Journal of Cellular Biochemistry 102:801–810 (2007)

PROSPECTS

Control of Beta Globin Genes

Milind C. Mahajan, Subhradip Karmakar, and Sherman M. Weissman*

Department of Human Genetics, Yale University School of Medicine, New Haven, Connecticut

Abstract The developmental changes in expression of the beta like genes from embryonic to adult stages of human life are controlled at least partially at the level of the promoter sequences of these genes and their binding factors, and competition for promoter specific interactions with the locus control region (LCR). In recent years, the control of beta globin genes has also been investigated at the level of chromatin structure involving the chemical modification of histones and their remodelling by DNA dependent ATPases (SMARCA) containing protein complexes. The role of intergenic RNA is also being investigated with renewed interest. Although a wealth of information on the structure/function relationship of the LCR and globin promoters has been gathered over more than two decades, the fundamental nature of the control of these genes at the molecular level is still not completely understood. In the following pages, we intend to briefly describe the progress made in the field and discuss future directions. J. Cell. Biochem. 102: 801–810, 2007.

Key words: LCR; globin genes; intergenic RNA; epigenetics

Hemoglobin is a heme containing tetrameric protein comprising two alpha and two beta-like chains. The α- and β-globin gene loci contain several genes arranged in tandem whose expression is developmentally regulated. Major genes of the α cluster are a ζ gene expressed in the embryonic stage and $\alpha-1$ and $\alpha-2$ genes expressed throughout fetal and post-natal life. Both these genes share a strong enhancer (HS40) situated at −40 kb from the ζ gene. The β-globin locus contains one embryonic ε gene, two fetal Aγ and Gγ genes one pseudogene (ψ) and two adult genes called delta (δ) and beta (β). All these genes share the same enhancer HS2 situated in the locus control region (LCR) at 10 kb upstream of the ε-gene. Remarkably, despite the differences in the control regions, alpha like and beta like genes are expressed at equal levels throughout development. Mutations in the α and β genes and their regulatory sequences that alter their expression levels and/or protein structure lead to thalassemia, a condition that is one of the more frequently occurring genetically inherited disorders. The β-globin locus is an excellent model system to study the molecular mechanism of several aspects of control mechanisms of gene expression, as it contains a multiplicity of different regulatory sequences (Fig. 1).

DEVELOPMENTAL CONTROL OF GLOBIN GENES

Studies deleting portions of the LCR sequences, changing the gene order, and varying the distance between the promoter and LCR have led to the suggestion that there is both competition for the gene specific LCR-promoter interactions and autonomous gene silencing mechanisms for the globin gene switching [Stamatoyannopoulos, 2005]. Several in vivo and in vitro DNase footprinting studies revealed the presence of erythroid specific protein binding sites on the beta like globin promoters and the LCR [Reddy et al., 1994; Ikuta et al., 1996]. Over the years, the prominent erythroid sequence specific DNA binding proteins have been cloned, and efforts have been made towards the biochemical analysis of the DNA–protein complexes at a limited number of such protein binding sites of the promoter sequences and DNase hypersensitive sites (HS) of the LCR. Several partially or largely erythroid and sequence-specific DNA binding proteins have been studied including GATA 1, EKLF/KLF1 and NF-E2. These proteins interact with the beta

Grant sponsor: NIH; Grant number: 1PO1 HL63357; Grant sponsor: NHLBI; Grant number: N01-HV-28186.

*Correspondence to: Sherman Weissman, 333, Cedar Street, PO Box 208005, New Haven, CT 06510.
E-mail: sherman.weissman@yale.edu

Received 25 June 2007; Accepted 28 June 2007

DOI 10.1002/jcb.21507

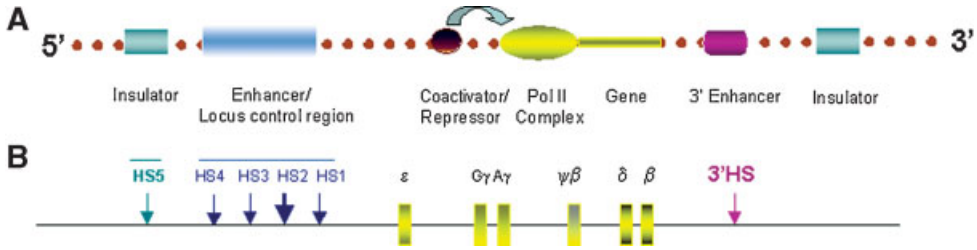

Fig. 1. Beta globin locus and its regulatory sequences serve as a poster boy for studying the regulation of gene expression. **A**: A cartoon describing a gene locus and its multiple regulatory sequences. **B**: Structure of the beta globin locus consisting of beta like genes arranged in tandem and its regulatory sequences.

globin promoters and the LCR in a cell-type specific manner. GATA-1 and EKLF are essential for the erythroid maturation as well as for globin production [Weiss et al., 1994; Bieker, 2005]. Although EKLF is present at all the stages of development, it has been proposed to be a preferential adult globin transcription activator [Bieker, 2005]. A helicase like transcription factor (HLTF) binds to a sequence at −140 bp to the β-globin promoter. Over expression of HLTF activates the β-globin transcription in K562 cells [Mahajan and Weissman, 2002]. On the other hand, even though NF-E2 binds to the core enhancer sequence of the HS2 and participates in the recruitment of Pol II to its cognate promoter [Johnson et al., 2001], the phenotype of the NF-E2 knockout mice suggests that it is not an essential protein either for erythropoiesis or for globin transcription [Andrews, 1998].

