and globin gene expression in mice lacking the transcription factor NF-E2. *Proc Natl Acad Sci U S A* 92: 8690-8694.
Shivdasani R.A., Rosenblatt M.F., Zucker-Franklin D., Jackson C.W., Hunt P., Saris C.J., and Orkin S.H. 1995. Transcription factor NF-E2 is required for platelet formation independent of the actions of thrombopoietin/MGDF in megakaryocyte development. *Cell* 81: 695-704.
Sipos L., Mihaly J., Karch F., Schedl P., Gausz J., and Gyurkovics H. 1998. Transvection in the Drosophila Abd-B domain: extensive upstream sequences are involved in anchoring distant cis-regulatory regions to the promoter. *Genetics* 149: 1031-1050.
Smale S.T. 2001. Core promoters: active contributors to combinatorial gene regulation. *Genes Dev* 15: 2503-2508.
Spana C., Harrison D.A., and Corces V.G. 1988. The Drosophila melanogaster suppressor of Hairy-wing protein binds to specific sequences of the gypsy retrotransposon. *Genes Dev* 2: 1414-1423.
Spann T.P., Goldman A.E., Wang C., Huang S., and Goldman R.D. 2002. Alteration of nuclear lamin organization inhibits RNA polymerase II-dependent transcription. *J Cell Biol* 156: 603-608.
Spector D.L. 2003. The dynamics of chromosome organization and gene regulation. *Annu Rev Biochem* 72: 573-608.
Spilianakis C.G. and Flavell R.A. 2004. Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol* 5: 1017-1027.
Splinter E., Grosveld F., and de Laat W. 2004. 3C technology: analyzing the spatial organization of genomic loci in vivo. *Methods Enzymol* 375: 493-507.
Stamatoyannopoulos G. and Grosveld F. 2001. Hemoglobin switching. In *The molecular basis of blood diseases* (ed. G. Stamatoyannopoulos, P. Majerus, R. Perlmutter, and H. Varmus), pp. 135-182. W.B. Saunders, Philidelphia.
Stein G.S., Zaidi S.K., Braastad C.D., Montecino M., van Wijnen A.J., Choi J.Y., Stein J.L., Lian J.B., and Javed A. 2003. Functional architecture of the nucleus: organizing the regulatory machinery for gene expression, replication and repair. *Trends Cell Biol* 13: 584-592.
Strahl B.D. and Allis C.D. 2000. The language of covalent histone modifications. *Nature* 403: 41-45.
Strahl B.D., Grant P.A., Briggs S.D., Sun Z.W., Bone J.R., Caldwell J.A., Mollah S., Cook R.G., Shabanowitz J., Hunt D.F., and Allis C.D. 2002. Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol* 22: 1298-1306.
Strouboulis J., Dillon N., and Grosveld F. 1992. Developmental regulation of a complete 70-kb human beta-globin locus in transgenic mice. *Genes Dev* 6: 1857-1864.
Su W., Jackson S., Tjian R., and Echols H. 1991. DNA looping between sites for transcriptional activation: self-association of DNA-bound Sp1. *Genes Dev* 5: 820-826.
Sun J., Brand M., Zenke Y., Tashiro S., Groudine M., and Igarashi K. 2004. Heme regulates the dynamic exchange of Bach1 and NF-E2-related factors in the Maf transcription factor network. *Proc Natl Acad Sci U S A* 101: 1461-1466.
Svejstrup J.Q. 2004. The RNA polymerase II transcription cycle: cycling through chromatin. *Biochim Biophys Acta* 1677: 64-73.
Talbot D. and Grosveld F. 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: 1391-1398.
Talbot D., Philipsen S., Fraser P., and Grosveld F. 1990. Detailed analysis of the site 3 region of the human beta-globin dominant