SSP and DRED are among the other DNA binding protein complexes that are implicated in the gene specific transcription activation and repression [Tanabe et al., 2002; Zhao et al., 2006]. The stage selector element at the γ promoter aids in the promotion of the LCR-γ promoter interaction in the fetal liver. Sp1 and a stage selector protein (SSP) interact with this specific sequences (the stage selector element SSE) in the γ-globin promoters. The SSP is a heterodimer consisting of a ubiquitous protein CP2 and an erythroid specific 22 kDa subunit called NF-E4 [Zhao et al., 2006]. This dimer preferentially activates gamma globin expression in the fetal liver. A 14 kDa isoform of NF-E4 that is made by alternate initiation of translation functions as a repressor of the γ-gene transcription.

Naturally occurring mutations at the beta like globin promoters have been studied to understand their transcriptional control. For example, some point mutations in the γ-globin promoter induce γ-globin production even in the adult stage leading to a clinical condition known as hereditary persistence of fetal hemoglobin (HPFH). Proteins have been identified that bind selectively to the wild type or HPFH mutant sequences, but their role in hemoglobin switching is uncertain. One such HPFH mutation occurs in a direct repeat element (DR1) that reduces the binding of the nuclear orphan receptors TR2 and TR4 containing repressor complex called DRED [Tanabe et al., 2002]. The TR2/TR4 heterodimer is a repressor for the ε- and γ-globin gene transcription. The HPFH mutation at their binding sequence may be responsible for the γ-globin production in the adult erythrocytes. On the other hand, a DNA methyltransferase 1 (DNMT1) containing protein complex binds to a DNA sequence with HPFH point mutation at −198 bp on the γ-gene promoter [Olave et al., 2007]. However, the significance of its binding to the onset of the HPFH phenotype is not clear.

Several of these transcription factors involved in the erythroid maturation and globin transcription exist in multi-protein complexes of varying size. Functionally, these protein complexes range form the chromatin remodeling factors to the transcriptional activators/repressors (Table I). EKLF recruits nucleosome remodeling SWI/SNF complex to the β-globin promoter [Armstrong et al., 1998]. The Ikaros transcription factor, in association with the NurD and SWI/SNF complexes, binds to an pyrimidine rich sequence upstream of the δ-globin gene [Bank et al., 2005]. Deletion of the Ikaros-complex binding sequence delays the switch from fetal to adult globin in transgenic mice. An LCR associated remodeling complex (LARC) that is more than 2 mDa in size is recruited to the HS2 region of the LCR [Mahajan et al., 2005].

TABLE I. Erythroid Transcription Factors and Associated Proteins

Transcription factor	Pulled down associated proteins	Method of pull down
GATA[a]-1	NuRd, ACF/WCRC, Fog-1, TAL-1, GFI-1b, Ldb, DNA repair proteins and topoisomerases	Biotinylted peptide tag
p18[a] Maf	SWI/SNF, NuRD, SAF-B, RAD 50, RAD 21, Bach1, p45 NF-E2, ZBP 89, Ldb-1, p75, hRHII, TIP 48, TIP 49, BRAF 35, HMG 2, and SmcB	Immuno precipitation
SSP	CP2, NF-E4	Yeast two hybrid
Complex	Associated proteins	Primary recruitment site
Biochemically purified multi-protein complexes		
E-RC 1	EKLF, Brg-1, and BAFs	β-Promoter
PYR	Ikaros, SWI/SNF, and NuRD complexes	Pyrimidine rich sequence upstream of δ-promoter
DRED	Nuclear orphan receptors, TR2, and TR4.	ε- and γ-globin promoters
LARC	hnRNP C1/C2, requiem, SWI/SNF, and MeCP 1	HS2
CTCF	Nudeophosmin, PARP, 40S ribosomal subunits, topoisomerase II, lamin A/C, histones H2A and H2Az, and Taf-1/set	Chicken HS4
198 HPFH	DNMT1, p52, SNEV, and RAP 74	γ-Globin promoter

[a]GATA-1 and p18 MAF exist is independent multiple complexes of varying protein compositions that activate and repress transcription.

As described above, a 450 kDa gamma promoter binding DRED complex and the stage selector protein (SSP) are among the smaller protein complexes recruited on the beta like promoters. Recently, a multi-protein complex containing DNMT1 and transcriptional co-activators p52, SNEV and RAP74 was purified [Olave et al., 2007]. This complex binds to the sequence with a HPFH point mutation at −198 bp in the γ-promoter. GATA-1 exists in several independent pre-formed protein complexes some of which are transcriptional activators while others are repressors [Grosveld et al., 2005]. It is not clear if any of these GATA-1 complexes binds to the beta globin locus. These complexes are suggested to be erythroid differentiation stage specific. Detailed analysis of the formation of these complexes and their association with the beta globin locus during primitive and definitive erythropoiesis is needed to fully delineate the role of GATA1 in globin gene regulation.

The insulator binding protein CTCF is associated with nucleolar nucleophosmin and PARP, 40S ribosomal subunits, topoisomerase II, lamin A/C, histones H2A and H2Az, and template activating factor Taf-1/set [Yusufzai et al., 2004]. The in vivo association of these CTCF binding proteins at the insulator sequence and their significance for the insulator activity is yet to be described.

Apart from existence of a transcription factor in several distinct protein complexes, exchange of transcription factors in a complex may occur during development and differentiation [Brand et al., 2004]. In the undifferentiated mouse erythroleukemia cells (MELs), an NF-E2p18/MafK-Bach1 heterodimer is associated with a large protein complex containing known repressors. Upon differentiation, Bach1 is replaced with p45 NF-E2, and other transcriptional activators that may be responsible for the activation of their cognate promoter [Brand et al., 2004]. Full biochemical analysis of these Bach1 and NF-E2 containing complexes in primitive and definitive erythropoiesis is needed. Several other large multi-protein complexes of unknown composition exist on the LCR and beta like gene promoter sequences (our unpublished observations). Biochemical analysis of these multi-protein complexes and their variants with respect to the protein composition during development and differentiation is required for further understanding of the control of the beta globin locus. Such analysis also

has the potential for the discovery of new functions of the LCR, mechanistic details of enhancer–promoter interaction and tissue and development specific promoter activation and inactivation.

CONTROL OF BETA GLOBIN LOCUS ACTIVITY BY LOCUS CONTROL REGION (LCR)

The locus control region of the beta globin genes takes part in several aspects of the control of beta globin locus such as enhancement of transcription, chromatin structure, insulation and positioning of the locus in the nucleus. The LCR spans 16 kb of DNA beginning 5 kb upstream of the ε-globin gene. It contains five DNase hypersensitive sites (HS) called HS1, HS2, HS3, HS4, and HS5 that contain phylogenetically conserved sequences [Stamatoyannopoulos, 2005]. Further upstream are the HS6 and HS7 that are not erythroid specific. Among these DNase hypersensitive sites, HS2 and HS3 have been investigated in some detail for their role in the enhancement of transcription and chromatin opening activity. The classical enhancer function of the LCR lies within the HS2 region and has tandem overlapping sites for the transcription factors AP1/NF-E2. An erythroid restricted p45 NF-E2 and ubiquitous p18 Maf are associated with this site. The LCR seems to enhance the transcription by recruiting NF-E2 to its cognate promoter [Johnson et al., 2001]. However, in the MELs, the deletion of the LCR does not alter the recruitment of the NF-E2 to the promoter, but it dramatically reduces the S-5 phosphorylation of RNA Pol II, a modification associated with early stages of transcription [Sawado et al., 2003]. These experiments suggest a role for the LCR in promoting the elongation of transcription by inducing the phosphorylation of the Pol II at the transcription initiation site. Phosphorylation at this site is accomplished by kinases that are part of the general transcription initiation factor TFIIH and the recruitment of this and other initiation factors to the globin promoter are yet to be investigated in this system.

Strong evidence for the dominant chromatin modifying activity of the LCR and/or the sequences beyond it comes from the occurrence of the naturally occurring Hispanic deletion of LCR and a large portion of DNA 5′ to LCR that keeps the rest of the chromatin in the closed configuration [Li et al., 2002]. Experimental evidence for the chromatin opening activity of the LCR comes from the suppression of position effect variegation in transgenic mice and biochemical purification of LARC that is recruited adjacent to the AP1/NFE2 site of the HS2 [Li et al., 2002; Mahajan et al., 2005]. LARC is a large combinatorial complex of SWI/SNF and MeCP1 complexes that are the chromatin remodeling complexes situated at the active gene and denser inactive chromatin, respectively. Interestingly, LARC contains hnRNP C1/C2 as an integral component and it seems to be responsible for the recruitment of the complex on HS2. A current model is that, on being recruited to the LCR, LARC spreads to the downstream beta like globin promoters. Remodeling complexes are also locally recruited on the promoters and adjoining sequences by transcription factors [Armstrong et al., 1998; Mahajan and Weissman, 2002; Bank et al., 2005]. It will be interesting to determine the sequence of the recruitment of these remodeling complexes on the β-globin locus and to establish which among these are the dominant remodellers. These studies are required to understand the mechanism of the tissue specific domain opening of the beta globin locus. So far, experiments with the LCR and LARC suggest it to play dominant role in the chromatin remodeling, followed by which the remodeling of the beta like globin promoters might occur. More studies are required to investigate this hypothesis.