control region. *Embo J* 9: 2169-2177.
Tanabe O., Katsuoka F., Campbell A.D., Song W., Yamamoto M., Tanimoto K., and Engel J.D. 2002. An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. *Embo J* 21: 3434-3442.
Tanimoto K., Liu Q., Bungert J., and Engel J.D. 1999. Effects of altered gene order or orientation of the locus control region on human beta-globin gene expression in mice. *Nature* 398: 344-348.
Tanimoto K., Liu Q., Grosveld F., Bungert J., and Engel J.D. 2000. Context-dependent EKLF responsiveness defines the developmental specificity of the human epsilon-globin gene in erythroid cells of YAC transgenic mice. *Genes Dev* 14: 2778-2794.
Tewari R., Gillemans N., Harper A., Wijgerde M., Zafarana G., Drabek D., Grosveld F., and Philipsen S. 1996. The human beta-globin locus control region confers an early embryonic erythroid-specific expression pattern to a basic promoter driving the bacterial lacZ gene. *Development* 122: 3991-3999.
Tewari R., Gillemans N., Wijgerde M., Nuez B., von Lindern M., Grosveld F., and Philipsen S. 1998. Erythroid Kruppel-like factor (EKLF) is active in primitive and definitive erythroid cells and is required for the function of 5'HS3 of the beta-globin locus control region. *Embo J* 17: 2334-2341.
Tilghman S.M., Tiemeier D.C., Seidman J.G., Peterlin B.M., Sullivan M., Maizel J.V., and Leder P. 1978. Intervening sequence of DNA identified in the structural portion of a mouse beta-globin gene. *Proc Natl Acad Sci U S A* 75: 725-729.
Tolhuis B., Palstra R.J., Splinter E., Grosveld F., and de Laat W. 2002. Looping and interaction between hypersensitive sites in the active beta-globin locus. *Mol Cell* 10: 1453-1465.
Torigoi E., Bennani-Baiti I.M., Rosen C., Gonzalez K., Morcillo P., Ptashne M., and Dorsett D. 2000. Chip interacts with diverse homeodomain proteins and potentiates bicoid activity in vivo. *Proc Natl Acad Sci U S A* 97: 2686-2691.
Trimborn T., Gribnau J., Grosveld F., and Fraser P. 1999. Mechanisms of developmental control of transcription in the murine alpha- and beta-globin loci. *Genes Dev* 13: 112-124.
Tsang A.P., Fujiwara Y., Hom D.B., and Orkin S.H. 1998. Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG. *Genes Dev* 12: 1176-1188.
Tsang A.P., Visvader J.E., Turner C.A., Fujiwara Y., Yu C., Weiss M.J., Crossley M., and Orkin S.H. 1997. FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation. *Cell* 90: 109-119.
Tuan D., Kong S., and Hu K. 1992. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. *Proc Natl Acad Sci U S A* 89: 11219-11223.
Tuan D., Solomon W., Li Q., and London I.M. 1985. The "beta-like-globin" gene domain in human erythroid cells. *Proc Natl Acad Sci U S A* 82: 6384-6388.
Tuan D.Y., Solomon W.B., Cavallesco R., Huang G., and London I.M. 1989a. Characterization of a human globin enhancer element. *Prog Clin Biol Res* 316A: 63-72.
Tuan D.Y., Solomon W.B., London I.M., and Lee D.P. 1989b. An erythroid-specific, developmental-stage-independent enhancer far upstream of the human "beta-like globin" genes. *Proc Natl Acad Sci U S A* 86: 2554-2558.
Tumbar T. and Belmont A.S. 2001. Interphase movements of a DNA chromosome region modulated by VP16 transcriptional activator. *Nat Cell Biol* 3: 134-139.
Tumbar T., Sudlow G., and Belmont A.S. 1999. Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. *J Cell Biol* 145: 1341-1354.
Turner B.M. 2002. Cellular memory and the histone code. *Cell* 111: 285-291.
Van de Corput M.P. and Grosveld F.G. 2001. Fluorescence in situ hybridization analysis of transcript dynamics in cells. *Methods* 25: 111-118.
Vazquez J., Belmont A.S., and Sedat J.W. 2001. Multiple regimes of constrained chromosome motion are regulated in the interphase Drosophila nucleus. *Curr Biol* 11: 1227-1239.
Verschure P.J., Van Der Kraan I., Enserink J.M., Mone M.J., Manders E.M., and Van Driel R. 2002. Large-scale chromatin organization and the localization of proteins involved in gene expression in human cells. *J Histochem Cytochem* 50: 1303-1312.
Verschure P.J., van der Kraan I., Manders E.M., Hoogstraten D., Houtsmuller A.B., and van Driel R. 2003. Condensed chromatin domains in the mammalian nucleus are accessible to large macromolecules. *EMBO Rep* 4: 861-866.
Verschure P.J., van Der Kraan I., Manders E.M., and van Driel R. 1999. Spatial relationship between transcription sites and chromosome territories. *J Cell Biol* 147: 13-24.
Vieira K.F., Levings P.P., Hill M.A., Crusselle V.J., Kang S.H., Engel J.D., and Bungert J. 2004. Recruitment of transcription complexes to the beta -globin gene locus in vivo and in vitro. *J Biol Chem*.
Visser A.E., Jaunin F., Fakan S., and Aten J.A. 2000. High resolution analysis of interphase chromosome domains. *J Cell Sci* 113 (Pt 14): 2585-2593.
Volpi E.V., Chevret E., Jones T., Vatcheva R., Williamson J., Beck S., Campbell R.D., Goldsworthy M., Powis S.H., Ragoussis J., Trowsdale J., and Sheer D. 2000. Large-scale chromatin organization of the major histocompatibility complex and other regions of human chromosome 6 and its response to interferon in interphase nuclei. *J Cell Sci* 113 (Pt 9): 1565-1576.
von Lindern M., Deiner E.M., Dolznig H., Parren-Van Amelsvoort M., Hayman M.J., Mullner E.W., and Beug H. 2001. Leukemic transformation of normal murine erythroid progenitors: v- and c-ErbB act through signaling pathways activated by the EpoR and c-Kit in stress erythropoiesis. *Oncogene* 20: 3651-3664.
Wadman I.A., Osada H., Grutz G.G., Agulnick A.D., Westphal H., Forster A., and Rabbitts T.H. 1997. The LIM-only protein Lmo2 is a bridging molecule assembling an erythroid, DNA-binding complex which includes the TAL1, E47, GATA-1 and Ldb1/NLI proteins. *Embo J* 16: 3145-3157.
Wai A.W.K., Gillemans N., Raguz-Bolognesi S., Pruzina S., Zafrana G., Meijer D., Philipsen S., and Grosveld F. 2003. HS5 of the human beta-globin locus control region: a developmental stage-specific border in erythroid cells. *EMBO J* 22.
Wall L., deBoer E., and Grosveld F. 1988. The human beta-globin gene 3' enhancer contains multiple binding sites for an erythroid-specific protein. *Genes Dev* 2: 1089-1100.