Interestingly, the Hispanic LCR deletion mutation also exhibits altered origin of replication of the DNA at the beta globin locus [Aladjem, 2004]. In the normal situation, the origin of DNA replication is situated 5′ to the β-globin gene and the firing of the replication takes place in the early S phase. However, in the Hispanic deletions, the origin of the replication shifts to the 3′ end of the β-globin gene in the late S phase. Additional evidence for the role of LCR in the control of DNA replication comes from the analysis of ectopic sites with LCR-globin gene fusion constructs. The LCR sequences inserted in various random locations within murine chromosomes can confer tissue-specific replication timing of adjacent globin genes. However, deletion of the LCR did not affect either the timing of the replication or the site of origin of replication [Aladjem, 2004]. These conflicting data with LCR deletion

studies suggest the role for the sequences beyond the LCR as evidenced by the effects of the Hispanic deletion on DNA replication and chromatin opening activity.

In addition to the enhancement of the transcription and chromatin remodeling, the LCR also harbors insulator activity. Typically, an insulator functions as a blocker of the enhancer–promoter interaction and a check for the spread of the heterochromatin. Although the HS5 region of the mouse and human beta globin LCR is suggested to be an insulator [Tanimoto et al., 2003], much of the information on the insulator properties of the LCR comes from the chicken HS4 sequence [Gaszner and Felsenfeld, 2006]. A zinc finger containing CTCF transcription factor interacts with insulator sequence at the HS4 region of the chicken LCR [Gaszner and Felsenfeld, 2006].

The LCR also participates in proper positioning of the beta globin locus in the transcriptionally active foci in erythroid cells [Ragoczy et al., 2006]. In immature erythroid cells, the beta globin locus is situated in the transcriptionally unfavorable nuclear periphery. During erythroid differentiation and maturation, the beta globin locus moves away from the periphery to the transcriptionally active inner location of the nucleus. Deletion of the LCR significantly effects the movement of the locus in the mature erythroid cells.

DNase hypersensitive sites HS1 to HS5 span about 16 kb of the LCR. A significantly large portion of this LCR sequence is phylogenetically conserved [Hardison et al., 1997]. Out of these, a small number of sequences at HS2, HS3, and HS5 are implicated in the function of the LCR. A high throughput screen for the analysis of the rest of the large chunk of the conserved region of the 16 kb LCR is necessary to thoroughly comb for the details of the known functions. Although this task seems enormous, advances being made in the genomics and proteomics could make such an exercise a distinct possibility.

EPIGENETICS OF THE BETA GLOBIN LOCUS

Genomic DNA exists in complexes with several proteins that together constitute the chromatin. Within the chromatin, most of the DNA is wrapped around nucleosomes that consist of a histone octamer made up of four basic histones. In addition to the sequence of the genome, DNA methylation and post-translational modification of histone tails alter the chromatin structure either for the activation or repression of the gene. Thus, the DNA sequences associated with the chemically modified histones and methylated bases constitutes the epigenome.

Several types of chemical modifications occur on histone tails. These include acetylation, methylation, ribosylation, sumoylation, ubiquitinylation, and phosphorylation [Kouzarides, 2007]. The basic amino acids arginine and lysine are the targets for acetylation and/or methylation and serine residues are the substrates for phosphorylation. The type of modification and position of the modified amino acid on histone tail reflect and influence whether the gene is inactive, primed or transcribing. Generally, acetylation of the histones reflects an open chromatin state that harbors active genes. On the other hand, methylation of histones can be linked to either activation or repression of the gene transcription. Methylation of lysines at positions 4, 36, and 79 on histone H3 correlate with the activity of genes, where as methylation of H3 at lysine 9 and 27 is predominantly associated with transcriptionally repressed chromatin [Berger, 2007]. Acetylation of lysines at Histone H3 and H4 as well as dimethylation of the Lys 4 of Histone H3 is the signature of the transcriptionally favorable environment. Trimethylation of the Histone H3 Lys 4 is present on the actively transcribing gene promoters. Transcriptionally silent heterochromatin contains di and trimethylated histone H3 Lys 9 [Berger, 2007]. Methylation of Lysine 36 of histone H3 generally marks the actively transcribing coding region beyond the transcription initiation site [Kim et al., 2007a,b]. In recent years, considerable attention has been given to the relation between histone modifications and the expression of beta globin genes (Table II).

As the acetylation of histones is most common with the active genes, early studies on the epigenetics of beta globin locus started with the investigation of its acetylation status during development and differentiation of red cells. In mouse erythroid cells, acetylation of histone H3 is discontinuous. High levels of acetylation were observed in the LCR and in the active β-major and β-minor genes in adult globin expressing MEL cells. The chromatin at the embryonic Ey and bh1 remained hypoacetylated consistent with the notion of association of histone H3 acetylation with active genes [Bulger et al.,

TABLE II. Histone Modifications at the Beta Globin Locus

Species	Transcriptionally active domains	Transcriptionally inactive domains
Chicken	H4-Ac, H3-AcK9/K14, H3-Me3K4, and H3-Me2K4	H3-MeK9, H3-Ac and H4-Ac, and H3-Me2K4
Mouse	H3-AcK9, H4-Ac, H3-Me2K4 H3-Me3K4, H3-Me3K79, and H3-Me3K9	H3-Me2K4, hypoacetylated H3, and H4
Human	H3-AcK9, H4-Ac, H3-Me2K4 H3-Me3K4, H3-Me3K9, and H3-Me3-K36	H3-Me2K4, hypoacetylated H3, and H4

Ac, acetylated; Me, methylated; K, lysine; H3, histone H3; H4, histone H4.