Wasylyk B. and Chambon P. 1983. Potentiator effect of the SV40 72-bp repeat on initiation of transcription from heterologous promoter elements. *Cold Spring Harb Symp Quant Biol* 47 Pt 2: 921-934.

Wasylyk B., Wasylyk C., Augereau P., and Chambon P. 1983. The SV40 72 bp repeat preferentially potentiates transcription starting from proximal natural or substitute promoter elements. *Cell* 32: 503-514.

Weatherall D.J. 2004. Thalassaemia: the long road from bedside to genome. *Nat Rev Genet* 5: 625-631.

Weiler K.S. and Wakimoto B.T. 1998. Chromosome rearrangements induce both variegated and reduced, uniform expression of heterochromatic genes in a development-specific manner. *Genetics* 149: 1451-1464.

Weintraub H. and Groudine M. 1976. Chromosomal subunits in active genes have an altered conformation. *Science* 193: 848-856.

Weiss M.J., Keller G., and Orkin S.H. 1994. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. *Genes Dev* 8: 1184-1197.

West A.G., Gaszner M., and Felsenfeld G. 2002. Insulators: many functions, many mechanisms. *Genes Dev* 16: 271-288.

Whyatt D.J., Karis A., Harkes I.C., Verkerk A., Gillemans N., Elefanty A.G., Vairo G., Ploemacher R., Grosveld F., and Philipsen S. 1997. The level of the tissue-specific factor GATA-1 affects the cell-cycle machinery. *Genes Funct* 1: 11-24.

Wijgerde M., Gribnau J., Trimborn T., Nuez B., Philipsen S., Grosveld F., and Fraser P. 1996. The role of EKLF in human beta-globin gene competition. *Genes Dev* 10: 2894-2902.

Wijgerde M., Grosveld F., and Fraser P. 1995. Transcription complex stability and chromatin dynamics in vivo. *Nature* 377: 209-213.

Williams R.R., Broad S., Sheer D., and Ragoussis J. 2002. Subchromosomal positioning of the epidermal differentiation complex (EDC) in keratinocyte and lymphoblast interphase nuclei. *Exp Cell Res* 272: 163-175.

Willoughby D.A., Vilalta A., and Oshima R.G. 2000. An Alu element from the K18 gene confers position-independent expression in transgenic mice. *J Biol Chem* 275: 759-768.

Winter R.B., Berg O.G., and von Hippel P.H. 1981. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor--operator interaction: kinetic measurements and conclusions. *Biochemistry* 20: 6961-6977.

Wolffe A.P. 1999. Architectural regulations and Hmg1. *Nat Genet* 22: 215-217.

Wong P.M., Chung S.W., Chui D.H., and Eaves C.J. 1986. Properties of the earliest clonogenic hemopoietic precursors to appear in the developing murine yolk sac. *Proc Natl Acad Sci U S A* 83: 3851-3854.

Wu C.T. and Morris J.R. 1999. Transvection and other homology effects. *Curr Opin Genet Dev* 9: 237-246.

Xin L., Liu D.P., and Ling C.C. 2003. A hypothesis for chromatin domain opening. *Bioessays* 25: 507-514.

Yasui D., Miyano M., Cai S., Varga-Weisz P., and Kohwi-Shigematsu T. 2002. SATB1 targets chromatin remodelling to regulate genes over long distances. *Nature* 419: 641-645.

Yoshida C., Tokumasu F., Hohmura K.I., Bungert J., Hayashi N., Nagasawa T., Engel J.D., Yamamoto M., Takeyasu K., and Igarashi K. 1999. Long range interaction of cis-DNA elements mediated by architectural transcription factor Bach1. *Genes Cells* 4: 643-655.

Yu W., Ginjala V., Pant V., Chernukhin I., Whitehead J., Docquier F., Farrar D., Tavoosidana G., Mukhopadhyay R., Kanduri C., Oshimura M., Feinberg A.P., Lobanenkov V., Klenova E., and Ohlsson R. 2004. Poly(ADP-ribosyl)ation regulates CTCF-dependent chromatin insulation. *Nat Genet* 36: 1105-1110.

Yudkovsky N., Ranish J.A., and Hahn S. 2000. A transcription reinitiation intermediate that is stabilized by activator. *Nature* 408: 225-229.

Yusufzai T.M. and Felsenfeld G. 2004. The 5'-HS4 chicken beta-globin insulator is a CTCF-dependent nuclear matrix-associated element. *Proc Natl Acad Sci U S A* 101: 8620-8624.

Yusufzai T.M., Tagami H., Nakatani Y., and Felsenfeld G. 2004. CTCF tethers an insulator to subnuclear sites, suggesting shared insulator mechanisms across species. *Mol Cell* 13: 291-298.

Zeng C., McNeil S., Pockwinse S., Nickerson J., Shopland L., Lawrence J.B., Penman S., Hiebert S., Lian J.B., van Wijnen A.J., Stein J.L., and Stein G.S. 1998. Intranuclear targeting of AML/CBFalpha regulatory factors to nuclear matrix-associated transcriptional domains. *Proc Natl Acad Sci U S A* 95: 1585-1589.

Zhao K., Hart C.M., and Laemmli U.K. 1995. Visualization of chromosomal domains with boundary element-associated factor BEAF-32. *Cell* 81: 879-889.

Zhou J., Ashe H., Burks C., and Levine M. 1999. Characterization of the transvection mediating region of the abdominal-B locus in Drosophila. *Development* 126: 3057-3065.

Zhou J. and Levine M. 1999. A novel cis-regulatory element, the PTS, mediates an anti-insulator activity in the Drosophila embryo. *Cell* 99: 567-575.

Zirbel R.M., Mathieu U.R., Kurz A., Cremer T., and Lichter P. 1993. Evidence for a nuclear compartment of transcription and splicing located at chromosome domain boundaries. *Chromosome Res* 1: 93-106.

Zong R.T., Das C., and Tucker P.W. 2000. Regulation of matrix attachment region-dependent, lymphocyte-restricted transcription through differential localization within promyelocytic leukemia nuclear bodies. *Embo J* 19: 4123-4133.

Summary

All multi-cellular organisms start as a single fertilised oocyte. During differentiation the number of cells is multiplied by means of cell division. These cells specialise into many different cell types like brain, blood and muscle cells. Yet, all these cells contain the same genetic information that is stored in billions of base pairs of DNA sequences, which are organised in large units called chromosomes. A DNA sequence that codes for a particular heritable trait (often a protein) is called a gene. The gene contentis the same in every cell (approximately 25.000 genes). The difference between the diverse cell types is therefore not dictated by the genomic make up of a cell but how this genomic makeup is used or to put it into other words, which genes are activated or repressed. The activation of genes is tightly regulated in time and space and a failure of proper transcriptional regulation results in all kinds of diseases e.g. cancer. Proper transcriptional activation of genes in higher eukaryotes involves a number of regulatory DNA elements; the promoter, localised close to the gene and frequently other elements like enhancers, which can be located quite a distance away from the actual gene that they control. One of the key questions in molecular biology is how these distant so-called *cis*-regulatory elements communicate over the vast distances with the promoter of a gene. Several mechanisms have been proposed for this communication: e.g. the looping model, the tracking model and the linking model. These models are discussed in detail in chapter 1.

A good example of a specialised cell in the mammalian system is the red blood cell or erythrocyte. The most abundant molecule in red blood cells is the oxygen and carbon dioxide carrying molecule haemoglobin, which consists of two α-globin proteins, two β-globin proteins and four heme molecules. The β-globin protein is encoded by genes organised in the β-globin locus. The mouse β-globin locus contains four β-like globin genes. Two of them, the embryonic β-globin genes, are expressed in the embryo while the other two, the adult β-globin genes, are active at a later stage during development and this continues into adult life. The β-like genes are arranged in order of their developmental expression from 5' to 3' along the DNA fibre. A crucial regulatory element named the locus control region (LCR) is located ~50 kb upstream of the adult βmajor promoter, and several distal upstream and downstream elements with unknown function exist.