2003]. This hypoacetylated state is proposed to be maintained by an HDAC-dependent surveillance mechanism [Im et al., 2002]. The human beta globin locus in the MELs also exhibits hyperacetylated histone H3 and H4 in the active LCR and adult β-globin genes. The rest of the locus harboring inactive ε- and γ-globin genes remains hypoacetylated [Schubeler et al., 2000]. In comparison to the hematopoietic progenitor cells, acetylation of Histone H4 of the human β-globin promoter increases in erythroid cells. Interestingly, acetylation of histone H3 decreases in the β-globin promoter in erythroid cells [Bottardi et al., 2003] implying that levels of acetylation is under tight control. On the other hand, chicken adult globin expressing erythroid cells have more or less uniformly hyeracetylated histones H3 and H4 throughout the beta globin locus [Litt et al., 2001a,b; Myers et al., 2003]. However, a 16 kb DNA region upstream of the HS4 boundary of the chicken beta globin locus remains unacetylated [Litt et al., 2001a,b].

Histone H3 methylation studies on the chicken globin locus are consistent with the common histone code for active and inactive genes [Litt et al., 2001a,b; Schneider et al., 2004]. In fact, these studies served as a model system among the metazoan genes that established the presence of the trimethylated Lys 4 of Histone H3 at the transcriptionally active promoters in embryonic as well as adult beta globin producing chicken erythroid cells, and the dimethylation of H3-Lys4 was distributed throughout the locus [Schneider et al., 2004]. Similarly, the methylation of Lys 79 on Histone H3 was significantly elevated on the active genes and was at undetectable or low levels in the inactive beta like genes in the adult as well as embryonic murine erythroid cells [Im et al., 2003]. The methylation and acetylation of Lys 9 were inversely correlated on the inactive and active genes of the chicken beta globin locus, respectively [Litt et al., 2001a,b]. Elevated levels of methylated histone H3 at Lys 9 were present on the developmentally inactive globin genes and on the constitutively condensed chromatin adjacent to the globin locus. A combination of histone H3 acetylation at Lys9 and methylation at Lys 4 occurs on the developmentally active beta globin genes [Litt et al., 2001a,b]. In human K562 cells too, correlation of di and trimethyl Lys 4 and acetylation on histone H3 on the active gamma and embryonic genes was observed [Kim et al., 2007a,b]. However, these modifications were accompanied by trimethylation of Lys 9 on histone H3 that is generally considered to be the signature of inactive closed chromatin. In murine erythroid cells also, the combination of trimethylated H3-Lys 4 and H3-Lys 9 residues at the active β-major gene were found [Vakoc et al., 2005]. The significance of this deviation from the general rule of occurrence of methylated H3-Lys 9 in the closed chromatin is yet to be worked out.

The mechanism of the chemically modified histone domain formation in a gene locus involves recruitment of appropriate modifying enzymes [Kouzarides, 2007]. The chromatin remodeling complexes recruited on the beta globin locus contain histone modifying enzymes [Bank et al., 2005; Mahajan et al., 2005]. Transcription factors are also reported to participate in the histone modification process on the beta globin locus. NF-E2 and GATA-1 influence the histone acetylation and methylation of the globin promoters, but the HS2 of the LCR shows these modifications in cells lacking NF-E2 or lacking GATA1 [Kiekhaefer et al., 2002; Im et al., 2003]. Interestingly, NF-E2 recruits CBP/p300 to the HS2 and acetylates the downstream promoter [Kim et al., 2007a,b]. Such site specific recruitment of the factors and their activity at several kilo/bases away requires either the spreading of the histone modification activity or the coming together of the two distant elements. LCR and its cognate

promoter are known to physically interact with each other [Brown, 2003]. However, in another study, deletion of LCR did not cause any change in the acetylation pattern of the globin promoters [Bulger et al., 2002]. The conflicting data obtained with the LCR deletion studies may be due to the recruitment of the LCR binding proteins on the alternate sites on the rest of the locus or to differences in the systems being tested. In addition, it seems probable that not all the functions of the LCR are being tested in current systems.

In summary, the chicken beta globin locus seems to follows the established norm of the histone modifications at active and inactive promoters at the different stages of development [Litt et al., 2001a,b; Schneider et al., 2004]. However, the murine and human beta globin loci exhibit complexities in the histone modification pattern [Vakoc et al., 2005; Kim et al., 2007a,b]. So far, epigenetics of beta globin locus has been investigated with relatively few histone modification types and patterns (Table II). A unique combination of histone modifications may be present on the beta globin regulatory domains. We speculate that these may be significant for the transcriptional switching of these genes during development. The challenge lies in identifying the specific modification mix among the possible combinations of modifications and their sites on all the four major histones. Rapid advances being made in the genomic, proteomic and bioinformatics tools should bring such a task within reach.