Most of our knowledge on the organisation of chromatin in the nucleus comes from light- and fluorescent microscopy studies, and although spectacular advances have been made in this field, optical constraints set limits to what can be resolved. Thus, it is as yet not possible to visualise the structural organisation of a single gene locus that spans, for example, 200 kilobases of genomic DNA. Intricate structural organisations are to be expected at this level of resolution, for example in cases where enhancers or other transcriptional regulatory elements communicate with distant promoters located on the same chromatin fibre. As discussed in chapter 4, several new techniques are being developed to study the structural organisation of a single gene locus and one of them is Chromosome conformation capture (3C) technology. 3C-technology gives a measure of the proximity between two given DNA fragments in the nuclear space. This method was originally developed to analyse the conformation of chromosomes in yeast. Chapter 5 describes how this procedure was optimised to study the conformation of chromosomal regions in (much more complex) mammalian cells and how it

was used to analyse the spatial organisation of a 200 kb region spanning the mouse β-globin locus in expressing erythroid and non-expressing brain tissue.

In brain, the β-globin locus adopts a seemingly linear conformation. However, in erythroid cells, the β-globin LCR, located 40-60kb away from the active genes, comes in close spatial proximity with these genes. The intervening chromatin with inactive β-like globin genes loops out. Moreover, two distant elements, being 130 kb apart from each other, participate in these interactions. The data shows that long-range gene regulation *in vivo* involves spatial interactions between distal transcriptional elements, with intervening chromatin looping out and is consistent with the looping model. Such spatial clustering of transcriptional regulatory elements is referred to as an Active Chromatin Hub (ACH).

In chapter 6 the structure of the β-globin locus in primitive erythroid cells, when the embryonic globin genes are active and the adult genes are silent, was analysed. The results show that a core ACH containing the β-globin regulatory elements is developmentally conserved, while β-like globin genes switch their interaction with this cluster during development, which correlates with the switch in their transcriptional activity. The human β-globin locus introduced as a transgene in mice behaves in a similar way. In erythroid progenitors that are committed to, but do not yet express β-globin, only a subset of regulatory sites cluster and form a structure called a Chromatin Hub (CH). Upon erythroid differentiation, a fully functional ACH is formed, containing all regulatory elements and the gene that becomes activated.

The role of special proteins, called transcription factors, in β-globin ACH formation is addressed in chapter 7. EKLF is an erythroid-specific transcription factor essential for expression of the adult β-like genes. In fetal livers obtained from EKLF knockout mice a locus conformation very similar to that observed in erythroid progenitors (a CH) is found. An EKLF protein fused to a ligand-binding domain was introduced to exclude that this observation is due to a failure of these cells to activate EKLF dependent secondary pathways. This fusion protein resides in the cytoplasm and is only able to translocate to the nucleus and bind to their DNA recognition sequences after addition of the drug 4-hydroxy tamoxifen. If the EKLF fusion protein is allowed to enter the nucleus the ACH is again formed. The ACH is also formed when a drug that inhibits protein synthesis and thus blocks secondary pathways is added. This data shows that EKLF is needed for the progression to and/or stabilisation of a fully functional ACH while formation of a CH is EKLF independent.

In chapter 8 it is proposed that ACH formation is key to establishing gene expression in repressive chromatin and to maintain hypersensitivity at genomic sites. Clustering of regulatory DNA elements, mediated by the affinity between proteins bound to these sites, may explain how genes on overlapping loci establish independent expression patterns. Also, it is suggested that clustering of *cis*-regulatory elements and active genes is essential for the high transcription rate of the β-globin genes. Transcription factors move rapidly through the nucleus, but efficient transcription requires a high local concentration of these factors at the genomic site. Local accumulation of *trans*-acting factors can be obtained by spatial clustering of their cognate binding sites. Therefore it is postulated that the β-globin ACH can be seen as a nuclear compartment dedicated to efficient (RNA polymerase II-mediated) transcription of the β-globin genes, analogous to the nucleolus being a compartment dedicated to RNA polymerase I-mediated transcription of rRNA genes.

Samenvatting

Alle multi-cellulaire organismen beginnen als een enkele bevruchte eicel. Gedurende de differentiatie wordt het aantal cellen vermeerderd door middel van celdeling. De cellen specialiseren zich tevens in verschillende celtypes zoals hersen-, bloed- en spiercellen. Toch bevatten al deze cellen dezelfde genetische informatie dat is opgeslagen in miljoenen base paren DNA, welke georganiseerd zijn in grote elementen die chromosomen worden genoemd. Een DNA sequentie die codeert voor een bepaalde overerfbare eigenschap (vaak een eiwit) wordt een gen genoemd. Het gehalte aan genen is gelijk in elke cel (ongeveer 25.000 genen). Het verschil tussen de diverse celtypes wordt daarom niet bepaald door de genomische opbouw van een cel maar juist hoe deze genomische opbouw gebruikt wordt, of anders gezegd; welke genen geactiveerd of juist onderdrukt worden. De activatie van genen wordt strak gereguleerd qua plaats en tijd en een gebrek aan juiste transcriptionele regulatie resulteert vaak in allerlei ziekten zoals b.v. kanker. Bij de juiste transcriptionele activatie van genen in hogere eukaryoten zijn verschillende regulerende DNA elementen betrokken; de promoter wordt vlakbij het gen gevonden terwijl andere elementen zoals enhancers zeer ver van het te activeren gen kunnen liggen. Een van de belangrijkste vragen in de moleculaire biologie is hoe deze zogeheten *cis*-regulerende elementen over deze aanzienlijke afstanden met de promoter van een gen kunnen communiceren. Verschillende mechanismen zijn voorgesteld voor deze communicatie: o.a. het looping model, het tracking model en het linking model. Deze modellen worden in hoofdstuk 1 in detail besproken.

Een goed voorbeeld van een gespecialiseerde cel in zoogdieren is de rode bloed cel. Het meest voorkomende molecuul in rode bloedcellen is het zuurstof en kooldioxide transporterende hemoglobine dat opgebouwd is uit twee α-globine eiwitten, twee β-globine eiwitten en vier heem groepen. Het β-globine eiwit wordt gecodeerd door genen in het β-globine locus. Het muizen β-globine locus bevat vier β-achtige genen. Twee daarvan, de embryonale β-globine genen, komen in het embryo tot expressie terwijl de twee andere genen, de volwassen β-globine genen actief worden in een later stadium van de ontwikkeling. De β-achtige genen zijn langs de DNA streng gerangschikt in de volgorde van hun expressie. Een vitaal regulerend element, de locus control region (LCR) genoemd, ligt ongeveer 50 kb stroom opwaarts van het volwassen βmajor gen, en verschillende ver weg gelegen elementen met een onbekende functie zijn geïdentificeerd.

De meeste kennis over de organisatie van chromatine in de cel kern is verkregen via licht- en fluorescentie microscopie studies, maar ondanks dat er spectaculaire vooruitgang geboekt is in dit veld wordt de resolutie begrensd door optische limitaties. Daarom is het momenteel niet mogelijk om de structurele organisatie van een individueel gen locus, dat bijvoorbeeld 200 kilobasen genomisch DNA omspant, te visualiseren. Een ingewikkelde structurele organisatie op dit niveau van resolutie is echter te verwachten b.v. waar enhancers of andere regulerende elementen met ver weg gelegen promoters communiceren. Zoals in hoofdstuk 4 bediscussieerd wordt zijn er diverse nieuwe technieken ontwikkeld om de structurele organisatie van een individueel gen locus te bestuderen en één daarvan is Chromosome conformation capture (3C) technologie. Deze methode was oorspronkelijk ontwikkeld om de structuur van chromosomen in gist te bestuderen. Hoofdstuk 5 beschrijft hoe deze procedure geoptimaliseerd werd om de structurele organisatie van chromosomale regio's in (de veel complexere) zoogdier cellen te bestuderen en hoe het toegepast werd om een chromosomale regio van 200 kb die het muizen β-globine locus bevat in expresserende

rode bloedcellen en niet expresserende hersen cellen te bestuderen.

In hersencellen heeft het β-globine locus een schijnbaar lineaire structuur. Maar in rode bloedcellen komt de β-globine LCR ruimtelijk in de naburigheid van het actieve gen. Het tussen liggende chromatine met de inactieve β-achtige genen vormt een lus. Bovendien zijn twee ver weg gelegen regulerende elementen, die 130 kb van elkaar liggen, ook betrokken bij deze interacties. Deze observaties laten zien dat lange afstand gen regulatie *in vivo* ruimtelijke interacties tussen verschillende regulerende elementen omvat terwijl het tussen liggende chromatine een lus vormt. Dit is in overeenstemming met het looping model. Deze ruimtelijke clustering van transcriptie regulerende elementen wordt een Actieve Chromatine Hub (ACH) genoemd.