INTERGENIC TRANSCRIPTION

In erythroid cells, along with the transcription of the appropriate beta like gene transcription, the sequences between the beta like globin genes and the LCR are also actively transcribed in the sense orientation in the direction of the cis-linked promoter [Tuan et al., 1992; Ashe et al., 1997; Xiang et al., 2006]. The extragenic transcription was initially described to originate from the HS2 region of the LCR [Tuan et al., 1992]. This RNA is nuclear and polyadenylated, but does not code for any protein [Ling et al., 2005]. These observations correlate with the recruitment of the RNA Pol II to HS2 [Johnson et al., 2003]. Subsequently, transcription was shown to originate from an endogenous retroviral long terminal repeat (ERV9) upstream of HS5 of the LCR [Plant et al., 2001]. The farthermost region that is transcribed into the embryonic globin gene in the human K562 cell is situated at -236 kb from the ε-globin gene [Xiang et al., 2006]. The genic and intergenic transcriptions seem to be different [Kim et al., 2007a,b]. The HS2 dependent intergenic transcription is also associated with the acetylation and Lys 4 methylation of the intervening sequence between the HS2 and its linked gene. However, the genic transcription does not participate in the histone modification [Kim et al., 2007a,b]. The intergenic RNA is processed by the dicer dependent turnover mechanism, unlike the typical mRNA processing of the genic hnRNA [Haussecker and Proudfoot, 2005], but upstream globin transcripts may also show splicing patterns based on the use of conventional splice sites [Xiang et al., 2006].

An obvious interpretation of the occurrence of this intergenic and extragenic transcription has been to look if it has any role to play in the control of globin genes and switching of the globin gene transcription during development. The significance of this extragenic and intergenic transcription has led to the proposal of several models for the LCR promoter interactions and structure of the chromatin. By initiating the transcription, the HS2 was suggested to track along the DNA to deliver the enhancer transcription factors to its cognate promoter. The elongation of the HS2 originated RNA was drastically inhibited when a transcriptional terminator was introduced between the HS2 and the downstream gene. It also inhibited the enhancement of the transcription of the linked ε-globin promoter [Ling et al., 2004]. Extragenic transcription occurs at all levels of development. However, the intergenic transcription between the globin genes is developmental stage specific and is correlated with the DNase hypersensitivity of the transcribed region [Gribnau et al., 2000]. This observation led to the proposal that the intergenic transcription establishes transcriptionally active and inactive domains in the beta globin locus. Interestingly, a naturally occurring Corfu deletion between the γ- and δ-globin genes activates the transcription of the γ-globin gene [Chakalova et al., 2005]. This particular Corfu deletion harbors site for termination of fetal stage intergenic RNA. More studies are needed to establish if the intergenic

RNA between γ and δ genes is involved in the silencing of the γ gene.

UNSOLVED ISSUES AND FUTURE CHALLENGES

There are obvious things we would like to know about the globin cluster that can at least be approached with current technology. For example, one would like to find the chromatin modifications that are unique for the globin locus sub-domains in normal embryonic, fetal and adult red cells, and obtain results of high throughput screening for the proteins binding to the entire locus during various stages of development.

One unsolved issue of general significance is the role of extragenic transcription in expression of globin genes. Whether the extragenic transcription is indeed needed for the regulation of transcription and formation of active/inactive chromatin domains or has some other role is still an open question. The globin cluster offers a challenging model for this study, as the intergenic transcription is controlled in a stage specific manner, and has been at least partially characterized at the level of mature transcripts.

Another puzzle is why the locus control region of the beta globin cluster is so complex, with multiple evolutionarily conserved regions, and so apparently different from the alpha globin control region. It may be relevant that the alpha globin cluster is located adjacent to widely transcribed genes while the human beta globin cluster is situated in the middle of a cluster of olfactory receptor genes that are presumably not expressed in most cell types, including erythrocytes.

A third issue is what mechanisms are responsible for the large differences in the relative accumulation of globin transcripts and progressive decline of most of the transcriptome at various stages of erythroid precursor development. Some of this may be accomplished by control of mRNA stability but the issue remains as to whether the same factors that control transcription in early erythroblasts operate in the later stages of erythroid development. The elucidation of the molecular mechanism that increases the globin transcript with concomitant decline in the rest of the transcripts in a differentiating erythrocyte will be fundamental to understanding the regulation of transcription during differentiation.

A related question is whether the movement of beta globin genes from heterochromatic regions into more open nucleoplasm during erythroid development is solely a passive consequence of transcription initiating within the cluster, or whether there are additional changes in the surrounding chromatin that either mediate or reflect this movement. Similarly, is the change from late replication of globin loci to earlier replication as erythroid cells develop only a passive consequence of increased activity of the globin promoters, or is it secondary to the translocation of the globin genes or more specific changes in the globin locus? There could also be special sequences or mechanisms that cause the change in relative timing of the replication of this locus.

Some of the more elegant studies of globin gene regulation are those indicating a switch back and forth from fetal to adult globin production at an individual locus. Presumptively this switch reflects changes in chromatin configuration but exactly what these are at the molecular scale, and what the factors are that determine the rate of switching remain to be explored.

Finally, in spite of the obvious nature and long term interest in the problem, we have to conclude that the mechanisms controlling the switching from embryonic to adult globin during organismal maturation are at best only partly understood.

ACKNOWLEDGMENTS

The authors apologize for curtailing many primary references due to the limitation of space. Instead, reviews have been quoted that cover the details of the said beta globin locus activity. This project has been funded in whole or in part with Federal funds from the National Heart, Lung, and Blood Institute, National Institutes of Health, under contract No. N01-HV-28186 and NIH Grant Number 1P01 HL63357.

REFERENCES

Aladjem MI. 2004. The mammalian beta globin origin of DNA replication. Front Biosci 9:2540–2547.

Andrews NC. 1998. The NF-E2 transcription factor. Int J Biochem Cell Biol 30:429–432.

Armstrong JA, Bieker JJ, Emerson BM. 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 HL, Monks J, Wijgerde M, Fraser P, Proudfoot NJ. 1997. Intergenic transcription and transinduction of the human beta-globin locus. Genes Dev 11:2494–2509.

Bank A, O'Neill D, Lopez R, Pulte D, Ward M, Mantha S, Richardson C. 2005. Role of intergenic human gamma-delta-globin sequences in human hemoglobin switching and reactivation of fetal hemoglobin in adult erythroid cells. Ann NY Acad Sci 1054:48–54.

Berger SL. 2007. The complex language of chromatin regulation during transcription. Nature 447:407–412.

Bieker JJ. 2005. Probing the onset and regulation of erythroid cell-specific gene expression. Mt Sinai J Med 72:333–338.

Bottardi S, Aumont A, Grosveld F, 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.

Brand M, Ranish JA, Kummer NT, Hamilton J, Igarashi K, Francastel C, Chi TH, Crabtree GR, Aebersold R, Groudine M. 2004. Dynamic changes in transcription factor complexes during erythroid differentiation revealed by quantitative proteomics. Nat Struct Mol Biol 11:73–80.

Brown KE. 2003. Chromatin folding and gene expression: New tools to reveal the spatial organization of genes. Chromosome Res 11:423–433.

Bulger M, Sawado T, Schubeler D, 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 MA, Hamilton J, Farrell CM, Hardison RC, 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.

Chakalova L, Osborne CS, Dai YF, Goyenechea B, Metaxotou-Mavromati A, Kattamis A, Kattamis C, Fraser P. 2005. The Corfu deltabeta thalassemia deletion disrupts gamma-globin gene silencing and reveals post-transcriptional regulation of HbF expression. Blood 105:2154–2160.

Gaszner M, Felsenfeld G. 2006. Insulators: Exploiting transcriptional and epigenetic mechanisms. Nat Rev Genet 7:703–713.

Gribnau J, Diderich K, Pruzina S, Calzolari R, Fraser P. 2000. Intergenic transcription and developmental remodeling of chromatin subdomains in the human beta-globin locus. Mol Cell 5:377–386.

Grosveld F, Rodriguez P, Meier N, Krpic S, Pourfarzad F, Papadopoulos P, Kolodziej K, Patrinos GP, Hostert A, Strouboulis J. 2005. Isolation and characterization of hematopoietic transcription factor complexes by in vivo biotinylation tagging and mass spectrometry. Ann NY Acad Sci 1054:55–67.

Hardison R, Slightom JL, Gumucio DL, Goodman M, Stojanovic N, 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.

Haussecker D, Proudfoot NJ. 2005. Dicer-dependent turnover of intergenic transcripts from the human beta-globin gene cluster. Mol Cell Biol 25:9724–9733.

Ikuta T, Papayannopoulou T, Stamatoyannopoulos G, Kan YW. 1996. Globin gene switching. In vivo protein-DNA interactions of the human beta-globin locus in erythroid cells expressing the fetal or the adult globin gene program. J Biol Chem 271:14082–14091.

Im H, Grass JA, Christensen HM, Perkins A, Bresnick EH. 2002. Histone deacetylase-dependent establishment and maintenance of broad low-level histone acetylation within a tissue-specific chromatin domain. Biochemistry 41:15152–15160.

Im H, Park C, Feng Q, Johnson KD, Kiekhaefer CM, Choi K, Zhang Y, Bresnick EH. 2003. Dynamic regulation of histone H3 methylated at lysine 79 within a tissue-specific chromatin domain. J Biol Chem 278:18346–18352.

Johnson KD, Christensen HM, Zhao B, Bresnick EH. 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 KD, Grass JA, Park C, Im H, Choi K, Bresnick EH. 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.

Kiekhaefer CM, Grass JA, Johnson KD, Boyer ME, Bresnick EH. 2002. Hematopoietic-specific activators establish an overlapping pattern of histone acetylation and methylation within a mammalian chromatin domain. Proc Natl Acad Sci USA 99:14309–14314.

Kim A, Kiefer CM, Dean A. 2007. Distinctive signatures of histone methylation in transcribed coding and noncoding human beta-globin sequences. Mol Cell Biol 27:1271–1279.