In hoofdstuk 6 wordt de structuur van het β-globine locus in primitieve rode bloed cellen, waar de embryonale genen actief zijn terwijl de volwassen genen niet actief zijn, geanalyseerd. De resultaten laten zien dat een kern van de ACH, die de regulerende β-globine elementen bevat, geconserveerd is terwijl de β-globine genen hun interactie wisselen gedurende de ontwikkeling en dit correleert met de wisseling in transcriptionele activiteit. Het humane β-globine locus, dat als transgen in de muis geïntroduceerd is, gedraagt zich op vergelijkbare wijze. In voorlopers van rode bloedcellen waarvan vaststaat dat ze β-globine gaan expresseren maar dat nog niet doen, clustert een deel van de regulatieve elementen en vormt een structuur die Chromatine Hub (CH) genoemd wordt. Na differentiatie tot rode bloed cel wordt een volledig functionele ACH gevormd dat alle regulerende elementen en het actieve gen bevat. De rol die speciale eiwitten, transcriptie factoren genoemd, spelen in de formatie van de β-globine ACH wordt behandeld in hoofdstuk 7. EKLF is een rode bloedcel specifieke transcriptie factor die essentieel is voor de expressie van de volwassen β-globine genen. In foetale levers van EKLF knock-out muizen wordt een structuur gevonden dat heel erg lijkt op de structuur die in de voorloper cellen is gevonden (een CH). Een EKLF eiwit gefuseerd met een ligand bindend domein werd geïntroduceerd om uit te sluiten dat deze observatie veroorzaakt wordt door het onvermogen van deze cellen om een secundair EKLF afhankelijk signaal te activeren. Dit fusie eiwit wordt vastgehouden in het cytoplasma en kan alleen als 4-hydroxy tamoxifen wordt toegevoegd naar de kern verhuizen. De ACH wordt dan weer gevormd en zelfs als tegelijkertijd de eiwit synthese geblokkeerd wordt, zodat secundaire effecten niet kunnen optreden, vindt ACH formatie plaats. Deze informatie laat zien dat EKLF nodig is voor de formatie en/of stabilisatie van een volledig functioneel ACH, en dat dit een primair effect is, terwijl CH formatie EKLF onafhankelijk is.

In hoofdstuk 8 wordt voorgesteld dat de formatie van een ACH de sleutel vormt tot gen expressie in onderdrukkend chromatine en om hypersensitiviteit op genomische locaties te behouden. Het clusteren van regulerende elementen, mogelijk gemaakt door de affiniteit tussen eiwitten die gebonden zijn aan deze elementen, kan verklaren hoe overlappende gen loci afzonderlijke expressie patronen tot stand brengen. Ook wordt gesuggereerd dat de clustering van de regulatieve elementen en actieve genen essentieel is voor de hoge transcriptie niveau van de β-globine genen. Transcriptie factoren bewegen zich snel door de kern terwijl efficiënte transcriptie een hoge locale concentratie van deze factoren op een genomische plek vereist. Locale accumulatie van *trans*-werkende factoren kan bewerkstelligd worden door de ruimtelijke clustering van hun bindings plaatsen. Daarom wordt gepostuleerd dat de β-globine ACH gezien kan worden als een kern compartiment dat toegewijd is aan (RNA polymerase II) transcriptie van de β-globine genen, zoals de nucleolus een kern compartiment is dat toegewijd is aan de RNA polymerase I transcriptie van rRNA genen.

Curriculum vitae

Name:	Robert-Jan Theo Sijtze Palstra
Born:	14 August 1968, ‘s-Gravenhage, The Netherlands
1981 –1984	MAVO, St Maartens College, Voorburg
1985 –1989	MLO, Van Leeuwenhoek instituut, Delft Analytical Chemistry
1989 –1995	HLO, Hogeschool Rotterdam en Omstreken, Delft Biochemistry
Scientific research:	
March 1997 - Feb 1998	Anatomisches Institut, Universtität Berne, Switserland ‘Cytoskeleton membrane interactions in Smooth Muscle Cells (Prof.dr. A. Draeger)
Mei 1998 – Sept 1999	Armauer Hansen Research Institute, Addis Ababa, Ethiopia ‘The role of the Ercc3 homologue in *Mycobacterium leprae*’ (Prof.dr. S. Britton)
Sept 1999 – Sept 2000	Molecular Medicine Unit, University of Leeds, UK ‘The influence of chicken lysozyme promoter mutations on the chromatin remodelling activity of its *cis*-regulatory elements’ (Prof.dr. C. Bonnifer)
Jan 2001 – Jan 2005	Cell biology, ErasmusMC, Rotterdam, The Netherlands ‘Long-range gene activation of the β-globin locus’ (Prof.dr. F. Grosveld and Dr. W. de Laat)

List of publications

Drissen R*, **Palstra RJ***, Gillemans N, Splinter E, Grosveld F, Philipsen S, and de Laat W.
The active spatial organization of the β-globin locus requires the transcription factor EKLF
Genes Dev. 2004 Oct; 18(20):2485-2490

Palstra RJ*, Tolhuis B*, Splinter E, Nijmeijer R, Grosveld F, de Laat W.
The β-globin nuclear compartment in development and erythroid differentiation.
Nat Genet. 2003 Oct; 35(2):190-194.

Tolhuis B*, **Palstra RJ***, Splinter E, Grosveld F, de Laat W.
Looping and interaction between hypersensitive sites in the active β-globin locus.
Mol Cell. 2002 Dec; 10(6):1453-1465.

Babiychuk EB, **Palstra RJ**, Schaller J, Kampfer U, Draeger A.
Annexin VI participates in the formation of a reversible, membrane-cytoskeleton complex
in smooth muscle cells.
J Biol Chem. 1999 Dec; 274(49):35191-35195.

* These authors contributed equally

Nawoord:

Vier jaar na aanvang van mijn AIO tijd ligt er nu een boekje en is er weer een avontuur voorbij. Time flies when you're having fun. Dit werk is echter geen één mans project en vele mensen hebben er op verschillende manieren aan bijgedragen aan wie ik allemaal mijn dank verschuldigd ben.

Allereerst moet ik natuurlijk mijn promotor Prof. Frank Grosveld bedanken. Frank, ik ben je zeer erkentelijk voor het feit dat je het hebt aangedurfd om met mij in zee te gaan. Het wetenschappelijk klimaat dat je weet te creëren is fantastisch en zeer inspirerend en ik heb het als een voorrecht beschouwt om in je groep te mogen werken. Aan mijn copromotor Dr. Wouter de Laat ben ik zeer veel dank verschuldigd. Beste Wouter, ik vind het heel tof dat je me de afgelopen vier jaar hebt willen begeleiden. Eerst nog met wat wilde long range activatie projecten maar al snel cumuleerde dit in een group effort om looping binnen het β-globine locus aan te tonen (en dat is gelukt!). De maanden van samenhorigheid en intense wetenschappelijke activiteit gedurende de zomer van 2002 zal ik niet licht vergeten en zal me altijd bijblijven als een schoolvoorbeeld van hoe mooi en spannend wetenschap kan zijn. Ik kijk er dan ook naar uit om in de komende jaren met je verder te werken.

Dit brengt me tot mijn andere partners in crime; Bas Tolhuis en Erik Splinter. Mannen, het was me een waar genoegen om met jullie samen te werken. De overstap van het navel staarderige werken aan mijn eigen kleine deel projectjes naar het symbiotische 3C-avontuur is me zeer goed bevallen. Erik, jou wil ik nog extra bedanken want zonder jouw gouden handjes en slimme kijk op de experimenten was dit boekje niet mogelijk geweest. Ik ben daarom erg blij dat je een van mijn paranimfen wilt zijn. Roy, ons laatste 3C project duurde wat langer dan gedacht maar heeft een heel mooi resultaat opgeleverd.

Jurgen, Daan, Petra en Marieke jullie zijn de nieuwe lichting en ik kijk uit naar onze samenwerking in het lab en alle mooie dingen waarvan ik zeker weet dat die nog komen gaan. Arnie, my old desk neighbour it is good fun having you around but I still haven't got the answer to the ultimate question "when is season 2 coming out?" I can't start mentioning all the other members of the department who contributed in different ways to this thesis (I might risk forgetting somebody) therefore to all of you: thank you very much!
Extra aandacht verdienen Marieke, Jasperina, Sjozef, Pim, Ton en Melle voor alle logistieke ondersteuning die altijd met veel enthousiasme gegeven wordt. Van de afdeling Hematologie wil ik Marieke von Linderen bedanken voor haar enthousiaste steun bij mijn onderzoek en haar bereidheid om altijd te helpen.

My time as a PhD student in Rotterdam had several prequels at different locations in the world, which have been instructional for my scientific career and personal life. Dear Annette, the time spend in your lab was really a lot of fun and educational. Dear Sven, I'm really grateful for the exciting time in Ethiopia, the trust you had in me, and the fun we had together (French wine is good, Ethiopian wine is Gooder!). Dear Conny, in your lab my passion for transcriptional regulation, enhancers and my quest for looping really started. I thank you very much for the great time in your lab and all the effort you did for me. I'm really pleased that you're an opponent in my committee.

Verder wil ik de mensen uit mijn persoonlijke omgeving bedanken. Vrienden, al de tijd dat we samen eindeloos klommen, dronken en discussieerden zijn heel erg belangrijk voor me geweest, allemaal heel erg bedankt en we moeten dat alles weer wat meer gaan doen. Jelle, jou moet ik even speciaal bedanken omdat je mijn paranimf wilt zijn. Onthoud dat het slechts een klein beetje moeilijker is dan ceremoniemeester, dus geen reden tot paniek.

Ma heel erg bedankt voor alle steun en interesse. Het feit dat je altijd achter me gestaan hebt, ook al moet ik je af en toe tot wanhoop gedreven hebben, betekent heel veel voor me. Beste Pa, ook al ben je er niet meer om het mee te maken ik weet dat je heel erg trots geweest zou zijn.

Tot slot, de meest belangrijke personen uit mijn leven: Brigitte en Lencho heel erg bedankt voor alle steun en liefde die jullie me geven. Ik ben maar een bofkont.