Kim A, Zhao H, Ifrim I, Dean A. 2007. {beta}-globin intergenic transcription and histone acetylation dependent on an enhancer. Mol Cell Biol

Kouzarides T. 2007. Chromatin modifications and their function. Cell 128:693–705.

Li Q, Peterson KR, Fang X, Stamatoyannopoulos G. 2002. Locus control regions. Blood 100:3077–3086.

Ling J, Ainol L, Zhang L, Yu X, Pi W, Tuan D. 2004. HS2 enhancer function is blocked by a transcriptional terminator inserted between the enhancer and the promoter. J Biol Chem

Ling J, Baibakov B, Pi W, Emerson BM, Tuan D. 2005. The HS2 enhancer of the beta-globin locus control region initiates synthesis of non-coding, polyadenylated RNAs independent of a cis-linked globin promoter. J Mol Biol 350:883–896.

Litt MD, Simpson M, Gaszner M, Allis CD, Felsenfeld G. 2001. Correlation between histone lysine methylation and developmental changes at the chicken beta-globin locus. Science 293:2453–2455.

Litt MD, Simpson M, Recillas-Targa F, Prioleau MN, Felsenfeld G. 2001. Transitions in histone acetylation reveal boundaries of three separately regulated neighboring loci. EMBO J 20:2224–2235.

Mahajan MC, Weissman SM. 2002. DNA-dependent adenosine triphosphatase (helicaselike transcription factor) activates beta-globin transcription in K562 cells. Blood 99:348–356.

Mahajan MC, Narlikar GJ, Boyapaty G, Kingston RE, Weissman SM. 2005. Heterogeneous nuclear ribonucleoprotein C1/C2, MeCP1, and SWI/SNF form a chromatin

remodeling complex at the beta-globin locus control region. Proc Natl Acad Sci USA 102:15012–15017.

Myers FA, Chong W, Evans DR, Thorne AW, Crane-Robinson C. 2003. Acetylation of histone H2B mirrors that of H4 and H3 at the chicken beta-globin locus but not at housekeeping genes. J Biol Chem 278:36315–36322.

Olave IA, Doneanu C, Fang X, Stamatoyannopoulos G, Li Q. 2007. Purification and identification of proteins that bind to the hereditary persistence of fetal hemoglobin-198 mutation in the gamma-globin gene promoter. J Biol Chem 282:853–862.

Plant KE, Routledge SJ, Proudfoot NJ. 2001. Intergenic transcription in the human beta-globin gene cluster. Mol Cell Biol 21:6507–6514.

Ragoczy T, Bender MA, Telling A, Byron R, Groudine M. 2006. The locus control region is required for association of the murine beta-globin locus with engaged transcription factories during erythroid maturation. Genes Dev 20:1447–1457.

Reddy PM, Stamatoyannopoulos G, Papayannopoulou T, Shen CK. 1994. Genomic footprinting and sequencing of human beta-globin locus. Tissue specificity and cell line artifact. J Biol Chem 269:8287–8295.

Sawado T, Halow J, Bender MA, 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.

Schneider R, Bannister AJ, Myers FA, Thorne AW, Crane-Robinson C, Kouzarides T. 2004. Histone H3 lysine 4 methylation patterns in higher eukaryotic genes. Nat Cell Biol 6:73–77.

Schubeler D, Francastel C, Cimbora DM, Reik A, Martin DI, 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.

Stamatoyannopoulos G. 2005. Control of globin gene expression during development and erythroid differentiation. Exp Hematol 33:259–271.

Tanabe O, Katsuoka F, Campbell AD, Song W, Yamamoto M, Tanimoto K, Engel JD. 2002. An embryonic/fetal beta-type globin gene repressor contains a nuclear receptor TR2/TR4 heterodimer. EMBO J 21:3434–3442.

Tanimoto K, Sugiura A, Omori A, Felsenfeld G, Engel JD, Fukamizu A. 2003. Human beta-globin locus control region HS5 contains CTCF- and developmental stage-dependent enhancer-blocking activity in erythroid cells. Mol Cell Biol 23:8946–8952.

Tuan D, Kong S, Hu K. 1992. Transcription of the hypersensitive site HS2 enhancer in erythroid cells. Proc Natl Acad Sci USA 89:11219–11223.

Vakoc CR, Mandat SA, Olenchock BA, Blobel GA. 2005. Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. Mol Cell 19:381–391.

Weiss MJ, Keller G, Orkin SH. 1994. Novel insights into erythroid development revealed through in vitro differentiation of GATA-1 embryonic stem cells. Genes Dev 8: 1184–1197.

Xiang P, Fang X, Yin W, Barkess G, Li Q. 2006. Non-coding transcripts far upstream of the epsilon-globin gene are distinctly expressed in human primary tissues and erythroleukemia cell lines. Biochem Biophys Res Commun 344:623–630.

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

Zhao Q, Zhou W, Rank G, Sutton R, Wang X, Cumming H, Cerruti L, Cunningham JM, Jane SM. 2006. Repression of human gamma-globin gene expression by a short isoform of the NF-E4 protein is associated with loss of NF-E2 and RNA polymerase II recruitment to the promoter. Blood 107:2138–2145.